

ИНСТИТУТ ЗА ФИЗИКУ

ПРИМЉЕНО:		29. 01. 2025	
Рад.јед.	б р о ј	Арх.шифра	Прилог
4801	111/1		

Научном већу
Института за физику у Београду


Предмет: Молба за покретање поступка за избор у звање истраживач сарадник

Молим Научно веће Института за физику у Београду да покрене поступак за мој избор у звање истраживач- сарадник.

У прилогу достављам:

1. Мишљење руководиоца лабораторије са предлогом чланова комисије за избор у звање,
2. стручну биографију,
3. преглед научне активности,
4. списак објављених научних радова,
5. потврду о уписаним докторским студијама,
6. дипломе са просеком оцена на основним и мастер студијама,
7. потврду о пријављеној теми докторске дисертације,
8. копије објављених научних радова.

С поштовањем,



Филип Арнаут, истраживач- приправник
Живка Давидовића 83/4, Звездара, Београд
Тел: +381 63 83 89 237
E-mail: filip.arnaut@ipb.ac.rs

ИНСТИТУТ ЗА ФИЗИКУ

ПРИМЉЕНО:		29. 01. 2025	
Рад.јед.	б р о ј	Арх.шифра	Прилог
0801	111/2		

Научном већу Института за физику у Београду

Предмет: Мишљење руководиоца лабораторије о избору Филипа Арнаута у звање истраживач- сарадник и предлог комисије

Филип Арнаут је запослен у Лабораторији за астрофизику и физику јоносфере, Института за физику у Београду, Универзитета у Београду од маја 2023. године. Докторска дисертација са називом „Примена модела вођених подацима у геофизици“ прихваћена је од стране Наставно- научног већа Рударско-геолошког факултета, Универзитета у Београду под менторством од стране др Весне Цветков, ред. проф. У досадашњем ангажовању у Институту за физику у Београду као и током својих докторских студија, Филип Арнаут објавио је више научних радова на SCI листи и презентовао је резултате својих истраживања на међународним скуповима и конгресима. Са обзиром да испуњава све предвиђене услове Правилником о стицању истраживачких и научних звања, сагласан сам са покретањем поступка и предлажем избор Филипа Арнаута у звање истраживач- сарадник.

За избор Филипа Арнаута у звање истраживач- сарадник предлажем комисију у следећем саставу:

- 1) др Александра Коларски, научни сарадник, Институт за физику у Београду, Универзитет у Београду,
- 2) др Александра Нина, виши научни сарадник, Институт за физику у Београду, Универзитет у Београду,
- 3) др Весна Цветков, редовни професор, Рударско-геолошки факултет, Универзитет у Београду.

Др Владимир Срећковић
Научни саветник
Руководилац Лабораторије за астрофизику и физику јоносфере
Институт за физику у Београду
Универзитет у Београду



У Београду
24. 1 2025. год.

БИОГРАФИЈА- Филип Арнаут

Филип Арнаут рођен је 14. јула 1997. године у Београду, где је завршио основну и средњу електротехничку школу „Раде Кончар“ смер електротехничар електронике. Основне академске студије на Рударско-геолошком факултету, Универзитета у Београду, одсеку за Геологију, студијски програм геофизика уписује 2016. године као студент финансиран од стране буџета Републике Србије. Основне академске студије завршава 2020. године са просеком од 8,29/10,00. Дипломски рад „Примена укритеног квадратног диспозитива за детекцију примарне зоне раста корена дрвета врбе“ одбранио је са оценом 10,00 на Катедри за геофизику тиме је стекао назив дипломирани инжењер геологије. У септембру 2020. године уписао је мастер академске студије на Рударско-геолошком факултету, Универзитета у Београду, на студијском програму геофизике као студент финансиран од стране буџета Републике Србије. Мастер академске студије завршава у јулу 2021. године са просеком од 9,36/10,00. Мастер рад „Корелабилност соларног ветра са сеизмичким догађајима у зони Балканског полуострва“ одбранио је са оценом 10,00 на Катедри за геофизику тиме је стекао назив мастер инжењер геологије. У октобру 2021. године уписао је докторске академске студије на Рударско-геолошком факултету Универзитета у Београду на студијском програму геологија, подобласт геофизика. Области истраживања у склопу докторске дисертације обухватају примену класификационих и регресионих метода машинског учења на просторне и временске гео- серије. У склопу наведеног истраживачког фокуса, просторне и временске астро- серије такође се убрајају, где се истражују подаци везани за Земљину ниску јоносферу. Тренутни просек оцена на докторским студијама износи 10,00/10,00.

Од 2018. године Филип Арнаут је активан члан Српског геолошког друштва, од ког добија прву награду на Конкурсу за најбољи рад младих геолога и студената у 2021. години за рад „Примена укритеног квадратног диспозитива за детекцију примарне зоне раста корена дрвета врбе“. За школску годину 2022/2023, добио је стипендију (енгл. *Norman and Shirley Domenico scholarship*) од стране Друштва истраживачких геофизичара (енгл. *Society of Exploration Geophysicists- SEG*). Награду за најбољи постер добио је на V Meeting on Astrophysical Spectroscopy A&M Data- Astronomy & Earth Observations (Палић, Република Србија, 12- 15. Септембар, 2023. године). Током своје научноистраживачке каријере Филип Арнаут је био члан организационог одбора на више међународних скупова.

Од 2019. године Филип Арнаут се бави научноистраживачким радом. У том периоду је објавио четрдесетичетири (44) научноистраживачка рада у домаћим и страним часописима као и презентација на научним конференцијама, од чега 11 на SCI листи. Истраживачко звање истраживач- приправник добио је у септембру 2022. године на Департману за геофизику, Рударско-геолошког факултета, Универзитета у Београду.

Филип Арнаут је учесник пројекта *Карактеризација и технолошки поступци за рециклажу и поновну употребу флотацijske јаловине рудника „Рудник“ (REASONING)* који је финансиран од стране Фонда за науку Републике Србије (PRIZMA). Филип Арнаут је такође члан два билатерална научна пројекта: *Space weather and monitoring the effects of extraterrestrial radiation* (билатерална сарадња у науци и технологији између Р. Србије и Аустрије за период 2024-2026) и *Detection of astrophysical and geophysical phenomena from VLF radio measurements using machine learning methods* (билатерална сарадња у науци и технологији између Р. Србије и Словачке за период 2024-2025).

Преглед научне активности Филипа Арнаута

Научна активност кандидата Филипа Арнаута заснива се на примени модела вођених подацима у геофизици. Досадашњи научноистраживачки рад се може класификовати као:

1) примена класификационих метода машинског учења за класификацију некавалитетног врло ниско фреквентног (VLF) јоносферског сигнала:

Радио таласи врло ниских фреквенција тзв. VLF сигнали су представљени подацима високе резолуције (један минут), при чему се приликом истраживања анализирају подаци са више парова предајник-пријемник. У зависности од временског ограничења жељеног периода истраживања количина података коју истраживач мора обрадити може бити изразито велика. Подаци VLF сигнала поред свих сметњи које су својствени за већину сигнала мерених у природи додатно су оптерећени и утицајем ефеката соларних флорова. Тестираће се методе класификације машинског учења на податке VLF сигнала са циљем аутоматизације отклањања непожељног дела сигнала, и

2) примена регресионих метода машинског учења за одређивање јоносферских параметара оштрине и висине границе Д слоја јоносфере:

Најнижи регион јоносфере, Д регион, је одређен са два јоносферска параметра, оштрином и висином границе. Током нормалних услова граница Д регион има релативно константну вредност оба параметра, док се током соларних флорова параметри Д региона мењају. Нумеричким моделовањем могу се одредити ови параметри за поремећени период, међутим то је веома изазован задатак али због комплексности самог процеса. Због тога ће се применити регресионе методе машинског учења на податке Д региона јоносфере са циљем аутоматизације добијања параметара Д региона током поремећених услова изазваних утицајем соларних флорова X- класе.

3) примена метода прогнозирања временских серија на податке концентрације загађујућих материја у ваздуху у Београду:

Прогнозирање временских серија концентрације загађујућих материја у ваздуху су значајна са аспекта добијања будућих вредности концентрације загађујућих материја у ваздуху на датој локацији. Таква информација је веома битна за особе које болују од болести дисајног тракта, као и за најмлађу и најстарију популацију које су највише угрожене високим загађењем ваздуха. За сада не постоји утемељена, стабилна и у потпуности тачна метода прогнозирања вредности квалитета ваздуха.

4) примена класификационих метода машинског учења за прогнозирање просторног положаја офиолита на простору источне Вардарске зоне:

Картирање литолошких јединица често захтева обиман теренски рад истраживача различитог профила, дуготрајну обраду података са терена и финалну израду карти на основу добијених података. Методе машинског учења омогућавају да се модел машинског учења „научи“ о просторном распореду одређених литолошких јединица на једном делу картираног подручја на основу доступних података, након чега се примењује усвојена логика на онај део подручја на коме се не зна просторни распоред дате литолошке јединице.

5) примена метода статистичких метода на податке магнетне суцептибилности прикупљене из три бушотине на јаловишту рудника „Рудник“:

Мерења магнетне суцептибилности високе резолуције уз праћење физичко-хемијских услова и дистрибуције минералних асоцијација, тешких метала и пратећих елемената по вертикалном профилу (са дубином) тела флотацијског јаловишта рудника „Рудник“ обезбеђује велики скуп података на који се примењују статистичке методе и методе науке о подацима.

Списак публикација

Категорија M22:

1. **Arnaut, F.**, Kolarski, A., Srećković, V.A., Mijić, Z., 2023. Ionospheric Response on Solar Flares through Machine Learning Modeling. *Universe*, 9, 474. <https://doi.org/10.3390/universe9110474>
2. **Arnaut, F.**, Kolarski, A., Srećković, V.A., 2023. Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. *Universe*, 9, 436. <https://doi.org/10.3390/universe9100436>
3. **Arnaut, F.**, Đurić, D., Đurić, U., Samardžić-Petrovic, M., Peshevski, I. 2024, Application of geophysical and multispectral imagery data for predictive mapping of a complex geo-tectonic unit: a case study of the East Vardar Ophiolite Zone, North-Macedonia. *Earth Science Informatics*. <https://doi.org/10.1007/s12145-024-01243-4>
4. **Arnaut, F.**, Đurđević, V., Kolarski, A., Srećković, V.A., Jevremović, S., 2024. Improving Air Quality Data Reliability through Bi-Directional Univariate Imputation with the Random Forest Algorithm. *Sustainability*, 16 (17). <https://doi.org/10.3390/su16177629>

Категорија M23:

1. **Arnaut, F.**, Kolarski, A., Srećković, V.A., 2024. Machine Learning Classification Workflow and Datasets for Ionospheric VLF Data Exclusion. *Data*, 9, 17. <https://doi.org/10.3390/data9010017>
2. **Arnaut, F.**, Cvetkov, V., Đurić, D., Samardžić-Petrović, M., 2023. Short-term forecasting of PM10 and PM2.5 concentrations with Facebook's Prophet Model at the Belgrade-Zeleno brdo. *Geofizika*, 40, 162–177. <https://doi.org/10.15233/gfz.2023.40.7>
3. Kolarski, A., Srećković, V.A., **Arnaut, F.**, 2023. Low intensity solar flares' impact: numerical modeling. *Contributions of the Astronomical Observatory Skalnaté Pleso*, 53. <https://doi.org/10.31577/caosp.2023.53.3.176>
4. Kolarski, A., Srećković, V.A., Langović, M., **Arnaut, F.**, 2023. Energetic solar flare events in relation with subionospheric impact on 6-10 September 2017: data and modeling. *Contributions of the Astronomical Observatory Skalnaté Pleso*, 53. <https://doi.org/10.31577/caosp.2023.53.3.138>
5. **Arnaut, F.**, Sretenović, B., Cvetkov, V., 2022. Improvement of 1D geoelectric sounding by narrowing the equivalence range and identification, quantification and reduction of lateral effects using the tri-potential technique. *Geofizika*, 39, 297–320. <https://doi.org/10.15233/gfz.2022.39.15>
6. Jevremović, S., Kachadoorian, C., **Arnaut, F.**, Kolarski, A. and Srećković, V.A., 2024. Sustainable Transportation Characteristics Diary—Example of Older (50+) Cyclists. *Data*, 9(11), p.123.

Категорија M24:

1. **Arnaut, F.**, Vuckovic, D., Vasiljevic, I., Cvetkov, V., 2021. Correlability of solar wind with seismic events in the Balkan peninsula zone. *Annales geologiques de la Peninsule balkanique*, 82, 69–83. <https://doi.org/10.2298/gabp211029008a>

Категорија M31:

1. **Arnaut, F.**, 2024. *Analysis of possible relationship between earthquakes and Solar flare events during Solar cycle 24*. In: *Building Bridges Between Climate Science and Society Through a Transdisciplinary Network*, 10-14 September 2024, Kopaonik Mt, Serbia.

2. **Arnaut, F.** (2024). Daytime and nighttime VLF signal classification utilizing machine learning methods. *Meeting on new trends in Astronomy & Earth Observation. Book of Abstracts and Contributed Papers.*

Kategorija M33:

1. **Arnaut F.**, Kolarski, A., 2023. Multilayer Perception Hyperparameter Fine-Tuning for Ionospheric VLF Amplitude Data Exclusion. V Meeting on Astrophysical Spectroscopy - A&M DATA – Astronomy & Earth Observations, 12-15 September, Palić, Republic of Serbia.
2. **Arnaut, F.**, Kolarski, A., 2023. Feature Importance Analysis in Random Forest Regression for Air Quality Forecasting in Belgrade, Serbia. International Conference on Recent Trends in Geoscience Research and Applications 2023, 23-27 October, 2023, Belgrade, Republic of Serbia and virtual.
3. **Arnaut, F.**, 2024. Solar Wind and Seismic Activity in the Balkan Peninsula: a 2019-2023 Progress Report. In: Building Bridges Between Climate Science and Society Through a Transdisciplinary Network, 10-14 September 2024, Kopaonik Mt, Serbia.
4. Cvetkov, V., Arnaut, F., & Životić, D. (2024). Comparative analysis of correlation coefficients in mineralogical and geophysical data from the mine tailing site “Rudnik,” Republic of Serbia. *V Congress of Macedonian Geologists.*
5. Cvetkov, V., Simić, V., Petrović, S., Arnaut, F., Kostović, M., Radulović, D., Stojanović, J., Jovanović, V., Todorović, D., Nikolić, N., Senćanski, J., Bogdanović, G., & Marilović, D. (2024). Project reasoning: Characterization and technological procedures for recycling and reusing of the Rudnik mine flotation tailings. *V Congress of Macedonian Geologists.*

Kategorija M34:

1. **Arnaut F.**, Kolarski, A., 2023. Alternative Evaluation Metrics for Machine Learning Model Selection in Ionospheric VLF Amplitude Data Exclusion. V Meeting on Astrophysical Spectroscopy - A&M DATA – Astronomy & Earth Observations, 12-15 September, Palić, Republic of Serbia.
2. **Arnaut, F.**, Kolarski, A., 2023. Machine Learning Approach for Distinguishing Daytime and Nighttime Ionospheric Conditions on VLF Signals Related to Solar Flares During 2011. XX Serbian Astronomical Conference, 16-20 October, 2023, Belgrade, Republic of Serbia.
3. **Arnaut, F.**, Kolarski, A. and Srećković, V., 2024. *Comparative Analysis of Random Forest and XGBoost in Classifying Ionospheric Signal Disturbances During Solar Flares* (No. EGU24-2046). Copernicus Meetings.
4. **Arnaut, F.**, Kolarski, A., and Srećković, V. A. (2024). X-ray-induced changes in near-Earth plasma: A machine learning perspective. *VI Conference on Active Galactic Nuclei and Gravitational Lensing*, June 02-06, 2024, Zlatibor Mt., Serbia. Available at: <https://doi.org/10.69646/aob24004>
5. Kolarski, A., Nina, A., Srećković, V. A., and **Arnaut, F.** (2024). Monitoring solar flares and gamma ray bursts: Multi-instrumental approach investigation. *VI Conference on Active Galactic Nuclei and Gravitational Lensing*, June 02-06, 2024, Zlatibor Mt., Serbia. Available at: <https://doi.org/10.69646/aob24004>
6. Srećković, V. A., Kolarski, A., **Arnaut, F.**, Dimitrijević, M. S., Christova, M. D., and Bezuglov, N. N. (2024). New molecular data for astrochemical modelling. *VI Conference on Active Galactic Nuclei and Gravitational Lensing*, June 02-06, 2024, Zlatibor Mt., Serbia. Available at: <https://doi.org/10.69646/aob24004>

7. Srećković, V.A., Kolarski, A., Langović, M., **Arnaut, F.**, Mijić, Z., Jevremović, S., Barović, J. and Kounchev, O., 2024. Novel Research in Astrophysics and Geophysics. In: Building Bridges Between Climate Science and Society Through a Transdisciplinary Network, 10-14 September 2024, Kopaonik Mt, Serbia.
8. Srećković, V.A., Kolarski, A., Langović, M., **Arnaut, F.** and Jevremović, S., 2024. *Dataset for low ionosphere modeling*. In: Building Bridges Between Climate Science and Society Through a Transdisciplinary Network, 10-14 September 2024, Kopaonik Mt, Serbia.
9. Kolarski, A., **Arnaut, F.**, Jevremović, S., Langović, M. and Komatina, S., 2024. *Modern society and climate change: oil and space industries perspectives*. In: Building Bridges Between Climate Science and Society Through a Transdisciplinary Network, 10-14 September 2024, Kopaonik Mt, Serbia.
10. Srećković, V. A., Kolarski, A., & **Arnaut, F.**, 2024. Exploring Solar Radiation: Influence, Diagnostics, Prediction. In: XIV Serbian-Bulgarian Astronomical Conference. November 2024. DOI: 10.69646/14sbac47a.
11. Kolarski, A., Nina, A., Srećković, V. A., **Arnaut, F.**, & Langović, M. (2024). Earth's lower ionosphere under energetic events: solar flares and gamma ray bursts as drivers for VLF signal perturbations. *International Meeting on Data for Atomic and Molecular Processes in Plasmas: Advances in Standards and Modelling. Book of Abstracts and Contributed Papers* . <https://doi.org/10.69646/aob241116>
12. Arnaut, F., Kolarski, A., Langović, M., Srećković, V. A., & Jevremović, S. (2024). Machine learning classification difficulties of VLF amplitude variations around the terminator. *International Meeting on Data for Atomic and Molecular Processes in Plasmas: Advances in Standards and Modelling. Book of Abstracts and Contributed Papers*. <https://doi.org/10.69646/aob241116>
13. Jevremović, S., Langović, M., Kolarski, A., & Arnaut, F. (2024). Annual PM2.5 and PM10 variations on Belgrade's Mostar interchange – Traffic impact. *Meeting on new trends in Astronomy & Earth Observation. Book of Abstracts and Contributed Papers*.
14. Langović, M., Srećković, V. A., Kolarski, A., Arnaut, F., & Jevremović, S. (2024). Impacts of seismic activity on human mobility. *Meeting on new trends in Astronomy & Earth Observation. Book of Abstracts and Contributed Papers*.
15. Srećković, V. A., Mijić, Z., Kolarski, A., Langović, M., Arnaut, F., Jevremović, S., Barović, J., Kounchev, O., & Simeonov, G. (2024). Novel research in astronomy & earth observation. *Meeting on new trends in Astronomy & Earth Observation. Book of Abstracts and Contributed Papers*.

Категорија M51:

1. **Arnaut F.**, 2021. Primena ukrštenog kvadratnog dispozitiva za detekciju primarne zone rasta korena drveta vrbe. Zapisnici Srpskog geološkog društva za 2020. godinu, 29-43
2. **Arnaut F.**, Sretenović B., 2021. 2D Electrical imaging investigations at the Neolithic settlement „Pločnik”, Bulletin of the Natural History Museum, 14, 46-64.
3. Sretenović, B., **Arnaut, F.**, Vasiljević, I., Cvetkov, V., 2019. 2D Geoelectrical Resistivity Tomography Application at the Former City Waste Dump “Ada Huja” - Eco-geological Problem. Podzemni Radovi, 34, 59-76.
4. Simić, V., Petrović, S., **Arnaut, F.**, Cvetkov, V., Kostović, M., Radulović, D., Stojanović, J., Jovanović, V., Todorović, D., Nikolić, N., Senčanski, J., Bogdanović, G., and Marilović, D. (2024). PRIZMA project: Characterisation and technological procedures for recycling and reusing of the Rudnik mine flotation tailings. *Reports of the Serbian Geological Society for the year 2023 / Comptes rendus des séances de la*

Société serbe de géologie pour les années 2023 / Записници Српског геолошког друштва (за 2023. годину), Belgrade, 2024.

Категорија М52:

1. **Arnaut F.**, Sretenović B., 2021. Landslide Investigations Using the geoelectrical Scanning Method and Ground Penetrating Radar -case Studies. Tehnika, 72(4), 419-425.

Категорија М53:

1. Medarević A., Borčić K., **Arnaut F.**, 2024. A retrospective study on road traffic injuries in the Republic of Serbia during the COVID-19 lockdown. Glasnik javnog zdravlja. 98, 265-279

Категорија М63:

1. **Arnaut F.**, Cvetkov, V., Đurić, D., 2023. Prognoziranje iznadprosečnih vrednosti kvaliteta vazduha u Novom Sadu korišćenjem Random Forest Modela. 10. Memorijalni naučni skup iz zaštite životne sredine "Docent Dr Milena Dalmacija", Novi Sad, 30.3.2023- 31.3.2023.
2. **Arnaut F.**, Sretenović B., Damnjanović V., Cvetkov V., 2022. Elektrometrijska ispitivanja vodoizvorišta Makiš. 9. Memorijalni naučni skup iz zaštite životne sredine „Docent Dr Milena Dalmacija”, Novi Sad 31.03.2022- 01.04.2022.
3. **Arnaut F.**, Cvetkov V., Đurić D., 2022. Primena jednostavnih metoda predviđanja na podatke vertikalnog elektrometrijskog sondiranja. XVIII Kongres Geologa Srbije. Divčibare 1- 4 Jun 2022, 33-34.

Остало:

1. **Arnaut F.**, Sretenović B., 2021. Some examples of electrical imaging (ERT) and ground penetrating radar (GPR) in solving civil engineering problems. Geologica Macedonica, 35(2), 73-84.
2. Sretenovic, B., **Arnaut, F.**, 2019. Inadequacy of 1D, 2D and 3D Resistivity Inverse Modelling in the Presence of Electrical Anisotropy. Earth Sciences, 8, 102. <https://doi.org/10.11648/j.earth.20190802.14>

УНИВЕРЗИТЕТ У БЕОГРАДУ-
РУДАРСКО-ГЕОЛОШКИ ФАКУЛТЕТ
11120 Београд 35, Ђушина 7, п.п. 35-62
Тел: (011) 3219-100, Факс: (011) 3235-539



UNIVERSITY OF BELGRADE,
FACULTY OF MINING AND GEOLOGY
Republic of Serbia, Belgrade, Djusina 7
Phone: (381 11) 3219-100, Fax: (381 11) 3235-539

Потврда о упису школске 2024/2025. године

Подаци о студенту

Индекс	Г804/21
Име и презиме	Филип Арнаут
Ниво квалификације	Докторске академске студије
Студијски програм	Геологија (ДГЛ)
Тренутни статус	Буџет

Упис године

Школска година	2024/2025.	Школарина	0 дин.
Статус	Самофинансирање	Уплата студента	0 дин.
Одабрано ЕСПБ	5 (0 + 5 + 0 + 0)		

Одабрани курсеви за школску 2024/2025. годину

Ознака	Назив	Статус	Семестар	ЕСПБ
20-ЗОДГЕ	Одбрана докторске дисертације	0	2	5
Укупно:				5

Београд, 15. октобар 2024.

Зорица Кисин



Република Србија
Универзитет у Београду

Оснивач: Република Србија

Дозволу за рад број 612-00-02666/2010-04 од 12. октобра 2011.
године је издало Министарство просвете и науке Републике Србије

Рударско-геолошки факултет, Београд

Оснивач: Република Србија

Дозволу за рад број 612-00-01125/2009-04 од 11. јануара 2010.
године је издало Министарство просвете Републике Србије

УБ



Диплома

Филиј, Душан, Арнауи

рођен 14. јула 1997. године, Београд, Република Србија, уписан школске
2016/2017. године, а дана 16. септембра 2020. године завршио је основне
академске студије, првог степена, на студијском програму Геофизика, обима
240 (двеста четрдесет) бодова ЕСПБ са просечном оценом 8,29 (осам и 29/100).

На основу тога издаје му се ова диплома о стеченом високом образовању и стручном називу
дипломирани инжењер геологије

Број: 11626300

У Београду, 2. фебруара 2021. године

Декан
Проф. др Зоран Глигорић

Ректор
Проф. др Иванка Појовић

00116452



Република Србија
Универзитет у Београду

Оснивач: Република Србија

Дозволу за рад број 612-00-02666/2010-04 од 12. октобра 2011.
године је издало Министарство просвете и науке Републике Србије

Рударско-геолошки факултет, Београд

Оснивач: Република Србија

Дозволу за рад број 612-00-01125/2009-04 од 11. јануара 2010.
године је издало Министарство просвете Републике Србије

УБ



Диплома

Филиј, Душан, Арнаућ

рођен 14. јула 1997. године, Београд, Република Србија, уписан школске
2020/2021. године, а дана 6. јула 2021. године завршио је мастер академске
студије, другој степену, на студијском програму Геофизика, обима 60
(шездесет) бодова ЕСПБ са просечном оценом 9,36 (девет и 36/100).

На основу тога издаје му се ова диплома о стиченом високом образовању и академском називу
мастер инжењер геологије

Број: 12487400

У Београду, 31. августа 2021. године

Декан
Проф. др Зоран Глијорић

Ректор
Проф. др Иванка Појовић

00125113

На основу члана 67. и 114. Статута Рударско-геолошког факултета Универзитета у Београду и члана 30. Правилника о докторским студијама на Рударско-геолошком факултету Универзитета у Београду, Наставно-научно веће Рударско-геолошког факултета Универзитета у Београду, на својој седници одржаној 18.04.2024. године, донело је

ОДЛУКУ

1. Усваја се извештај Комисије за оцену научне заснованости теме докторске дисертације **Филипа Арнаута, мастер инж. геологије**.
2. Одобрава се именованом израда докторске дисертације под насловом „Примена модела вођених подацима у геофизици“.
3. Тема докторске дисертација је оригинална идеја, од значаја за развој науке, примену њених резултата, односно развој научне мисли уопште.
4. За ментора се именује др Весна Цветков, ред. проф.
5. Сагласност на предлог теме докторске дисертације из ове одлуке даје Веће научних области техничких наука.



ДЕКАН

М. Аболмасов
др. Биљана Аболмасов, ред. проф.

Достављено:

- Већу научних области техничких наука
- Ментору
- Именованом
- Одељењу за студентска питања

ПУБЛИКАЦИЈЕ

КАТЕГОРИЈА М20

Article

Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation

Filip Arnaut ^{*}, Aleksandra Kolarski  and Vladimir A. Srećković 

Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; aleksandra.kolarski@ipb.ac.rs (A.K.); vlada@ipb.ac.rs (V.A.S.)

* Correspondence: filip.arnaut@ipb.ac.rs

Abstract: The process of manually checking, validating, and excluding data in an ionospheric very-low-frequency (VLF) analysis during extreme events is a labor-intensive and time-consuming task. However, this task can be automated through the utilization of machine learning (ML) classification techniques. This research paper employed the Random Forest (RF) classification algorithm to automatically classify the impact of solar flares on ionospheric VLF data and erroneous data points, such as instrumentation errors and noisy data. The data used for analysis were collected during September and October 2011, encompassing solar flare classes ranging from C2.5 to X2.1. The F1-score values obtained from the test dataset displayed values of 0.848; meanwhile, a more detailed analysis revealed that, due to the imbalanced distribution of the target class, the per-class F1-score indicated higher values for the normal data point class (0.69–0.97) compared to those of the anomalous data point class (0.31 to 0.71). Instances of successful and inadequate categorization were analyzed and presented visually. This research investigated the potential application of ML techniques in the automated identification and classification of erroneous VLF amplitude data points; however, the findings of this research hold promise for the detection of short-term ionospheric responses to, e.g., gamma ray bursts (GRBs), or in the analysis of pre-earthquake ionospheric anomalies.

Keywords: machine learning; solar flares; VLF ionospheric data; anomaly classification; anomaly detection; class imbalance



Citation: Arnaut, F.; Kolarski, A.; Srećković, V.A. Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. *Universe* **2023**, *9*, 436. <https://doi.org/10.3390/universe9100436>

Academic Editor: Dmitrii Kolotkov

Received: 8 August 2023

Revised: 14 September 2023

Accepted: 28 September 2023

Published: 30 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the study of the terrestrial ionosphere in relation to various extreme phenomena has become a very interesting topic. Using various observation techniques, many scientists around the world have made significant progress in studying the imprints of extreme phenomena in the ionosphere [1,2]. As a result, the ionosphere is recognized as a useful “tool” for studying the disturbances caused by such phenomena. Modern communication systems, including satellite and navigation systems, as well as various radio signals, traverse the atmosphere and lower ionosphere. It is worth noting that these signals may encounter disturbances during periods of solar perturbations. X-ray solar flares can be classified based on their influence on VLF wave propagation via the Earth–ionosphere waveguide [3]. The significance of modeling the lower ionosphere cannot be overstated in relation to diverse technological, research, and industrial fields [4].

VLF ionospheric data modelling, like any other type of modelling, involves a preliminary data pre-processing phase. During this phase, data filtering and domain-specific transformations are usually applied, along with the elimination of erroneous data points. In the context of VLF ionospheric amplitude analysis, the methodology bears resemblance to the aforementioned approach. However, it diverges in the sense that the elimination of two distinct categories of erroneous data points is required. The first category pertains to instrumentation errors, wherein the VLF receiver produces flawed measurements. The second category involves the researcher’s decision to either exclude or annotate the data

points affected by the solar flare influences on the VLF data. The process of manually eliminating and/or annotating erroneous data in VLF data analysis is subject to an additional consideration, namely, the high measurement resolution, i.e., measurements taken at one minute intervals. The volume of data gathered during a given period of investigation can be overwhelming, requiring significant time and effort to manually sift through it and exclude. Consequently, the automation of this process is considered to be highly advantageous.

One crucial element in the field of space weather physics pertains to the interrelationship among solar flare (SF) occurrences, the ionospheric reaction to such events, and Coronal Mass Ejections (CME) [2]. The importance of SF occurrences, including those classified as lower M-class rather than X-class, holds considerable significance and is currently a research focus. SF events classified as M-class exhibit a lower correlation with CMEs compared to X-class flares; however, a subset of M-class flares has been found to be associated with faster CMEs [5], which aligns with the concept known as Big Flare Syndrome [5,6]. Moreover, there exists a significant correlation between X-class SFs of great intensity and CMEs, which have been observed to induce disruptions to satellite and navigation systems [7].

The current research in the fields of SFs, CMEs, and the ionosphere is focused on the application of machine learning (ML) techniques. Classification methods have been employed in the classification of lightning waveforms [8], as well as in the classification of radar returns [9–13] and auroral image classification [14,15]. Similar to the pre-processing phase of VLF data analysis, manual radar return classification entails human intervention and is a labor-intensive procedure [9]. The resemblance between the two aforementioned processes underscores the justification for employing ML classification techniques in order to automatically eliminate and/or classify erroneous ionospheric VLF data.

In order to streamline the manual exclusion or labeling of inaccurate ionospheric VLF data, the application of ML classification techniques has been considered. The aim of this study was to employ pre-existing ionospheric VLF amplitude data (Worldwide Archive Of Low-Frequency Data And Observations—WALDO, <https://waldo.world/>, accessed on 24 March 2023) that have been previously utilized in research, as well as soft range X-ray irradiance data (Geostationary Operational Environmental Satellite—GOES, <https://www.ncei.noaa.gov/>, accessed on 24 March 2023), with the purpose of investigating the feasibility of automatically labeling erroneous data and the effects of increased solar flare activity on measured VLF data. The task involved the utilization of the Random Forest (RF) method for ML classification. The data used for this task consisted of a total of 19 transmitter–receiver (T-R) pairs which are situated in North America. The data spanned the time period between September 2011 and October 2011, during which, solar flare activity was observed. In addition, in the manuscript, we discuss how the presented research serves as a potential method for the automated labeling of data in various fields of space science. The used datasets, results, and a post-processing workflow can be found on Zenodo: <https://zenodo.org/record/8220971>, accessed on 7 August 2023 (Supplementary Materials).

2. Materials and Methods

The present study made use of data obtained in 2011, specifically data gathered during solar flare events that took place in September (ranging from C2.5 to X2.1) and October (ranging from C5.5 to M1.5). The data employed in this study had already undergone a process of “labeling”, wherein erroneous data points were previously excluded (Figure 1). This circumstance presented a distinctive opportunity to utilize the dataset for the purpose of training a ML model, which, in turn, could potentially automate the process of labeling in subsequent instances. Figure 1 illustrates four instances in which a researcher was required to manually exclude erroneous data points from the dataset in preparation for a subsequent analysis. Figure 1a illustrates the impact of noisy data or errors in instrumentation on VLF data. Conversely, Figure 1b demonstrates the effects of SFs on VLF data, including the presence of outlier data points (single data points that significantly deviate from the rest of the measured data points). Figure 1c,d exhibit a confluence of SF impacts, anomalous data

points, and instrumentation errors. The aforementioned instances have a direct impact on the processing time required, i.e., the “data cleaning”. This is particularly significant due to the high measurement rate and the time spans of VLF data. Consequently, this may result in large datasets and necessitate a substantial amount of time for the manual exclusion of such data.

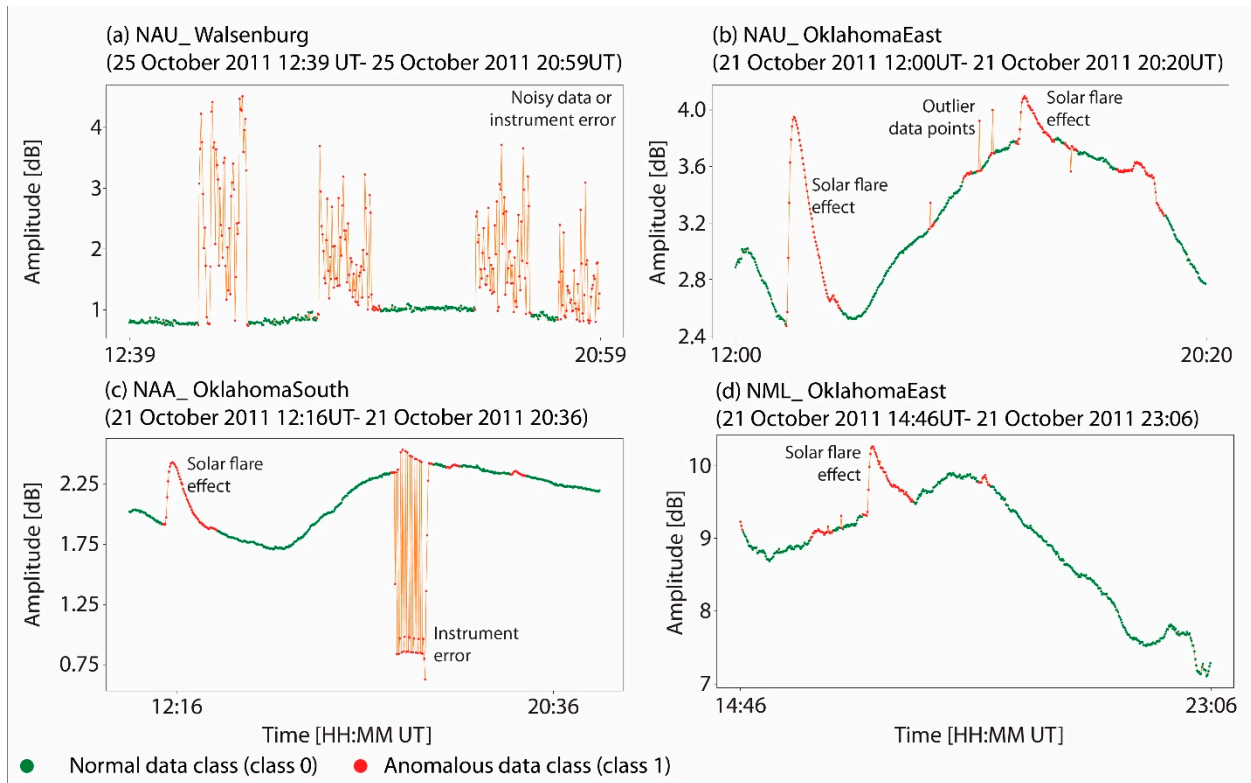


Figure 1. Illustrations of erroneous VLF data; (a) example of instrumentation error or noisy data; (b) example of solar flare effects on VLF data with outlier data points; (c) example of solar flare effects and instrumentation error; and (d) example of solar flare effects and outlier data points.

The dataset employed in this study comprised five VLF transmitters, namely NPM, NLK, NML, NAA, and NAU, along with four VLF receivers, specifically Walsenburg, Oklahoma South, Oklahoma East, and Sheridan (Figure 2). There was a total of 19 T-R pairs.

In standard ML workflows, a crucial step involves the pre-processing of data (Figure 3). During this phase, the data are transformed in order to meet to the requirements of the ML methods, i.e., dataset formats. Additionally, features are extracted, which will serve the purpose of revealing important patterns and characteristics within the dataset. Furthermore, the complete dataset is divided into separate training and testing datasets. In cases where the classes within the training dataset are imbalanced, it is common practice to balance them. This balancing process enhances the evaluation metrics and predictive capabilities of the model.

The samples, in which the excluded data points were appropriately labeled, were merged with the original dataset, comprising both the excluded and non-excluded data points. This integration resulted in the creation of a unified database, serving as the foundation for the ML modeling process. The database was updated with soft-range X-ray irradiance data [16], VLF transmitter, and VLF receiver information [17], as well as local receiver time data. The primary features of the database were the VLF amplitude data and X-ray data. These features served as the basis for calculating the other features. Additionally, the transmitter, receiver, and local receiver time were considered as secondary features and played crucial roles in establishing the core of the database. The target variable was encoded as binary data, where the data points from the labeled samples that were

excluded were assigned a value of 1 in the target variable, indicating anomalous data points. Conversely, the data points that were retained in the labeled sample were assigned a value of 0 in the target variable, representing normal data points. A process of filtering the X-ray and VLF data was conducted to eliminate any data points that lacked a measured X-ray variable or VLF amplitude variable.

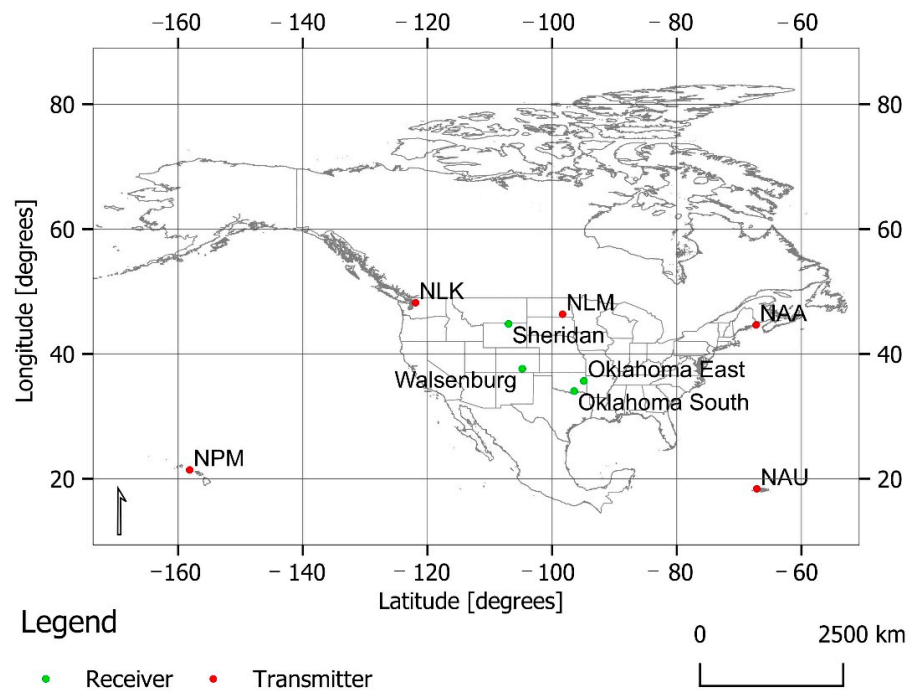


Figure 2. Spatial distribution of VLF transmitters and receivers utilized in this research (base map data obtained from DIVA-GIS, <https://www.diva-gis.org/gdata>, accessed on 1 April 2023).

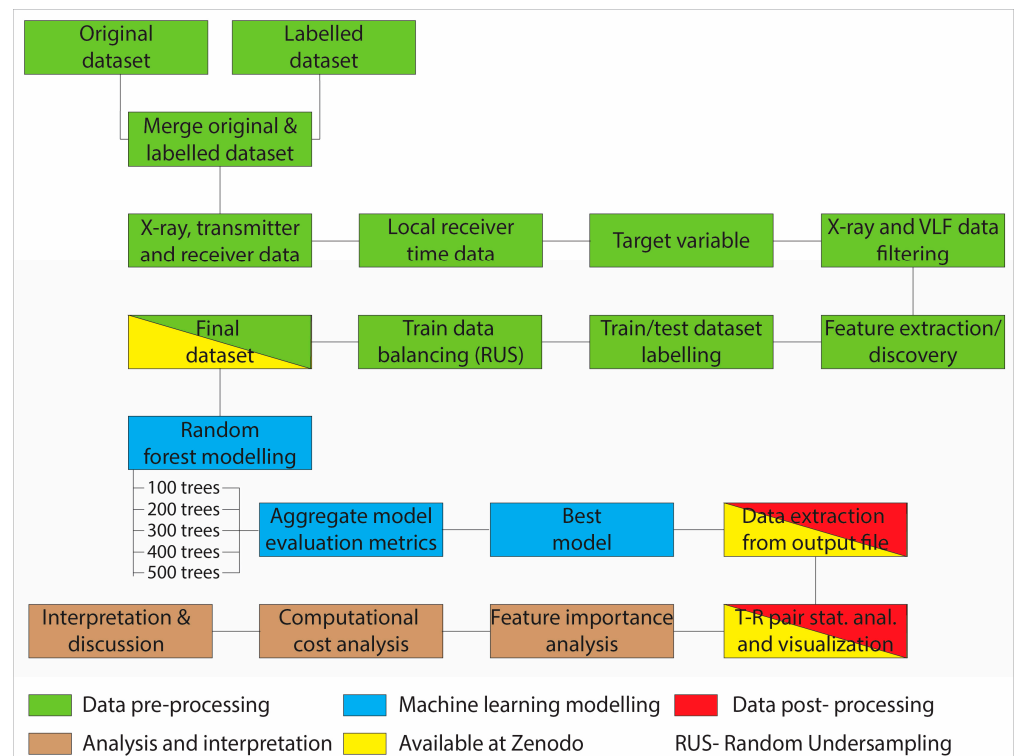


Figure 3. Pre-processing, modelling, and post-processing workflow.

The process of feature extraction, also referred to as discovery, pertains to the identification of features which are denoted as tertiary features in this study. The tertiary features were calculated based on the primary features, namely the VLF amplitude and X-ray data. The tertiary features were classified as statistical features, as they contained relevant information regarding the rolling window statistics. These statistics included the standard deviation, mean, and median values of the rolling window, which were computed for various window sizes. Specifically, the window sizes considered were 5 (short-term), 20 (mid-term), and 180 (long-term) minutes, representing different time dependencies within the dataset. Furthermore, the data were augmented by incorporating a lagged signal for time intervals ranging from 1 to 5 min. Additionally, the rate of change, as well as the first and second differentials, were calculated for the primary features' data. The most recent additions corresponded to a set of binary features, which encoded whether a given data point exceeded the mean or median value of the VLF amplitude data. The primary objective of the tertiary features was to determine whether any statistical parameters contained valuable information for the ML classification task. The total number of features was 41, as tertiary features were computed for both the VLF amplitude data and X-ray data, in addition to those exceeding the mean or median values. An overview of these features is presented in Table 1.

Table 1. Hierarchical feature classification for RF modelling.

Feature	Class	Class 2	Feature	Class	Class 2
VLF amplitude	P	Measured	RW median	T	Calculated
X-ray	P	Measured	Lagged signal	T	Calculated
Receiver local time	S	Measured	Rate of change	T	Calculated
Transmitter	S	/	First difference	T	Calculated
Receiver	S	/	Second difference	T	Calculated
RW st. dev.	T	Calculated	Higher than mean	T	Calculated
RW mean	T	Calculated	Higher than median	T	Calculated

RW—Rolling window; st. dev—Standard deviation; P—Primary; S—Secondary; and T—Tertiary.

Once the features were generated for the entire database, it was divided into separate training and testing databases. Subsequently, the training and testing data points were appropriately labeled to facilitate their utilization in the JASP software [18]. The last stage of data pre-processing involved addressing the imbalance in the training dataset, as datasets with imbalanced class distributions can introduce bias towards the majority class [19–21]. Imbalanced classification tasks pertain to an inherent disparity or disproportion between classes in binary classification problems. The methods employed to address this class imbalance can be categorized into two main approaches: under-sampling and oversampling. Under-sampling involves applying methods to the majority class to reduce the number of instances, while oversampling entails techniques applied to the minority class to increase the number of minority instances. The present study employed the random under-sampling technique [22], which involves the random removal of instances from the majority class [23,24]. The issue of imbalanced classification is commonly observed in anomaly detection scenarios, as highlighted by [25]. In the present study, a similar situation arose, where one class represented the normal category, while the other class pertained to the anomalous data category, specifically the excluded data class.

The Random Forest (RF) algorithm was proposed by Breiman in 2001 [26]. It has been extensively utilized in various scientific and industrial domains for classification and regression tasks over the past two decades. The RF algorithm is known for its ability to avoid overfitting due to the implementation of averaging or voting [27] and the utilization of the law of large numbers [26]. This algorithm has gained significant popularity due to its simplicity, as it only requires the specification of the number of trees (the decision trees used to control the complexity of the model) as a hyperparameter.

The initial RF classification in this research was conducted using five models, distinguished solely by the varying number of trees employed. The quantity of trees varied between 100 and 500, increasing by increments of 100 trees. The evaluation metrics of the aggregate model were analyzed to determine the optimal model. In cases where there was no clear best model, the model parsimony method could be employed. This method selected the model with the fewest hyperparameters, indicating that simplicity was a desirable quality in the best model.

The study employed several classification evaluation metrics, including accuracy, precision, recall, false positive rate, AUC, F1-score, Matthew's correlation coefficient (MCC), and statistical parity parameter.

- (a) The accuracy parameter can be defined as the proportion of instances that were correctly classified out of the total number of instances, encompassing both true positive and true negative classifications.
- (b) The precision parameter is defined as the proportion of correct positive predictions out of the total number of predicted positives, specifically, the ratio of true positives to the sum of true positives and false positives.
- (c) The recall parameter, which is alternatively referred to as the true positive rate, quantified the proportion of correctly identified positive instances in relation to the combined count of true positives and true negatives. In other words, recall assessed the fraction of positive instances that were accurately classified [28].
- (d) The false positive rate refers to the proportion of incorrect positive predictions relative to the overall number of instances in the negative class. Specifically, it quantified the ratio of misclassified instances that were predicted as being in the negative class, but were actually part of the positive class, to the total number of instances in the negative class.
- (e) The Area Under the Receiver Operating Characteristic Curve (AUC) is a commonly employed, single-number metric for evaluating classification performance [29]. It is suggested as a comprehensive measure in ML classification tasks [30], and is a widely used measure for assessing the overall classification performance of a classifier [28]. Compared to accuracy, the AUC is considered to be a superior metric [30]. Moreover, the AUC parameter was employed to determine the model's ability to differentiate between the positive and negative classes, which refers to the model's discrimination performance [31]. From a practical standpoint, AUC values of 0.5 indicated that the model lacked the ability to distinguish between classes and performed random classification. Conversely, AUC values closer to 1 were more desirable, since they represent better classification models.
- (f) The F1-score evaluation metric calculated the harmonic mean between the recall (true positive rate) and precision. In the context of imbalanced binary classification tasks, the F1-score has been identified as a more desirable metric compared to accuracy [28,32].
- (g) The Matthew's correlation coefficient (MCC) is considered to be an even more favorable alternative to both F1-score and accuracy, due to its reliability and the requirement for a high MCC score to indicate a successful performance across all four categories of the confusion matrix (true positive, false positive, true negative, and false negative) [33]. Moreover, it has been observed that both accuracy and F1-score can exhibit inflated values compared to the actual value when dealing with imbalanced datasets [33], and that the MCC should be used as a standard for imbalanced datasets [34].
- (h) The statistical parity parameter, which is the simplest parameter out of all those mentioned earlier, indicated the proportion of classified instances per-class out of all the classified instances. In other words, the per-class values of the statistical parity parameter should closely align with the actual distribution of the data in the test set for a good classification model. The metrics that were previously presented were employed based on specific requirements, and furthermore, on a per-class basis, to ensure a comprehensive statistical analysis.

3. Results

The workflow consisted of multiple stages in the processing of the data, with the primary stages being: the pre-processing of the data (including the division of the data into training and testing sets, as well as balancing the classes within the training dataset), modeling using RF with different numbers of trees, the selection of the best model, and the evaluation of the metric statistics for each T-R pair. These stages are presented in more detail in the following sections.

3.1. Data Pre-Processing

The research utilized a full set of training data consisting of 19 T-R pairs that collected data during solar flare events in September 2011, measured with 1 min intervals. The dataset comprised a total of 135,308 data points prior to balancing, with a class distribution of 22% for the anomalous data (designated as 1 in the database) and 78% for the normal data (class 0). Following the implementation of the random under-sampling technique, the class distribution achieved a balanced state, with an equal distribution of 50% for each class. The resulting training dataset consisted of a total of 59,344 data points. The database's testing phase encompassed solar flare events that were documented in October 2011. It consisted of 19 T-R pairs and a total of 180,071 data points.

3.2. Random Forest Modelling

The RF modeling was conducted using a range of trees, ranging from 100 to 500, with increments of 100 trees. Figure 4 depicts the accuracy, precision, F1-score, and AUC parameter for all five models. The values of each model for all four evaluation metrics exhibited a notable degree of similarity. The out-of-bag classification accuracy demonstrated convergence in the random forest model with 50 trees, suggesting that the subsequent models would yield comparable outcomes. The RF model, consisting of 100 trees, was selected as the optimal model for the study. The initial rationale behind this observation was that, despite the close resemblance of the four evaluation metrics, the RF model with 100 trees exhibited marginally superior outcomes in terms of accuracy and F1-score. Another factor considered was that models with a smaller number of trees require less computational resources, i.e., model parsimony. The ideal model is the one that utilizes the fewest (hyper)parameters and is the simplest. Furthermore, the statistical parity evaluation metric revealed that the RF model, employing a total of 100 trees and utilizing a balanced dataset, exhibited the greatest degree of success in predicting the class ratio within the testing dataset. Specifically, the RF model correctly identified 16.3% of the instances as anomalous, which closely aligned with the actual proportion of 15% anomalous instances present in the testing dataset. The other RF models exhibited a range of anomalous instances, with percentages varying from 16.7% to 17%.

In order to determine the optimal model, additional testing was conducted by varying the number of trees in the RF model. Specifically, the RF model was tested with 125, 150, and 175 trees, in addition to the initial model with 100 trees. The outcomes exhibited a high degree of comparability across the models, with minimal variations in their overall classification efficacy. Consequently, the RF model with 100 trees maintained its status as the superior model.

Within the testing dataset for the RF model, separation between various T-R pairs and the calculated evaluation metrics was conducted (Figure 5). The analysis involved interpreting the T-R pairs separately to enhance the understanding of the model's ability to distinguish between the normal and anomalous data points. The evaluation metrics of accuracy revealed that the T-R pair NML–Oklahoma South exhibited a distinct outlier status, with values of approximately 0.6. Regarding precision, both NPM–Walsenburg and NAA–Walsenburg were outliers, as they achieved a high precision score of approximately 0.94. In terms of the F1-score, NPM–Walsenburg exhibited relatively increased values. A median value of 0.87 for the F1-score suggested that half of the models outperformed a score of 0.87 when evaluated using the F1-score metric. An additional parameter employed

in this analysis was the MCC, which exhibited a varied range, spanning from 0.7 (NPM–Walsenburg) to 0.14 (NLK–Oklahoma South), with a median value of 0.45. Based on the analysis conducted using individual T-R pairs and focusing solely on the overall evaluation metrics, it can be concluded that NPM–Walsenburg exhibited the most favorable evaluation metrics among all the T-R pairs, whereas NML–Oklahoma South demonstrated the least favorable performance.

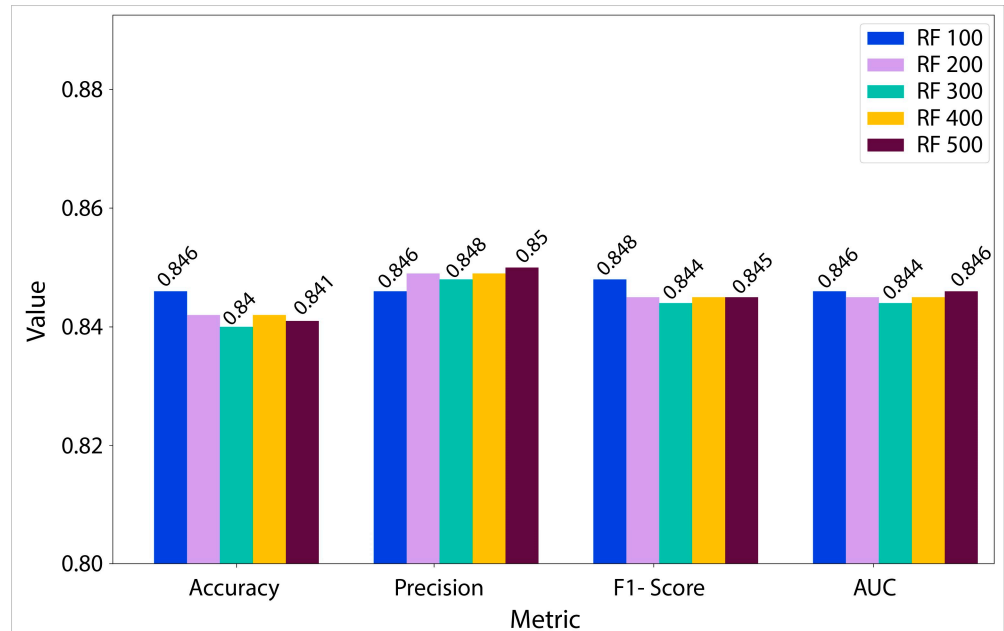


Figure 4. Comparison of Random Forest models with varying numbers of trees based on accuracy, precision, F1-Score, and the Area Under the Receiver Operating Characteristic Curve (AUC).

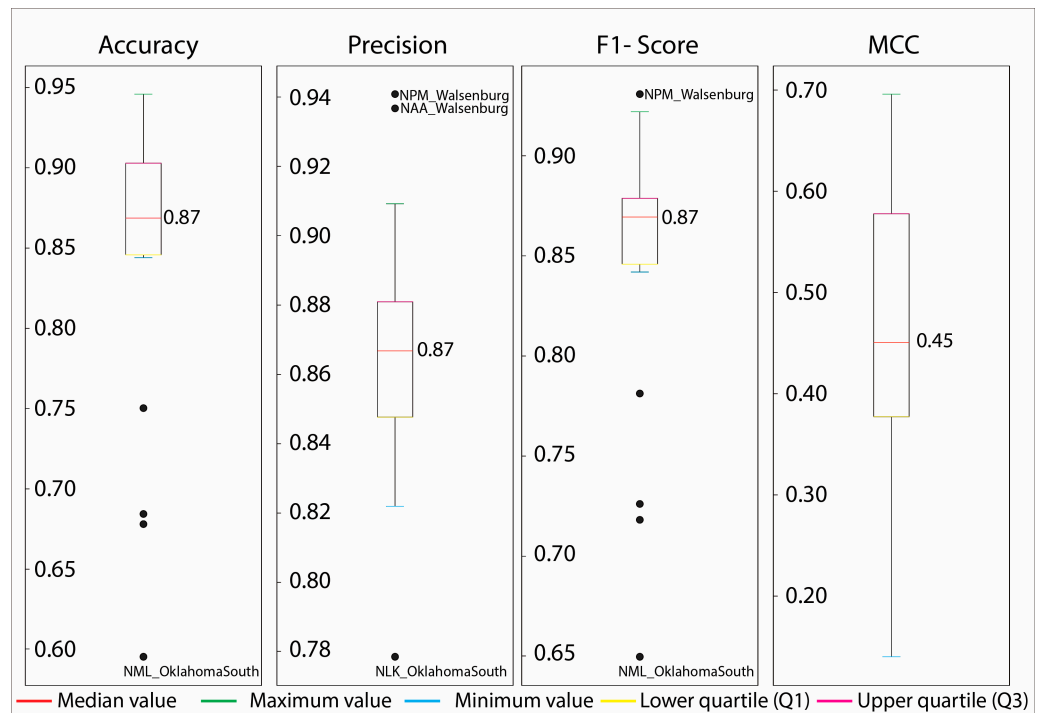


Figure 5. Accuracy, precision, F1-Score, and Matthew’s correlation coefficient for individual transmitter–receiver pairs.

A more comprehensive examination was conducted by employing per-class evaluation metrics, which involved the calculation of the evaluation metrics for both the 0 (normal) and 1 (anomalous) classes within the testing database (Figure 6). The primary aim of computing the per-class evaluation metrics was to assess whether the model exhibited a significantly disparate capacity to predict one class compared to the others. Due to the imbalanced nature of this ML problem, the discrepancy may not be readily apparent when considering the overall evaluation metrics. Furthermore, the evaluation metrics for individual classes can offer additional understanding of the variability in the evaluation metrics and the effectiveness of the employed model in terms of classification.

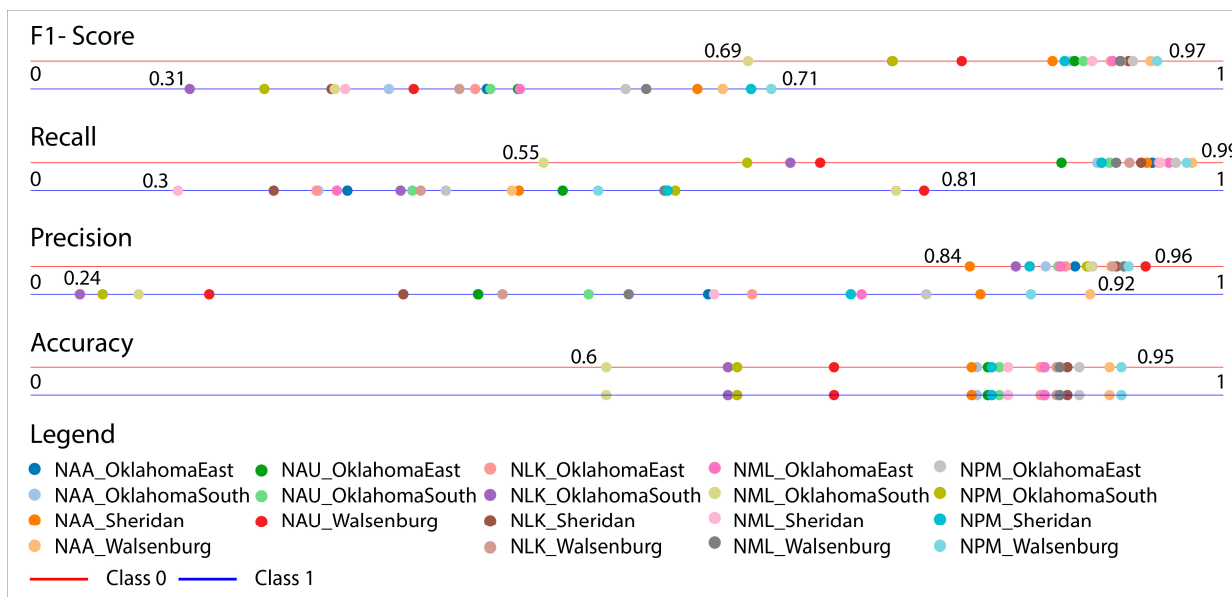


Figure 6. Per-class accuracy, precision, recall, and F1-Score values for individual transmitter–receiver pairs.

The F1-score offered initial insights by revealing that the range of values for class 0 spanned from 0.69 to 0.97, whereas the range for class 1 exhibited a lower minimum of 0.31 and a lower maximum of 0.71. This analysis offers initial observations regarding the classification efficacy of the model. Specifically, it revealed that the geometric mean of the true positive rate and precision was comparatively lower for the anomalous data class. For instance, the T-R pair NPM–Walsenburg achieved the highest overall F1-score of 0.94, as shown in Figure 5. However, the F1-score for the anomalous data class for the same T-R pair was lower at 0.71, whereas the F1-score for the normal data class was 0.97. Due to the imbalance in the ML task at hand, it was observed that the mean of both values was 0.94. Consequently, a comprehensive statistical analysis was warranted to further investigate this matter. In contrast, the F1-scores for class 1 and class 0 in the NML–Oklahoma South model, which is considered to be the poorest model according to Figure 5, were relatively low at 0.41 and 0.69, respectively. In all of the models that were constructed, it was observed that the F1-score for the anomalous data class was consistently lower than the F1-score for the normal data class. This discrepancy necessitated a per-class statistical analysis of the evaluation metric parameter due to the skewed distribution of the imbalanced problem in relation to the calculated evaluation metrics. The precision parameter revealed a notable disparity between the two classes. Specifically, class 0 exhibited a range of 0.12 (ranging from 0.84 to 0.96), whereas class 1 demonstrated a significantly broader range of 0.68. This discrepancy suggested that all the models achieved a relatively satisfactory precision range for the non-anomalous day class (class 0), but exhibited considerable variation in performance for the anomalous day class.

3.3. In-Depth Analysis of Selected Transmitter–Receiver Pair Classifications

Figure 7 gives an instance of the classification accuracy, showcasing the best two T-R pairs, namely NPM–Walsenburg and NAA–Walsenburg. In both visual representations, the upper panel presents X-ray irradiance data and the middle panel corresponds to the true labeling of the data, which was the manual classification performed by a researcher. Conversely, the lower panel represents the classification achieved through the utilization of an ML model.

The NPM–Walsenburg instance was selected due to its demonstration of an error in the signal measurement, which the model effectively detected and classified as anomalous data. It is important to acknowledge that not all the data points within the interrupted signal were categorized as anomalous; rather, a subset of data points was identified as normal. However, it is worth noting that only a limited number of instances of such occurrences were observed. Furthermore, during the start of the signal, a minor solar flare event was accurately identified as anomalous, aligning with the manual classification performed by the researchers. The classification of the amplitude signal as non-anomalous before the interruption of the signal and after the occurrence of the solar flare event was incorrect.

In contrast, the NAA–Walsenburg case study presented a collection of six solar flare events that were identified as anomalous. The RF model accurately detected five of these events, albeit with limited success in accurately determining their duration. Regrettably, one event was entirely omitted by the model.

Both instances exemplified real-world situations that researchers encounter during the processing of ionospheric VLF data, namely, signal interruption and the impact of solar flare events. In both scenarios, the RF model demonstrated a satisfactory classification performance. This approach allows researchers to save time by manually adjusting the data labels instead of fully processing the signal and labeling the data.

Figure 8a exhibits a time section commencing on 19 October 2011 at approximately 13:47 UT time and concluding on the same date at around 22:07 UT time. This time section encompassed a duration of 500 min, i.e., 500 data points. The results of the conducted manual labelling, which involved the interpretation of six anomalous time spans of the VLF amplitude signal, are presented in the middle panel of Figure 8a. Meanwhile, the RF classification of the signal resulted in a solitary, extensive anomalous region, which represented a wholly inaccurate classification of the signal. Consequently, the researcher would need to manually assign a label to the signal. The aforementioned observation can also be applied to the time interval depicted in Figure 8b, specifically that on the 20 October 2011. This interval exhibited a duration of 500 min, similar to the previous example. In this particular instance, a prominent solar flare event is observable in the middle panel of Figure 8b. The observed event was accurately recognized as anomalous, although it is worth noting that a significant portion of the signal was also classified as anomalous, indicating a discrepancy in the data labeling between the researcher and the RF model.

In both aforementioned instances, the inaccurate categorization performed by the RF model served as an illustration of the subpar classification that can also be observed from the model. In both of the instances, the utilization of automatic labelling was not feasible due to the model's inadequate classification performance, necessitating manual intervention.

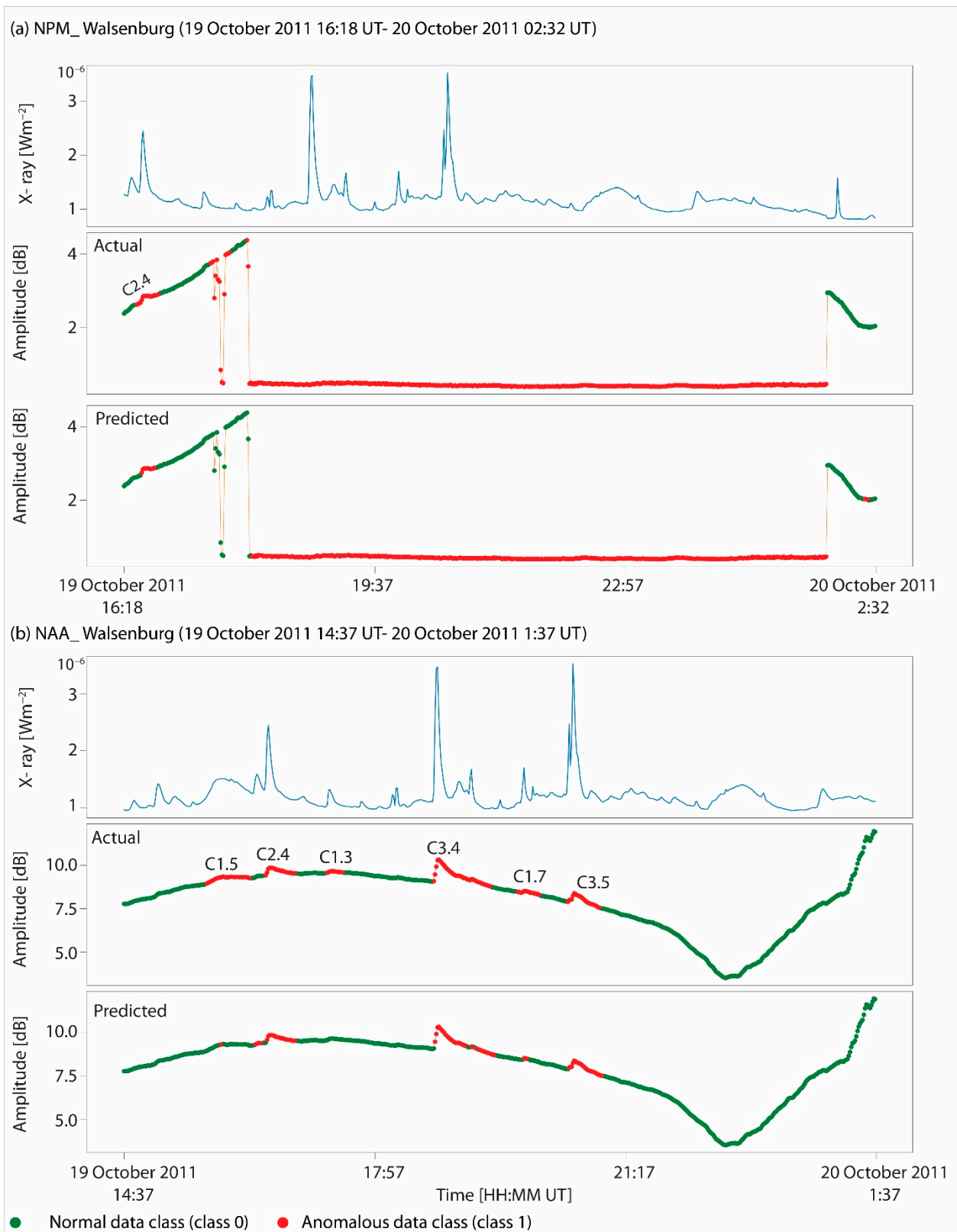


Figure 7. Examples of good classification; (a) transmitter–receiver pair NPM–Walsenburg, top panel—X-ray irradiance (from GOES), middle panel—true classification of the VLF amplitude signal, and bottom panel—output classification of the RF modelling; and (b) transmitter–receiver pair NAA–Walsenburg, top panel—X-ray irradiance, middle panel—true classification of the VLF amplitude signal, and bottom panel—output classification of the RF modelling.

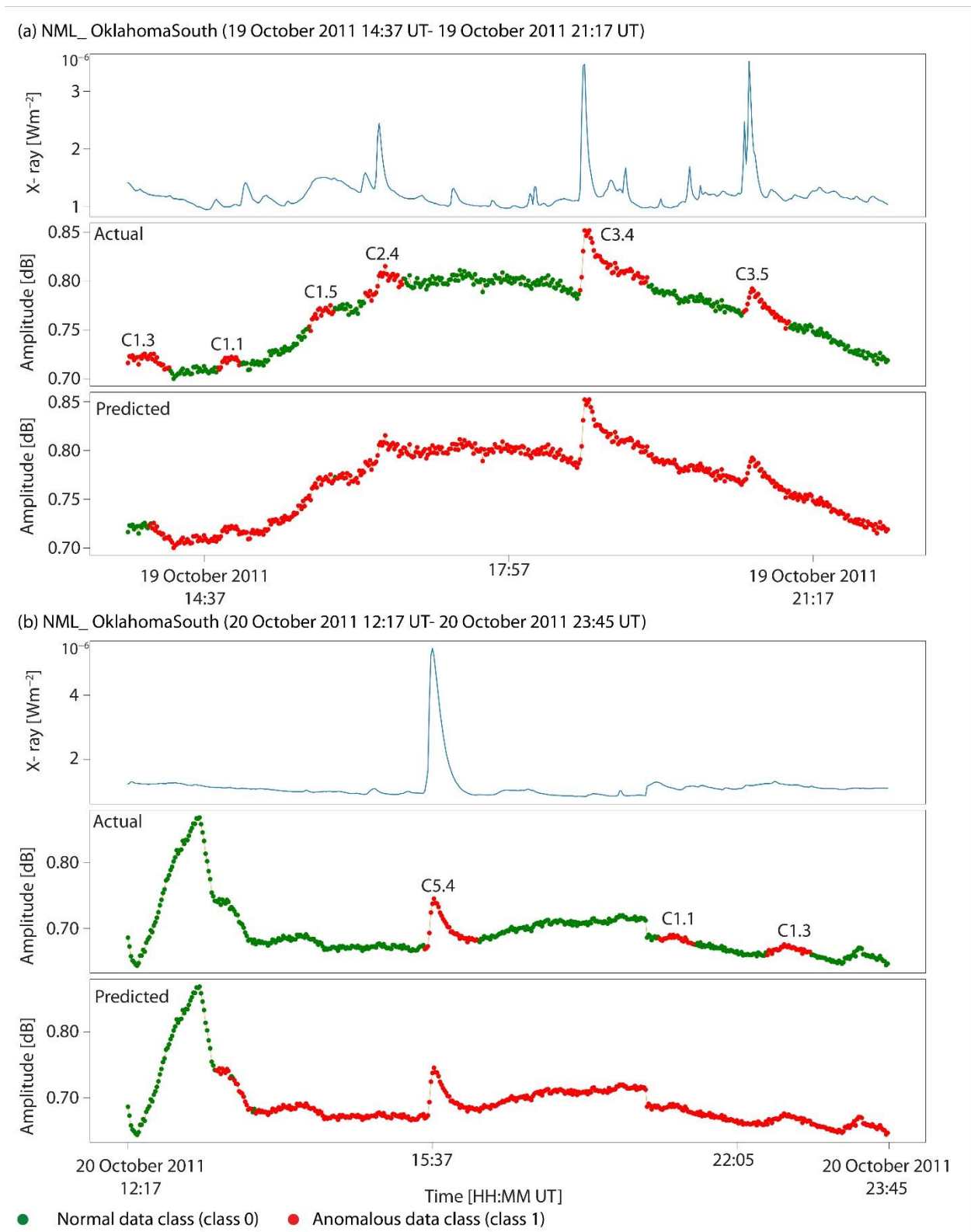


Figure 8. Examples of poor classification; (a) transmitter–receiver pair NML–Oklahoma South, top panel—X-ray irradiance (from GOES), middle panel—true classification of the VLF amplitude signal, and bottom panel—output classification of the RF modelling; and (b) transmitter–receiver pair NML–Oklahoma South, top panel—X-ray irradiance, middle panel—true classification of the VLF amplitude signal, and bottom panel—output classification of the RF modelling.

Model outputs, including solar flares of greater intensity, encompassing both M- and X-class events, are given in Figure 9. Figure 9a displays the T-R pair NAU–Oklahoma East on the 21 October from 11:59 to 20:19 UT. A total of six solar flares were observed, with five falling under the C class category, while one was classified as an M1.3 class solar flare. The RF model utilized in the analysis of the M1.3 solar flare demonstrated a reasonably accurate start time when the researcher opted to exclude the VLF data points. However, it exhibited a poorer classification of the end time for the excluded data. It is important to acknowledge that the M1.3 solar flare persisted and transitioned into a C1.6 solar flare, which was erroneously misclassified by the model. Regarding the additional solar flares depicted in Figure 9a, it is noteworthy that a C2.8 solar flare and a C.16 solar flare were both classified satisfactorily. Moreover, Figure 9a serves as a compelling illustration of outlier detection, showcasing four distinct occurrences of abrupt spikes (outlier data points) in the VLF signal. Notably, these instances were not accurately identified as anomalous by the RF model.

However, when the training and testing samples were rotated to enable the classification of X-class solar flares, the outcomes depicted in Figure 9b were achieved. Figure 9b illustrates an interesting instance and serves as a robust evaluation of the RF model. There were two primary factors contributing to the current event. Firstly, there was a significant solar flare of X-class (X2.1) magnitude. Secondly, there was a notable interruption in the signal, which persisted for a relatively extended period of time. The model successfully identified the majority of data points in the interrupted signal as anomalous. However, it incorrectly classified the beginning of the interrupted signal as non-anomalous. However, it should be noted that the model performed reasonably well in predicting the occurrence of an X-class solar flare. It accurately identified the beginning of the data exclusion period and correctly classified the entire signal as anomalous, with the exception of a slightly shorter duration of anomalous classification at the end compared to the researcher's classification.

Both of these examples provided much needed insights into the model's capabilities and limitations. Future research could potentially enhance this predictive power by conducting fine-tuning experiments and investigating various pre-processing techniques and other ML methods, which, in turn, could potentially provide better classification outcomes.

An examination of the feature importance was conducted subsequent to the determination of the extent of the capabilities exhibited by the best overall model. The findings from the analysis of the feature importance indicated that it was possible to develop a model with a reduced number of features (20) while still maintaining a relatively high level of predictive power and reducing the computational time (cost) by around 50%. For further details about the analysis of the feature importance, refer to Appendix A.

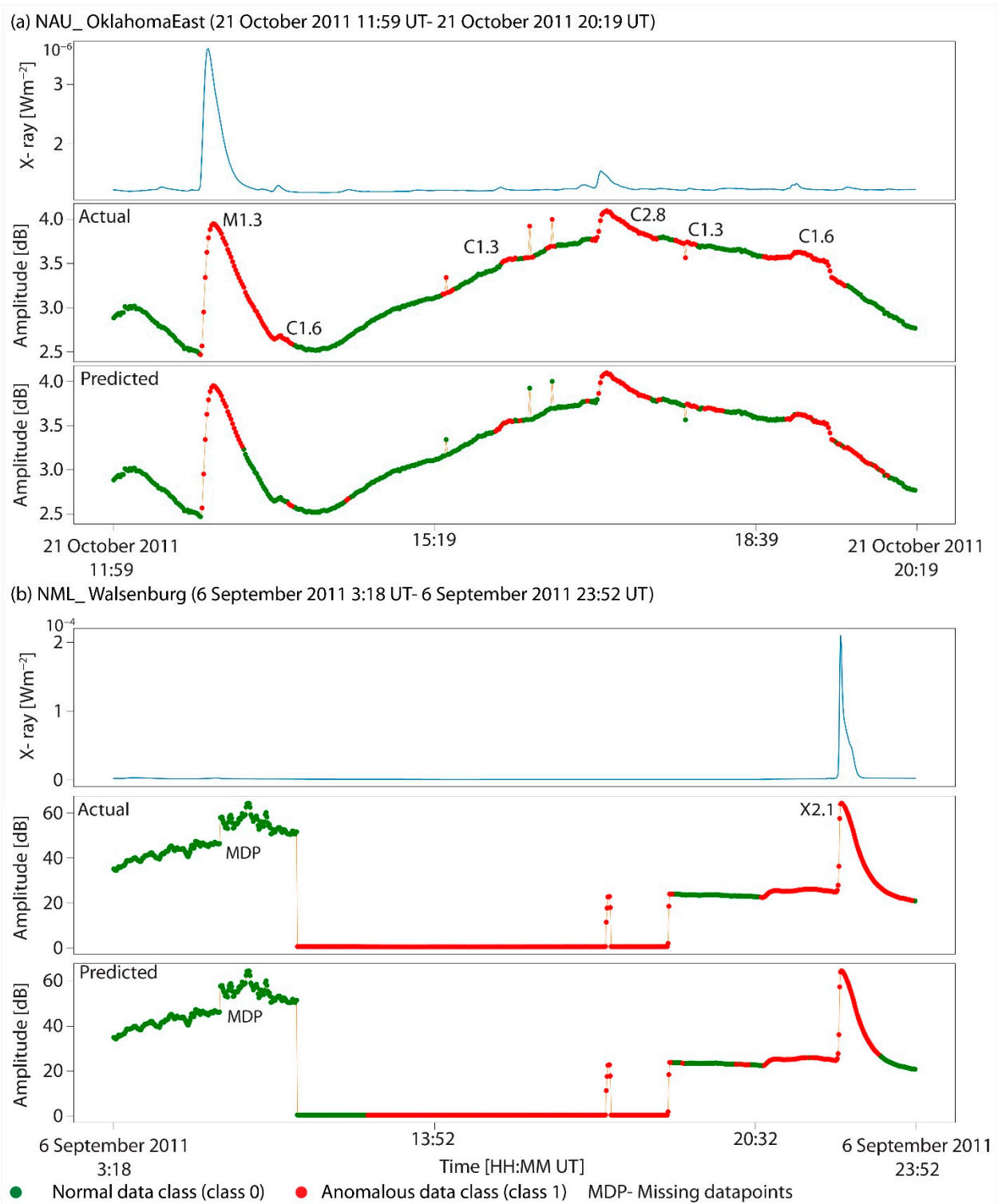


Figure 9. Examples of M- and X-class solar flares' classification; (a) transmitter–receiver pair NAU–Oklahoma East, top panel—X-ray irradiance (from GOES), middle panel—true classification of the VLF amplitude signal, and bottom panel—output classification of the RF modelling; (b) transmitter–receiver pair NML–Walsenburg, top panel—X-ray irradiance, middle panel—true classification of the VLF amplitude signal, and bottom panel—output classification of the RF modelling.

4. Discussion

The research presented in this study serves as an investigation into the potential for automating the manual labeling process of VLF ionospheric data. The statistical evaluation metrics produced models that have the potential for future refinement. However, for the purposes of this research paper, and when considering each individual T-R pair, there were several instances where the researcher would likely require minimal manual adjustments to the ML classification. Conversely, certain instances arose in which the RF classification produced entirely unsatisfactory outcomes, necessitating the researcher to undertake manual data relabeling. The utilization of ML techniques for the automated classification of VLF ionospheric data has the potential to reduce the amount of time researchers spend manually labeling and excluding data. However, it also presents opportunities for further research, which should aim to acquire supplementary data from a larger number of T-R pairs, as well as data from different time sections.

This endeavor has the potential to enhance the classification efficacy of the method. Furthermore, there are several other potential strategies that can be employed to enhance this classification accuracy. For instance, exploring alternative under-sampling and over-sampling techniques, such as the Synthetic Minority Oversampling Technique (SMOTE) proposed by [35], could be beneficial. Unlike the random under-sampling method utilized in this study, SMOTE oversamples the minority class, thereby increasing the amount of data available in the training dataset. Furthermore, it is possible to evaluate various other ML classification techniques, including Support Vector Machines, K-nearest neighbors, and Artificial Neural Networks, among others. The aforementioned techniques have the potential to exert a substantial influence on the overall success rate of the method currently being investigated. Consequently, this could significantly reduce the amount of time researchers spend manually labeling and excluding erroneous data points from their datasets. Furthermore, the issue presented in this study pertained to a binary classification problem, wherein the two categories encompassed normal and anomalous data. Future research may also aim to address multi-class problems, specifically the classification of day and night VLF signals. This would involve a three-class problem, distinguishing regular and anomalous day and night signals. In conclusion, it is worth exploring the application of hybrid approaches that incorporate non-ML techniques, such as time series analyses and forecasting. These hybrid methods can be evaluated independently or in combination with ML methods, allowing for a direct comparison between ML and non-ML approaches, or the development of hybrid methodologies.

Regarding the RF method, it demonstrated its effectiveness and can be considered to be a highly favorable choice when addressing unfamiliar ML applications. In future research, it would be beneficial to conduct a comparative analysis of methods, even those deemed as “more complex” than the RF method.

5. Conclusions and Perspectives

Here, solar flares data, based on space- and ground-based observations, were investigated. As noted, the process of manually labeling solar flares data and subsequently excluding certain data is a demanding and laborious task that can be automated through the application of ML techniques. The present study employed the RF classification algorithm to effectively categorize anomalous data points, including instances of instrumentation errors and the effects of solar flares. The primary findings of this research can be succinctly summarized as follows:

- The results of the RF classification analysis indicated that the model with 100 trees performed the best overall. However, it is worth noting that models with higher numbers of trees also demonstrated satisfactory evaluation metric statistics. The reason for selecting the RF model with 100 trees was due to the preference for a simpler model with fewer hyperparameters, as this requires less computational time.
- The per-class statistics indicated that the F1-scores for all T-R pairs were higher for the non-anomalous data class, ranging from 0.69 to 0.97, meanwhile the F1-scores

for the anomalous data class were lower, ranging from 0.31 to 0.71. Furthermore, the evaluation metrics employed for the per-class analysis, specifically the recall and precision, exhibited similar characteristics, namely, lower values of the metrics for the anomalous data class. This suggested that the model exhibited a relatively weaker ability to accurately classify instances belonging to the anomalous data class compared to those belonging to the non-anomalous data class.

- Instances of effective classification were demonstrated through the NAA–Walsenburg and NPM–Walsenburg T-R pair examples. Both examples demonstrated satisfactory classifications, accompanied by satisfactory evaluation metric statistics. However, it should be noted that the T-R pair NML–Oklahoma South served as an illustrative case of poor classification. In this instance, the RF model incorrectly identified the majority of the time section as anomalous, thereby rendering the classification ineffective and requiring manual labeling.
- The presented research serves as an examination the potential of employing ML techniques for the automated labeling and/or exclusion of solar flares data.
- Future research should aim to investigate various oversampling and under-sampling techniques, as well as different ML methods, in order to determine if there is a more effective approach for this task. The RF algorithm yielded satisfactory outcomes in this context, affirming the feasibility of detecting VLF signal anomalies.
- As a perspective, this study could be used in the detection of the short-term responses of the ionosphere to gamma-ray bursts and in the analysis of pre-earthquake ionospheric anomalies, as well as in various fields of space science, where signals and data are very difficult to examine.
- Additional perspectives from this study can be found in broader applications of ML classification (or regression) methods to ionospheric VLF data. One potential application pertains to the automatic detection of daytime and nighttime signals. This involves the utilization of classification methods to address multi-class problems, which include the classification of normal daytime signals, anomalous daytime signals, nighttime signals, and transition zones related to terminators. Applying such methods and testing other use cases might yield significant insights into the issues pertaining to VLF ionospheric research and data.

Supplementary Materials: The training and testing datasets used in this study are available online at: <https://zenodo.org/record/8220971>, accessed on 7 August 2023.

Author Contributions: Conceptualization, F.A. and A.K.; writing—original draft preparation, F.A. and A.K.; writing—review and editing F.A., A.K. and V.A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

Data Availability Statement: Publicly available datasets were analyzed in this study. National Centers for Environmental Information (NCEI) Available online: <https://www.ncei.noaa.gov/>, accessed on 24 March 2023. Worldwide archive of low-frequency data and observations (WALDO) Available online: <https://waldo.world/>, accessed on 24 March 2023.

Acknowledgments: VLF data are provided by the WALDO database (<https://waldo.world>, accessed on 1 January 2023), operated jointly by the Georgia Institute of Technology and the University of Colorado Denver, using data collected from those institutions as well as Stanford University, and has been supported by various US government grants from the NSF, NASA, and the Department of Defense.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Feature Importance Analysis

Appendix A.1. Methodology

The final stage of the ML modeling phase involved conducting an analysis of the feature importance and evaluating the computational costs associated with the model. The analysis of the feature importance was conducted using the parameters of Mean Decrease in Node Accuracy (MDNA) and Total Increase in Node Purity (TINP). The MDNA parameter quantified the average reduction in the model's accuracy for each individual feature, while the TINP parameter quantified the average increase in the node purity for each individual feature in the model. In order for a feature to be deemed significant, it should maximize both the MDNA and the TINP. The ranking of the features and the analysis of the computational costs were conducted based on the scores of the MDNA and TINP. The number of features was systematically reduced from the full set of 41 to subsets of 20, 10, and 5. The duration of training and testing the model was recorded to assess the time taken. The examination of the evaluation metrics and the computational time required by the model was considered as a computational cost analysis, which involved balancing the trade-off between maximizing certain evaluation metrics and the time needed for the model training and testing. The reduction in the computational time, while maintaining similarity to the evaluation metrics of the model that used all features, was considered advantageous for future research.

Appendix A.2. Results

The analysis of the feature importance was a crucial step in refining the model, eliminating irrelevant features, and optimizing the overall performance of the model, which was derived from the RF modelling process. Here, the MDNA and TINP parameters were employed to determine the feature importance. Figure A1 presents a visual representation of the top 20 most informative features. Among the various features examined, it was found that both the MDNA and TINP yielded valuable insights. Notably, the local receiver time emerged as the most informative feature, followed by a statistical measure termed the rolling standard deviation of the VLF amplitude. This particular measure utilized the preceding 180 data points for its analysis. An additional noteworthy observation from the feature importance ranking was that the rolling window statistics comprised 60% (12 out of 20) of the top 20 most informative features. Figure A1 did not exhibit certain features, which were regarded as the least informative. The list of features with low information content included the second derivative of the X-ray signal, the binary class that exceeded the mean or median, and the first and second derivatives of the VLF amplitude and X-ray data.

The sequential RF modeling approach was employed based on the insights obtained from the feature importance analysis. This involved constructing a series of models using the same methodology as that of the previous best model, which utilized 100 trees in the RF model. However, the subsequent models were constructed with reduced numbers of features. Specifically, the models were created using the 20 most informative features (half of the original set), the 10 most informative features (a quarter of the original set), and the 5 most informative features (an eighth of the original set). The model was subsequently compared to the previous best model, which was the RF model that utilized all 41 features. Additionally, a computational cost analysis was conducted to evaluate the model's training and testing time.

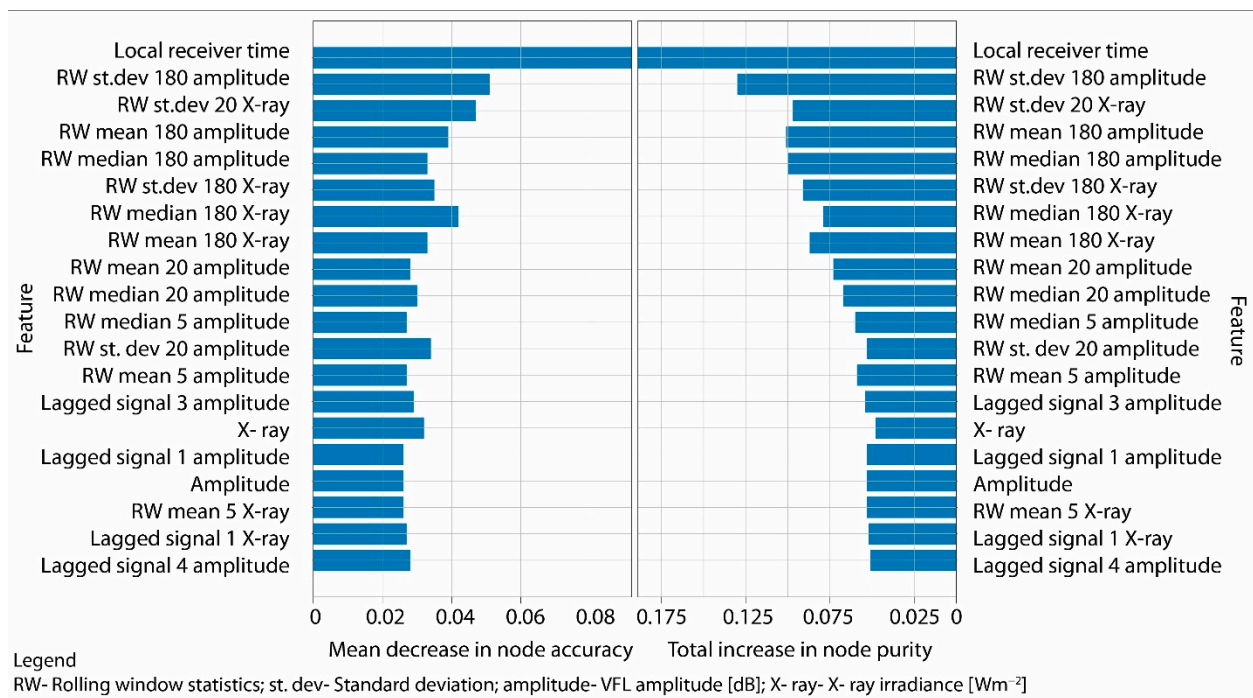


Figure A1. Top 20 most informative features from the RF modelling based on the mean decrease in node accuracy (left) and total increase in node purity (right).

Table A1 presents the model’s performance using different feature sets, including the full set of 41 features, as well as subsets of 20, 10, and 5 features. The evaluation metrics considered for each class included the F1-score, AUC, recall, false positive rate, precision, accuracy, statistical parity, and the duration of training and testing the model. In terms of the F1-score, the model that utilized all 41 features demonstrated the highest performance for the anomalous data class, achieving a value of 0.508. This observation also held true for the AUC parameter, which showed the performance of both normal and anomalous data scores. The recall values, particularly for the anomalous data display, indicated that the model employing the fewest features exhibited an unexpectedly higher recall value compared to the other models, with a value of 0.663. The relationship between the precision and accuracy metrics was not consistent in the model that used the fewest features. Specifically, the evaluation metrics for the anomalous data class showed a decreasing trend as the number of features decreased. The statistical parity parameter revealed that the model with the lowest number of features exhibited an overestimation of the anomalous data class by approximately 9.7%, representing the most significant disparity observed among all the constructed models. In general, the model that employed the least number of features may be rejected, despite exhibiting the highest recall value. This was because it significantly overestimated the anomalous data class to a greater extent compared to the other models. In light of the assessment criteria, a choice can be reached regarding the models that employed 20 and 10 features. Both models exhibited the highest difference in the F1-scores for the anomalous data class, with a value of 3.3%. In the case of the three models, the difference between the highest and lowest AUC values for both classes were only slightly above 1%. Additionally, the highest difference in the false positive rate for the anomalous data class was recorded at 1.9% between all three models. The primary distinction was observed in the training duration of the models. Specifically, the model employing the complete feature set required slightly over two minutes (124 s) to complete the training and testing processes. In contrast, the model utilizing 20 features necessitated approximately half the time (59 s). Similarly, the model employing only 10 features required approximately 34 s for the training and testing phases.

Table A1. Selected evaluation metric statistics for models utilizing different numbers of features and the Random Forest model with 100 trees.

EM/No. Features	Class	41 Features	20 Features	10 Features	5 Features
F1-score	0	0.908	0.904	0.899	0.876
	1	0.508	0.475	0.489	0.502
AUC	0	0.847	0.839	0.844	0.84
	1	0.845	0.836	0.841	0.838
Recall	0	0.902	0.899	0.883	0.827
	1	0.53	0.489	0.538	0.663
FP Rate	0	0.47	0.511	0.462	0.337
	1	0.098	0.101	0.117	0.173
Precision	0	0.915	0.909	0.915	0.933
	1	0.488	0.461	0.449	0.404
Accuracy	0	0.846	0.837	0.831	0.802
	1	0.846	0.837	0.831	0.802
Statistical Parity *	0	0.837	0.84	0.82	0.753
	1	0.163	0.16	0.18	0.247
Time	/	124	59	34	24

* The true class distributions in the test dataset were 0.85 for class 0 and 0.15 for class 1; EM—evaluation metric.

As a result, the most optimal solution was achieved by employing models that incorporated only half or a quarter of the features present in the complete set of models. The evaluation metrics presented were highly comparable, with minimal compromises made in terms of statistical integrity, while significantly reducing the time required for the training and testing of the model (by 50% or 72%). In this particular instance, the disparity did not result in a significant time reduction in absolute terms (from approximately 65 to 90 s). However, when working with a substantially larger volume of data, this discrepancy could lead to a more substantial time reduction, consequently reducing the computational expenses.

References

- Barta, V.; Natras, R.; Srećković, V.; Koronczay, D.; Schmidt, M.; Šulic, D. Multi-instrumental investigation of the solar flares impact on the ionosphere on 05–06 December 2006. *Front. Environ. Sci.* **2022**, *10*, 904335. [\[CrossRef\]](#)
- Kolarski, A.; Veselinović, N.; Srećković, V.A.; Mijić, Z.; Savić, M.; Dragić, A. Impacts of Extreme Space Weather Events on September 6th, 2017 on Ionosphere and Primary Cosmic Rays. *Remote Sens.* **2023**, *15*, 1403. [\[CrossRef\]](#)
- Grubor, D.P.; Šulić, D.M.; Žigman, V. Classification of X-ray Solar Flares Regarding Their Effects on the Lower Ionosphere Electron Density Profile. *Ann. Geophys.* **2008**, *26*, 1731–1740. [\[CrossRef\]](#)
- Kolarski, A.; Srećković, V.A.; Mijić, Z.R. Response of the Earth's Lower Ionosphere to Solar Flares and Lightning-Induced Electron Precipitation Events by Analysis of VLF Signals: Similarities and Differences. *Appl. Sci.* **2022**, *12*, 582. [\[CrossRef\]](#)
- Miteva, R.; Samwel, S.W. M-Class Solar Flares in Solar Cycles 23 and 24: Properties and Space Weather Relevance. *Universe* **2022**, *8*, 39. [\[CrossRef\]](#)
- Kahler, S.W. The Role of the Big Flare Syndrome in Correlations of Solar Energetic Proton Fluxes and Associated Microwave Burst Parameters. *J. Geophys. Res.* **1982**, *87*, 3439. [\[CrossRef\]](#)
- Srećković, V.A.; Šulić, D.M.; Vujčić, V.; Mijić, Z.R.; Ignjatović, L.M. Novel Modelling Approach for Obtaining the Parameters of Low Ionosphere under Extreme Radiation in X-Spectral Range. *Appl. Sci.* **2021**, *11*, 11574. [\[CrossRef\]](#)
- Wang, J.; Huang, Q.; Ma, Q.; Chang, S.; He, J.; Wang, H.; Zhou, X.; Xiao, F.; Gao, C. Classification of VLF/LF Lightning Signals Using Sensors and Deep Learning Methods. *Sensors* **2020**, *20*, 1030. [\[CrossRef\]](#)
- Sigillito, V.; Wing, S.; Hutton, L.; Baker, K. Classification of Radar Returns from the Ionosphere Using Neural Networks. *Johns Hopkins APL Tech. Dig.* **1989**, *10*, 262–266.
- Dhande, J.; Dandekar, D.R. PSO Based SVM as an Optimal Classifier for Classification of Radar Returns from Ionosphere. *Int. J. Emerg. Technol.* **2011**, *2*, 1–3.
- Oo, A.N. Classification of Radar Returns from Ionosphere Using NB-Tree and CFS. *Int. J. Trend Sci. Res. Dev.* **2018**, *2*, 1640–1642. [\[CrossRef\]](#)
- Ameer Basha, G.; Lakshmana Gupta, K.; Ramakrishna, K. Expectation of Radar Returns from Ionosphere Using Decision Tree Technique. In *Advances in Data Science and Management*; Springer Nature: Singapore, 2020; pp. 209–214. [\[CrossRef\]](#)
- Adhikari, S.; Thapa, S.; Shah, B.K. Oversampling Based Classifiers for Categorization of Radar Returns from the Ionosphere. In *Proceedings of the 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, 2–4 July 2020. [\[CrossRef\]](#)

14. Shang, Z.; Yao, Z.; Liu, J.; Xu, L.; Xu, Y.; Zhang, B.; Guo, R.; Wei, Y. Automated Classification of Auroral Images with Deep Neural Networks. *Universe* **2023**, *9*, 96. [[CrossRef](#)]
15. Lian, J.; Liu, T.; Zhou, Y. Aurora Classification in All-Sky Images via CNN–Transformer. *Universe* **2023**, *9*, 230. [[CrossRef](#)]
16. National Centers for Environmental Information (NCEI). Available online: <https://www.ncei.noaa.gov/> (accessed on 24 March 2023).
17. Worldwide Archive of Low-Frequency Data and Observations (WALDO). Available online: <https://waldo.world/> (accessed on 24 March 2023).
18. JASP—A Fresh Way to Do Statistics. Available online: <https://jasp-stats.org/> (accessed on 1 April 2023).
19. Prusa, J.; Khoshgoftaar, T.M.; Dittman, D.J.; Napolitano, A. Using Random Undersampling to Alleviate Class Imbalance on Tweet Sentiment Data. In Proceedings of the 2015 IEEE International Conference on Information Reuse and Integration, San Francisco, CA, USA, 13–15 August 2015. [[CrossRef](#)]
20. Kulkarni, A.; Chong, D.; Batarseh, F.A. Foundations of Data Imbalance and Solutions for a Data Democracy. In *Data Democracy*; Academic Press: Cambridge, MA, USA, 2020; pp. 83–106. [[CrossRef](#)]
21. Devi, D.; Biswas, S.K.; Purkayastha, B. A Review on Solution to Class Imbalance Problem: Undersampling Approaches. In Proceedings of the 2020 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2–4 July 2020. [[CrossRef](#)]
22. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *ACM SIGKDD Explor. Newsl.* **2004**, *6*, 20–29. [[CrossRef](#)]
23. Hasanin, T.; Khoshgoftaar, T. The Effects of Random Undersampling with Simulated Class Imbalance for Big Data. In Proceedings of the 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, 6–9 July 2018. [[CrossRef](#)]
24. Saripuddin, M.; Suliman, A.; Syarmila Sameon, S.; Jorgensen, B.N. Random Undersampling on Imbalance Time Series Data for Anomaly Detection. In Proceedings of the 2021 the 4th International Conference on Machine Learning and Machine Intelligence, Virtual, 1–3 December 2021. [[CrossRef](#)]
25. Mishra, S. Handling Imbalanced Data: SMOTE vs. Random Undersampling. *Int. Res. J. Eng. Technol.* **2017**, *4*, 317–320.
26. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
27. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random Forests for Classification in Ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)]
28. Hossin, M.; Sulaimani, M.N. A Review on Evaluation Metrics for Data Classification Evaluations. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1–11. [[CrossRef](#)]
29. Hand, D.; Till, R. A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Mach. Learn.* **2001**, *45*, 171–186. [[CrossRef](#)]
30. Jin, H.; Ling, C.X. Using AUC and Accuracy in Evaluating Learning Algorithms. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 299–310. [[CrossRef](#)]
31. Rosset, S. Model Selection via the AUC. In Proceedings of the Twenty-First International Conference on Machine Learning—ICML, Banff, AB, Canada, 4–8 July 2004. [[CrossRef](#)]
32. Joshi, M.V. On Evaluating Performance of Classifiers for Rare Classes. In Proceedings of the 2002 IEEE International Conference on Data Mining, Maebashi City, Japan, 9–12 December 2002. [[CrossRef](#)]
33. Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genom.* **2020**, *21*, 6. [[CrossRef](#)] [[PubMed](#)]
34. Chicco, D.; Tötsch, N.; Jurman, G. The Matthews Correlation Coefficient (MCC) Is More Reliable than Balanced Accuracy, Bookmaker Informedness, and Markedness in Two-Class Confusion Matrix Evaluation. *BioData Min.* **2021**, *14*, 13. [[CrossRef](#)]
35. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-Sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Application of geophysical and multispectral imagery data for predictive mapping of a complex geo-tectonic unit: a case study of the East Vardar Ophiolite Zone, North-Macedonia

Filip Arnaut¹ · Dragana Đurić² · Uroš Đurić³ · Mileva Samardžić-Petrović³ · Igor Peshevski⁴

Received: 12 July 2023 / Accepted: 28 January 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

The Random Forest (RF) and K nearest neighbors (KNN) machine learning (ML) algorithms were evaluated for their ability to predict ophiolite occurrences, in the East Vardar Zone (EVZ) of central North Macedonia. A predictive map of the investigated area was created using three data sources: geophysical data (digital elevation model, gravity and geomagnetic), multispectral optical satellite images (Landsat 7 ETM+ and their derivatives), and geological data (distance to fault map and ophiolite outcrops map). The research included a comparison and discussion on the statistical and geological findings derived from different training dataset class ratios in relation to a testing dataset characterized by significant class imbalance. The results suggest that the precise selection of a suitable class balance for the training dataset is a critical factor in achieving accurate ophiolite prediction with RF and KNN algorithms. The analysis of feature importance revealed that the Bouguer gravity anomaly map, total intensity of the Earth's magnetic field reduced to the pole map, distance to fault map, band ratio BR3 map obtained from multispectral satellite images, and digital elevation model are the most significant features for predicting ophiolites within the EVZ. KNN showed poorer results compared to RF in terms of both the evaluation metrics and visual analysis of prediction maps. The methods applied in this research can be applied for predictive mapping of complex geo-tectonic units covered by dense vegetation, and may indicate the presence of these units even if they were not previously mapped, particularly when geophysical data are used as features.

Keywords Random Forest · K nearest neighbors · Remote Sensing · Geophysical data · Predictive mapping · East Vardar Zone

Introduction

In machine learning algorithms (ML), the automatic inductive approach was used to recognize patterns in data, and the learned pattern relationships were then applied to other

similar data or the same datasets but in different domains to generate predictions for data-driven classification and regression problems (Cracknell and Reading 2014). In situations involving the prediction of spatially dispersed categories in extremely complex processes, these algorithms have proven to be immensely useful (Kanevski et al. 2009). The Random Forest (RF) algorithm is widely used for predictive

Communicated by H. Babaie.

✉ Filip Arnaut
filip.arnaut@ipb.ac.rs

Dragana Đurić
dragana.djuric@rgf.bg.ac.rs

Uroš Đurić
udjuric@grf.bg.ac.rs

Mileva Samardžić-Petrović
mimas@grf.bg.ac.rs

Igor Peshevski
pesevski@gf.ukim.edu.mk

¹ University of Belgrade, Institute of Physics Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

² University of Belgrade, Faculty of Mining and Geology, Dušina 7, 11000 Belgrade, Serbia

³ University of Belgrade, Faculty of Civil Engineering, Bulevar Kralja Aleksandra 73/1, 11000 Belgrade, Serbia

⁴ Ss. Cyril and Methodius University in Skopje, Faculty of Civil Engineering MK, Bulevar Partizanski Odredi 24, 1000 Skopje, North Macedonia

or geological mapping (Zuo and Carranza 2023). The mineral prospectivity mapping studies conducted in the Philippines demonstrated that the RF model exhibited superior performance compared to other models in predicting the presence of gold deposits (Carranza and Laborte 2015a). This was also observed in cases where there were only a few occurrences of minerals, such as in the predictive mapping of porphyry Cu deposits (Carranza and Laborte 2015b). Furthermore, the RF method offers a notable advantage in the form of feature importance quantification. This quantification, as noted by Rodriguez-Galiano et al. (2014), provides crucial information that aligns with geological expectations. Machine learning methods have demonstrated their efficacy as a valuable tool in geological mapping, because of that, applications in ophiolite mapping could provide satisfactory results.

Remote sensing techniques and satellite imagery, characterized by different spatial and spectral resolutions, have been extensively studied for their application in geological mapping through the analysis and processing of digital images (Cracknell and Reading 2014). Multispectral satellite data, such as Landsat 7 ETM+, Landsat 8, Landsat 9, or Sentinel-2, have been effectively employed for mapping lithology and geological structures in well-exposed geological areas where spectral analysis is optimal (Aliyu et al. 2021; Lorenz 2004; Harris et al. 2005, 2008, 2009, 2014; Schetselaar and Ryan 2008; Leverington 2010; Leverington and Moon 2012; Behnia et al. 2012; Harris and Grunsky 2015; Albert and Ammar 2021). While there were positive outcomes in the application of remote sensing for lithological mapping of surface occurrences, challenges arise in characterizing geological materials in areas covered with dense vegetation or sediments (Leverington and Moon 2012; Harris and Grunsky 2015; Kuhn et al. 2018; Ge et al. 2022).

In contrast to prior studies that focused on predicting land cover or vegetation classes using multispectral imagery and machine learning (ML) algorithms (Huang et al. 2002; Foody and Mathur 2004; Ham et al. 2005; Waske and Braun 2009; Cracknell and Reading 2014), this study emphasizes the use of ML methods to map complex geological units, irrespective of whether they are exposed outcrops or not. The approach employed in this study involved utilizing different data sets, considering their level, resolution, and physical properties (Kuhn et al. 2018), including airborne magnetic data, terrestrial gravity data, and multispectral satellite data (Landsat 7 ETM+). These data sets underwent ML algorithm processing to generate continuous surface raster maps of ophiolite complex data, predicting and mapping the complex geological unit—the East Vardar ophiolite zone.

Geophysical data derived from measured values were utilized to assess the spatial distribution of properties, primarily for phenomena not observable at the surface, as commonly done in geophysics (Cracknell and Reading 2014).

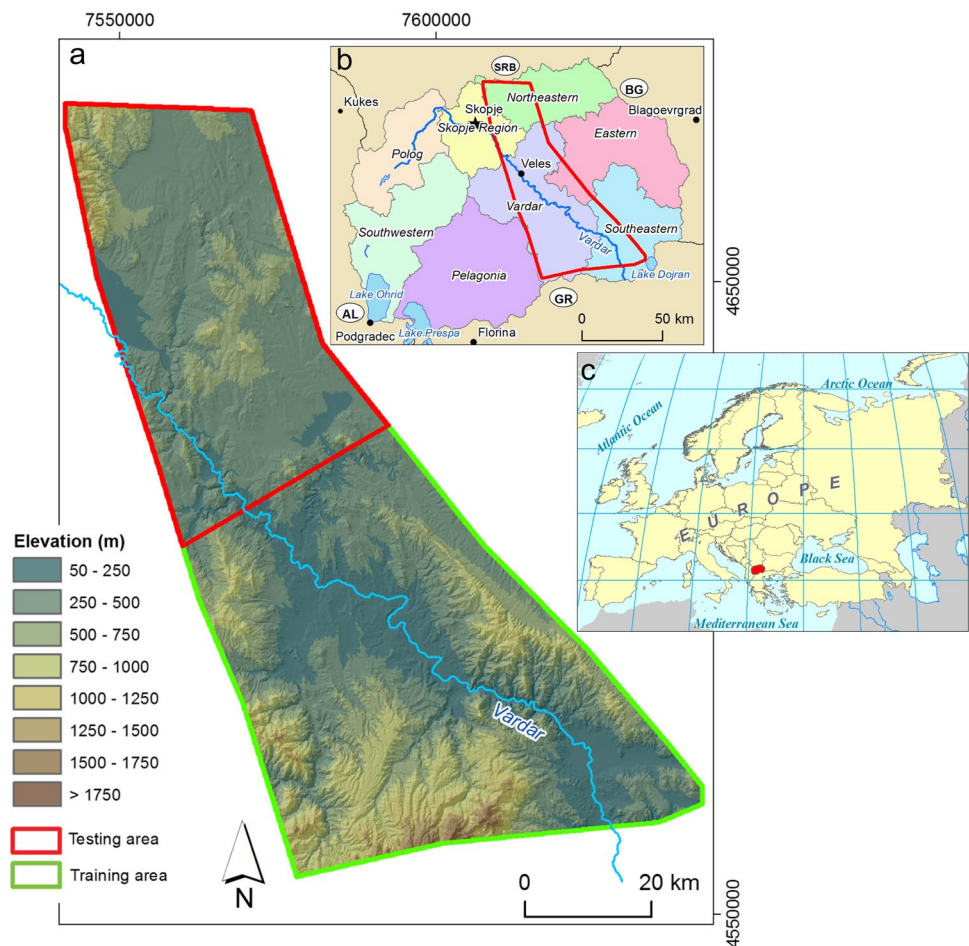
Physical parameters in geophysics are typically associated with specific lithologies. The data used in this application included multiple remotely sensed data sets, encompassing airborne geophysics (magnetics and elevation), terrestrial geophysics (gravity data), and satellite multispectral data. The primary task involved a high-dimensional input space and a complex, variable set of data relationships. All data were represented as raster data and used in the ML algorithm, regardless of the values they represented.

The evaluated algorithms include RF and K nearest neighbors (KNN). According to Cracknell and Reading (2014) and their cited sources, the architecture of ML algorithms and the statistical distributions of data guide the training of classification models (typically by minimizing a loss function), which are then applied to the same input variables to predict classes present in the training data. The emphasis was on comparing the categorical lithology predictions that were generated by the ML algorithms RF and KNN with various data. In this study, both RF and KNN were used to combine satellite multispectral data with geophysical airborne and terrestrial data, as well as the distance to fault map, as features. Incorporating these data, particularly geophysical data (airborne and terrestrial), into ML predictions has not yet been realized, particularly for mapping complex lithological units as single areas of interest (Cracknell and Reading 2014). This research was focused on the improvement of a set of features (widely accessible remote sensing data supplemented with geophysical data and tectonic settings; distance to fault map) with the RF and KNN algorithms in areas with dense vegetation and complex geological units below the surface. An added element of this study pertains to the utilization of legacy geological maps as the basis for ophiolite test data (Pendžerkovski et al. 1963; Rakicević et al. 1965, 1969, 1973; Ivanovski and Rakicević 1970; Hristov et al. 1965, 1973; Karajovanovic and Hristov 1976; Dumurdzanov et al. 1981; Karajovanovic and Hadži-Mitrova 1982), which were created using conventional geological mapping methods i.e., surface mapping techniques. The utilization of RF and KNN algorithms presents the potential for predicting ophiolites that may not be noticeable at the terrain surface, specifically those which are situated in the subsurface and that cannot be mapped using conventional geological mapping methods. The study presented a ML task that involved binary classification, which entails the prediction of one of two classes: the non-ophiolite class (designated as 0) and the ophiolite class (designated as 1).

Study area

The study area is in the central part of North Macedonia, spans 5270 km² from north to south (Fig. 1), and contains the East Vardar Zone (EVZ). In general, the EVZ traverses

Fig. 1 **a)** Digital elevation model (DEM) of study area with highlighted testing and training area; **(b)** Geographical position of study area (red polygon) within North Macedonia; **(c)** The geographical position of North Macedonia



Romania, Serbia, North Macedonia, Greece, and Turkey (from north to south) (Petrović et al. 2015) and represents the easternmost part of the extremely complex tectonic unit in the central Balkan Peninsula, namely the Vardar Zone (VZ).

As one of several subparallel and NNW- SSE stretching belts that resulted from Mesozoic convergence and subsequent collision, the Vardar Zone is located between the Drina- Ivanjica Unit and the Serbo-Macedonian Massif (SMM). As an NNW- SSE extending belt, EVZ is generally situated between Kopaonik Unit (to the west) and SMM (to the east) (Fig. 2). The presence of the ophiolite complex is a defining feature of EVZ. Ophiolites are portions of the oceanic crust with middle ocean ridge (MOR) or supra-subduction zone (SSZ) affinity that were tectonically deposited on the continental margin during the closure of the ocean or the exhumation process. According to Moore's (1982) classification, EVZ ophiolites belong to the Cordilleran type of ophiolite, which are spatially and temporally associated with island and volcanic arches, pyroclastic rocks, subduction mechanism, and the accretionary mechanism (Petrović et al. 2015). EVZ ophiolites are a complex unit, composed of ingenious rocks (gabbro-dolerites, dolerites, dolerite dykes,

and rare basaltic pillow lavas) with geochemical and tectonic affinity to "island arc tholeiite" and MOR. Occasionally, serpentized harzburgites are found in small numbers at the EVZ margins. All ophiolitic rocks are found in association with intermediate and acid-calcalkaline granitic rocks (Resimić-Šarić et al. 2006). Overlying ophiolites in the study region are Tithonian limestone overstep sequence and/or Cretaceous sedimentary rocks (Dimitrijević 1997).

The western part of the study area corresponds to Kopaonik Unit (KU), also known as Paikon Unit in Greece (Đurić, 2023). The primary lithological components of the KU are shales, phyllites with quartzite, limestone, spilite, and carbonates from the upper Triassic. Part of the eastern portion of the KU is covered by unaltered Lower Cretaceous para-flysch (Schefer et al. 2007; Zelić et al. 2010). The eastern part of the study area consists of SMM metamorphic rocks, which are typically the oldest units in the study area: Precambrian and Paleozoic metamorphites. To obtain a comprehensive synthesis of the geological evolution of the EVZ, it is recommended to refer to the work of Boev et al. (2018).

To simplify the geological analysis and produce a continuous raster, the data used in this study were obtained

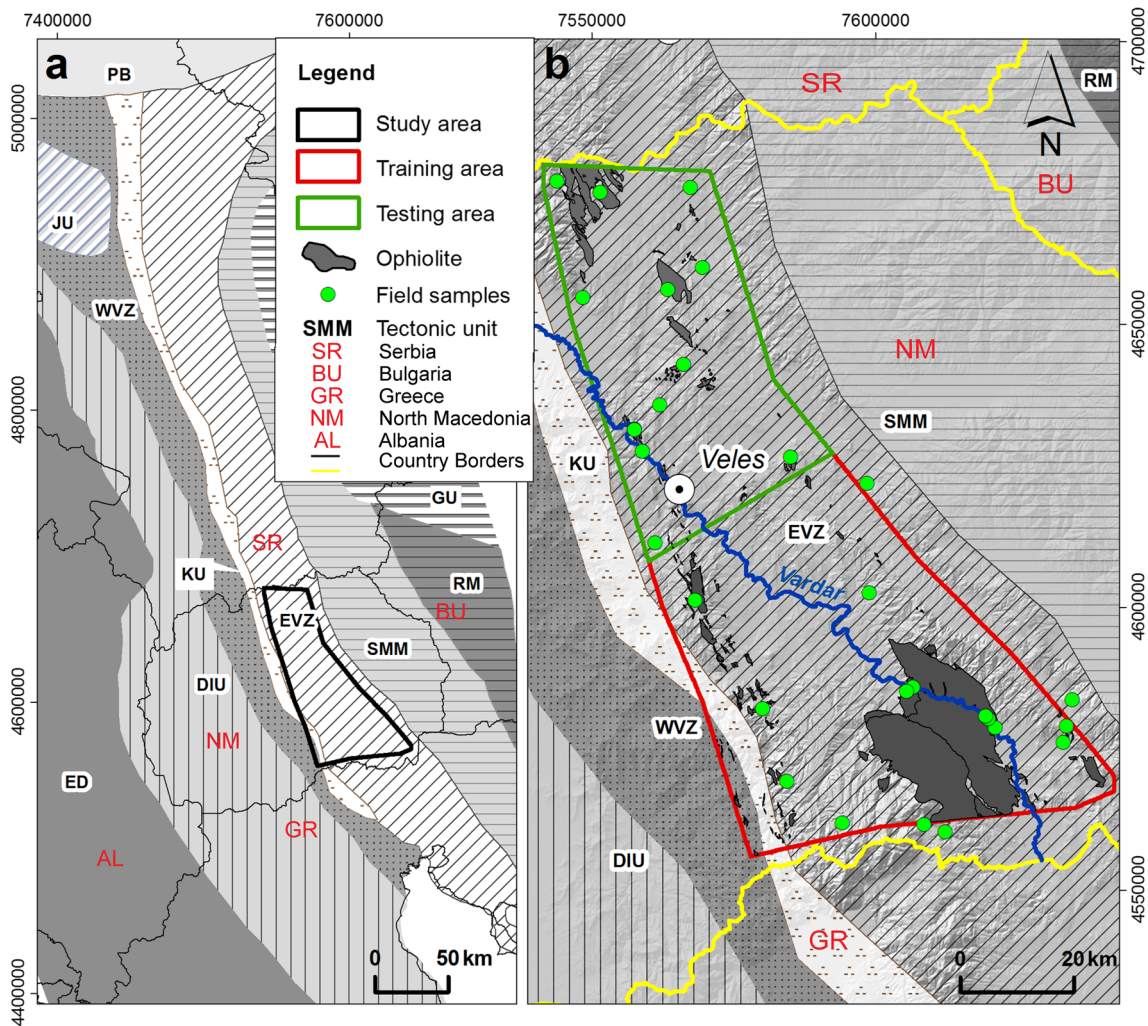


Fig. 2 a) Simplified geotectonic map of part of the Balkan Peninsula (modified from Robertson et al. 2009); b) Simplified geotectonic-geological map of the study area with testing and training areas outlined; Tectonic units: **ED**-External Dinaric Unit, **WVZ**- west Vardar Zone,

KU- Kopaonik Unit (Kopaonik Block and Ridge Unit), **EVZ**- East Vardar Zone, **JU**- Jadar Unit, **SMM**- Serbo-Macedonian Massif, **RM**- Rhodope Massif, **GU**- Getic Units, **PB**- Pannonian Block

from basic geological maps. The study grouped dolerite, gabbro, gabbro-dolerite, basaltic pillow lavas, and serpentinized harzburgites as a single ophiolite unit (as shown in Fig. 2), while all other geological units were grouped as non-ophiolite units.

Materials and methods

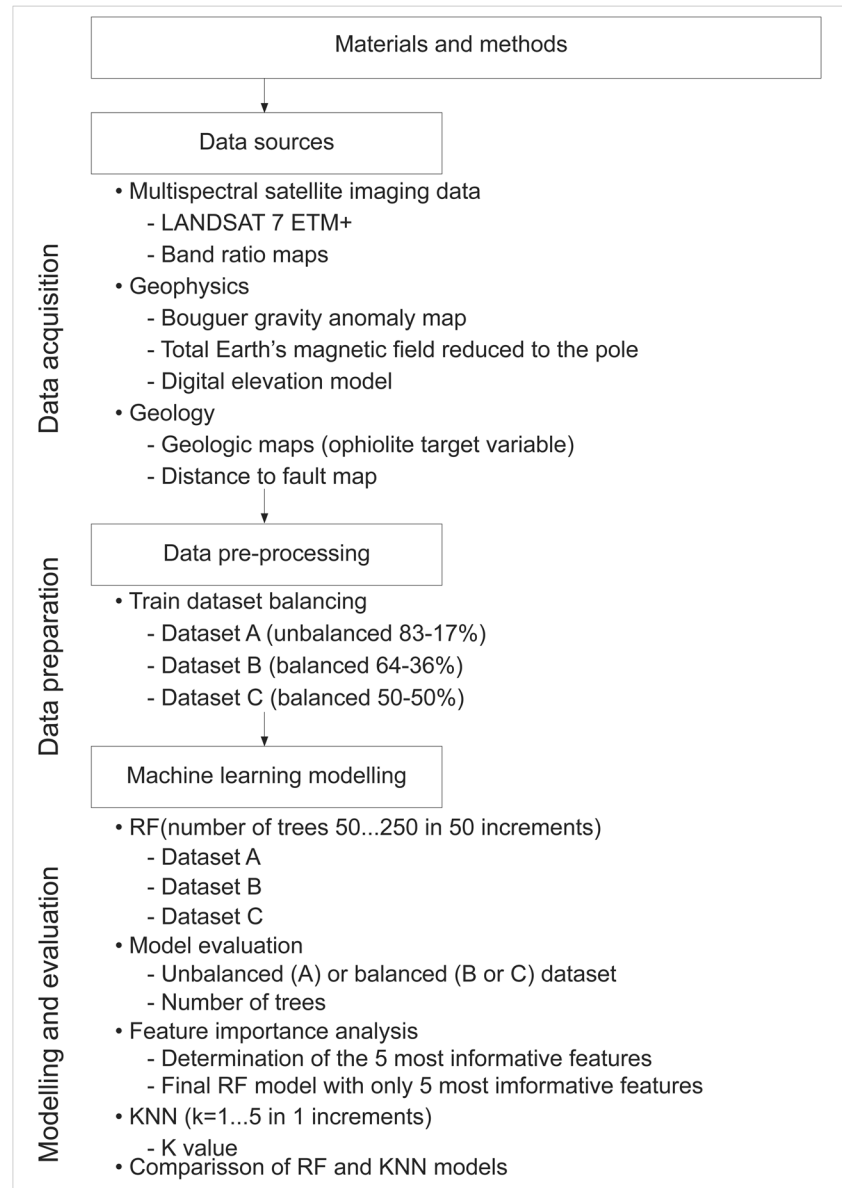
The materials and methods chapter, and the results and discussion chapter, follow Fig. 3, which displays the research workflow. Data sources, including multispectral satellite imaging, geophysical data, and geological data, are described in the materials and methods chapter. Data pre-processing is an essential step in preparing data for ML tasks, and it is considered a distinct result as it involves

transforming the original data into data sets for specific purposes in ML modelling (in this case displaying the different results with different class balances) and it is presented in the results section.

Data sources

A predictive map was generated for ophiolites in the study area by utilizing three distinct data sources. These sources include multispectral optical satellite images (Landsat 7 ETM+ and their derivatives), geophysical data (including gravity, geomagnetic, and digital elevation model), and geological data (such as the distance to fault map and field mapped ophiolite localities obtained from legacy geological maps).

Fig. 3 Data acquisition, preparation, machine learning modelling and evaluation workflow



Multispectral satellite imaging data

The Landsat 7 ETM+ satellite images of the investigated area (30 m resolution), bands 1, 2, 3, 4, 5 and 7 were used to predict distinct classes in according due to their established track record of delivering reliable and robust results in similar applications, as evidenced by previous studies (Al-Rawashdeh et al. 2006). The Landsat 7 satellite image underwent essential preprocessing steps, including geometric correction, atmospheric correction, and noise elimination. Additionally, basic processing procedures such as filtering, illumination enhancement, and contrast improvement were applied to enhance the overall quality and interpretability of the image. Band ratios, which are widely used for lithological mapping (Bolt and Bruggenwert 1976; Sposito 1989;

Farrand 1997; Longhi et al. 2001; Akhavi et al. 2001; Neville et al. 2003; Al-Rawashdeh et al. 2006), was also derived from Landsat 7 images using the previously mentioned ratios 3/1 (BR1), 5/4 (BR2), and 5/7 (BR3), chosen based on the spectral characteristics of the ophiolite complex. Table 1 provides the corresponding wavelengths for each band of the Landsat 7 ETM+ satellite.

Geophysical data

In addition to the DEM (Fig. 1c) that was obtained from the ASTER mission (30 m resolution), the Bouguer i.e., gravity anomaly map (BAM) and the total intensity of the Earth's magnetic field anomaly map, reduced to the pole

(RTP), were selected as geophysical input parameters i.e., features for predictive lithological mapping.

Gravity data (Bilibajkić et al. 1979) were obtained by detailed gravimetric surveys conducted between 1952 and 1984 by the Geophysical Institute in Serbia and Macedonia. The Cassinis Formula (1930) was utilized to calculate normal gravity values. Bullard A, a simple Bouguer correction, and free-air corrections of the first order were applied. The Bouguer anomaly map (Fig. 4a) was generated utilizing a density of 2.67 g/cm^3 , whereas the measurement density was approximately 1 point/km^2 .

From geomagnetic data (surveys in Serbia and North Macedonia), the total intensity of the Earth's magnetic field anomaly map (Fig. 4b) was calculated. The normal magnetic field was computed using the Geomagnetic Institute's (Grocka) formula for the epoch 1960.0. RTP was applied to these data, resulting in the Total Intensity EMF Anomaly Map being reduced to the pole, considering the Koenigsberger ratio (Q ratio) within the study area (29 localities shown on Fig. 2), which does not exceed the value of 0.001, indicating that the effects of the remanent magnetization can be disregarded (Petrović 2015; Cvetkov et al. 2016).

Geologic data

The published 1:100,000 SFR Yugoslavia Basic Geologic Maps (Pendžerkovski et al. 1963; Rakikevik et al. 1965, 1969, 1973; Ivanovski and Rakicević 1966; Hristov et al. 1965, 1973; Karajovanovic and Hristov 1976; Dumurdzanov et al. 1981; Dimitrijević, 1978; Karajovanovic and Hadži-Mitrova 1982) were used to obtain labelled samples (manually annotated i.e., classified instances of ophiolites) representing ophiolite classes for training and evaluation of RF and KNN predictions.

The strategic placement of ophiolite is intricately linked to the Mesozoic tectonic settings of the region. Moreover, the discernible regional fault pattern plays a pivotal role in this geological context, underscoring its profound significance. Using the same geological map sheets, data obtained by fieldwork, all faults were digitized and subsequently interpreted. Employing a well-established common criterion (Brockmann et al. 1977; Novak and Soulakellis 2000) based on satellite images (Fig. 5a), we determined these faults as regional faults that could be responsible for ophiolite emplacement within the area. The distance-to-fault raster is a map that shows the distance of each pixel in the study area to the nearest fault. In this case, it was created using a fault map and the Euclidian Distance tool (ArcGIS® 10.1). The fault map provides information about the location

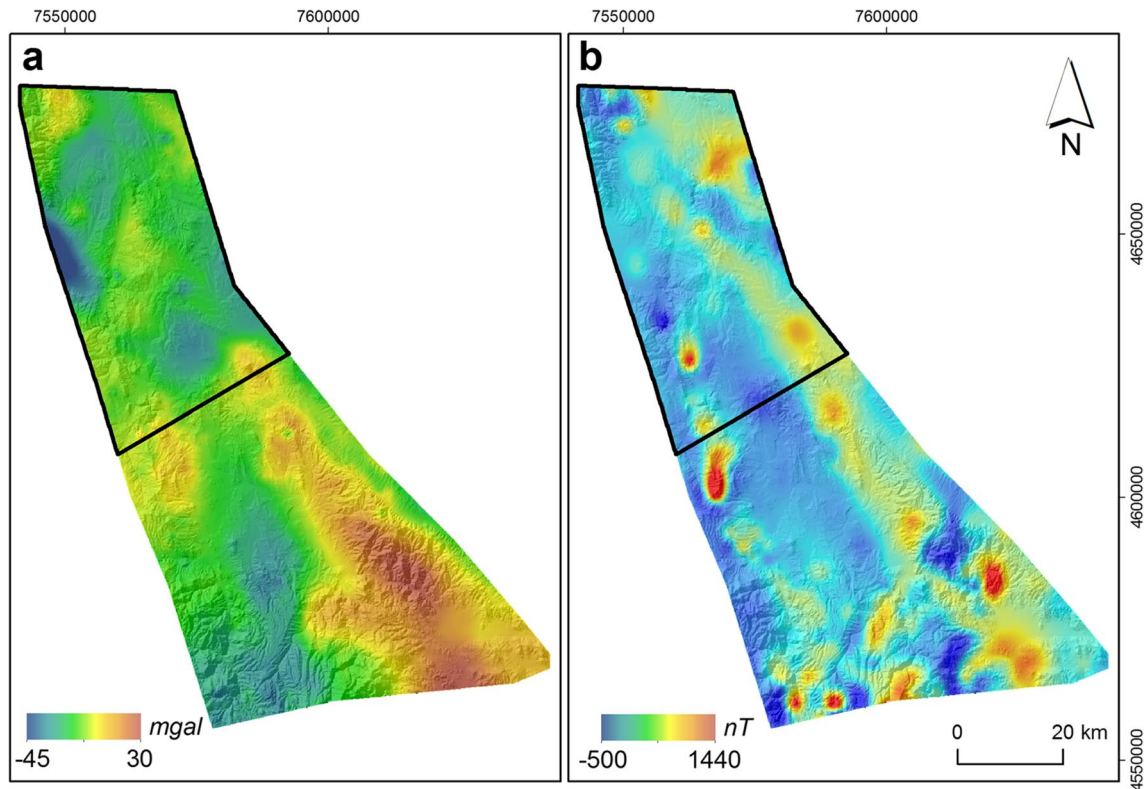


Fig. 4 Geophysical maps; **a**) Bouguer Anomaly Map; **b**) Total Intensity EMF Anomaly Map reduced to the pole; Black outline- testing area

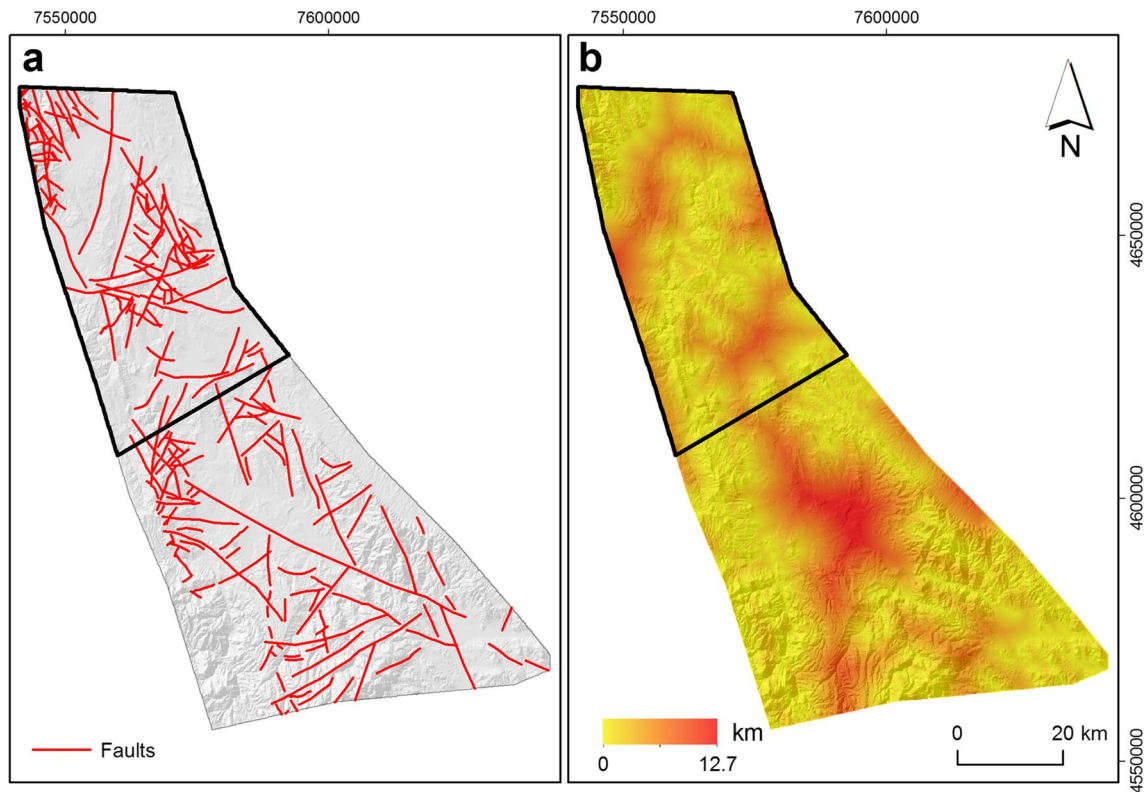


Fig. 5 a) Fault map; b) Distance to fault map; Black outline- testing area

and extent of known faults in the area, while the Euclidian Distance tool calculates the straight-line distance from each pixel to the nearest fault. The resulting raster map can be used as a feature in ML algorithms to help predict the distribution of geological units or other phenomena in the study area, considering the influence of nearby faults.

All the obtained data was transformed to the MGI Balkans Zone 7 projected coordinate system. The target variable, namely the ophiolite map, was characterized as a binary categorical class with two distinct values (0 for non-ophiolites and 1 for ophiolites). Table 1 provides a brief overview of the data utilized in this study, including the corresponding

Table 1 Utilized data, abbreviations, groups, and types used in this research

Parameter	Symbol	Data group	Data type	Usage	Other
Landsat 7 ETM + channel 1	C1	Satellite	Integer	Feature	0.45–0.52 μm
Landsat 7 ETM + channel 2	C2	Satellite	Integer	Feature	0.52–0.60 μm
Landsat 7 ETM + channel 3	C3	Satellite	Integer	Feature	0.63–0.69 μm
Landsat 7 ETM + channel 4	C4	Satellite	Integer	Feature	0.77–0.90 μm
Landsat 7 ETM + channel 5	C5	Satellite	Integer	Feature	1.55–1.75 μm
Landsat 7 ETM + channel 7	C7	Satellite	Integer	Feature	2.09–2.35 μm
Band ratio 1	BR1	Satellite	Float	Feature	C3/C1
Band ratio 2	BR2	Satellite	Float	Feature	C5/C4
Band ratio 3	BR3	Satellite	Float	Feature	C5/C7
Digital elevation model	DEM	Geophysics	Integer	Feature	/
Bouguer anomaly map	BAM	Geophysics	Float	Feature	/
Total intensity of the EMF reduced to the pole	RTP	Geophysics	Float	Feature	/
Distance to fault map	DF	Geology	Float	Feature	/
Ophiolite map	Ophiolites	Geology	Categorical	Target	/

abbreviations i.e., symbols, data groupings, and data types employed throughout the research.

Machine learning algorithms

The random forest algorithm

Breiman (2001) proposed and developed the Random Forest (RF) ML algorithm for regression and classification, which has widespread applications and popularity. It has been utilized by numerous environmentalists and geoscientists over the past two decades to predict classes and generate spatial data maps. It is a meta-classifier (estimator) that fits several user-defined decision tree classifiers on various subsamples of the training dataset. RF uses averaging to enhance predictive accuracy and prevent overfitting (Kuncheva 2004; Cutler et al. 2007; Youssef et al. 2016).

RF has gained popularity among researchers in various fields both including science and industry for classification and regression tasks. According to Puissant et al. (2014), when compared to other ML techniques, only a few parameters (the number of variables per node and the number of random trees) must be set by the end user, making data preparation easier and saving time during data processing. Since it is a non-parametric method, the data is less dependent on regular statistical (grid) distribution. According to Breiman (2001), it is not susceptible to noise or overfitting (unlike standard random trees) and is relatively fast in comparison to other classification methods, such as boosting techniques. In addition, the RF algorithm is available as a package within most of the popular open-source statistical software packages (R, Weka or JASP). According to Cracknell (2014), RF should be considered a first-choice algorithm for the prediction of classes that represent lithologies (in this case, ophiolite) using commonly available geological and geophysical data, and this was the primary reason for using this algorithm to predict ophiolite.

Tuning and validation are essential components of all ML-based models, including the RF algorithm. According to previous research (Breiman 2001; Catani et al. 2013; Harris and Grunsky 2015), the most effective method for determining the optimal range of the number of trees ($n^{\circ}T$) for an algorithm, is to run a certain (user-defined) number of model performance tests with sequentially increasing $n^{\circ}T$, until the Out-of-bag (OOB) error stops decreasing and stabilizes at a near value. The number of trees was incrementally increased from 50 to 250 with a step of 50 trees. Subsequently, all results were graphed on the test area and, alongside the evaluation metrics, were interpreted and discussed. The process of conducting ML classification on spatial data involves a crucial step of visual analysis. In this study, the occurrence of false positive predictions may suggest the existence of ophiolites that were not detected using traditional surface

mapping methodologies due to their subsurface location. For the false positive values to possess geological plausibility, it is crucial that they are situated in a region that is consistent with existing geological features, such as a previously identified and mapped ophiolite belt. Moreover, the EVZ ophiolites exhibit lithological and tectonic complexity, resulting in spatially varying compositions. All models were plotted and compared to a digitized ophiolite map derived from legacy geologic maps to validate the models visually.

K nearest neighbors

The K nearest neighbors (KNN) algorithm was developed in the middle of the twentieth century (Fix and Hodges 1951; Cover and Hart 1967). Since then, it has become a popular and simple ML algorithm that has been implemented in classification problems across many research domains.

The KNN classification algorithm is a non-parametric algorithm (Ge et al. 2018) that finds the K closest objects in the train data from which it predicts the test data label (Steinbach and Tan 2009). The choice of the K parameter has a significant impact on the KNN classification (James et al. 2013), and a high K value can result in overfitting (Cracknell and Reading 2014). For this study, the JASP software was utilized for both RF and KNN modelling.

Comparative studies employing both RF and KNN on remotely sensed geo-spatial data demonstrate that the RF algorithm outperforms the KNN algorithm (Cracknell and Reading 2014; Ge et al. 2020; Shayeganpour et al. 2021; Bachri et al. 2022).

The classification evaluation metrics used in this study encompassed standard ML classification evaluation metrics, such as accuracy, precision, F1-Score, area under the receiver operating characteristic curve (AUC), rates derived from the confusion matrix (true positive rate—TP, false positive rate—FP, true negative rate—TN, and false negative rate—FN), and statistical parity. Metrics such as accuracy and precision are used to measure the proportion of correctly classified instances out of the total number of instances (accuracy) or the proportion of correct positive predictions out of the total number of positive predictions (precision). The F1-Score is a metric used to evaluate classification performance and it is calculated as the harmonic mean between the true positive rate and precision. The F1-Score is considered a better metric than accuracy in imbalanced ML tasks (Joshi 2002; Hossin and Sulaiman 2015). The AUC value quantifies the model's capacity to differentiate between the classes in a classification ML task, such as its ability to differentiate between the ophiolite and non-ophiolite classes in this research. The statistical parity parameter indicates the distribution of predicted instances. The previously mentioned evaluation metrics (accuracy, precision and F1 score) should for a better model be represented by higher values,

as for the statistical parity metric it should closely resemble the original class distribution in the testing dataset. All the aforementioned evaluation metrics were used individually for each class to meet specific requirements in model evaluation.

The computational cost of the employed algorithms is an additional factor to be considered in classification modeling. Cracknell and Reading (2014) demonstrated that among the five algorithms presented, the support vector machine (SVM) algorithm exhibited the highest computational time, while the RF and artificial neural networks (ANN) algorithms shared the second position. A prevalent strategy for reducing computational time involves the reduction of the number of features utilized, with the least informative ones being disregarded. The present study yields feature selection tables through RF modeling, which are distinguished by two parameters: the mean decrease in node accuracy (MDNA) and the total increase in node purity (TINP). The results of the feature selection process i.e., feature importance analysis can possibly suggest that it is beneficial to construct an alternative RF model that employs a reduced number of features. This approach would yield a decrease in computational time and a comparison of the two models with different number of features utilized.

Data pre- processing

The research area was partitioned into two distinct areas, namely the training and testing areas. The training area spanned 3270 km² and encompassed a total of 328 560 sampling data points. On the other hand, the testing area covered 2000 km² and comprised 201 254 sampling data points, as indicated in Table 2. Three distinct sets, each comprising different proportions of training samples, were chosen for analysis. These sets were labeled as A (comprising the entire data sample), B (comprising two-thirds of the 0-class sample), and C (comprising an equal number of samples from both the 0 and 1 ophiolite classes). The primary objective was to assess whether a reduction in the quantity of sampling data points, coupled with an increase in the proportion of ophiolites and non-ophiolites in the training dataset, would yield improved predictions in the testing dataset. It is worth noting that the testing dataset is significantly imbalanced, as evidenced by Table 2.

Results

Random forest

The RF algorithm was applied to all three datasets (A, B, and C) with varying numbers of data points and distinct ratios of the two binary categorical classes pertaining to

Table 2 Datasets A, B and C with number of points and ratios for training and testing area

Dataset A					
	0	1	Sum	0 [%]	1 [%]
Training	273 106	55 454	328 560	83	17
Testing	189 116	12 138	201 254	93	7
Sum	462 222	67 592	529 814	87.2	12.8
Dataset B					
	0	1	Sum	0 [%]	1 [%]
Training	100 000	55 454	155 454	64.3	35.7
Testing	189 116	12 138	201 254	93	7
Sum	289 116	67 592	356 708	81	19
Dataset C					
	0	1	Sum	0 [%]	1 [%]
Training	55 454	55 454	110 908	50	50
Testing	189 116	12 138	201 254	93	7
Sum	244 570	67 592	312 162	78	22

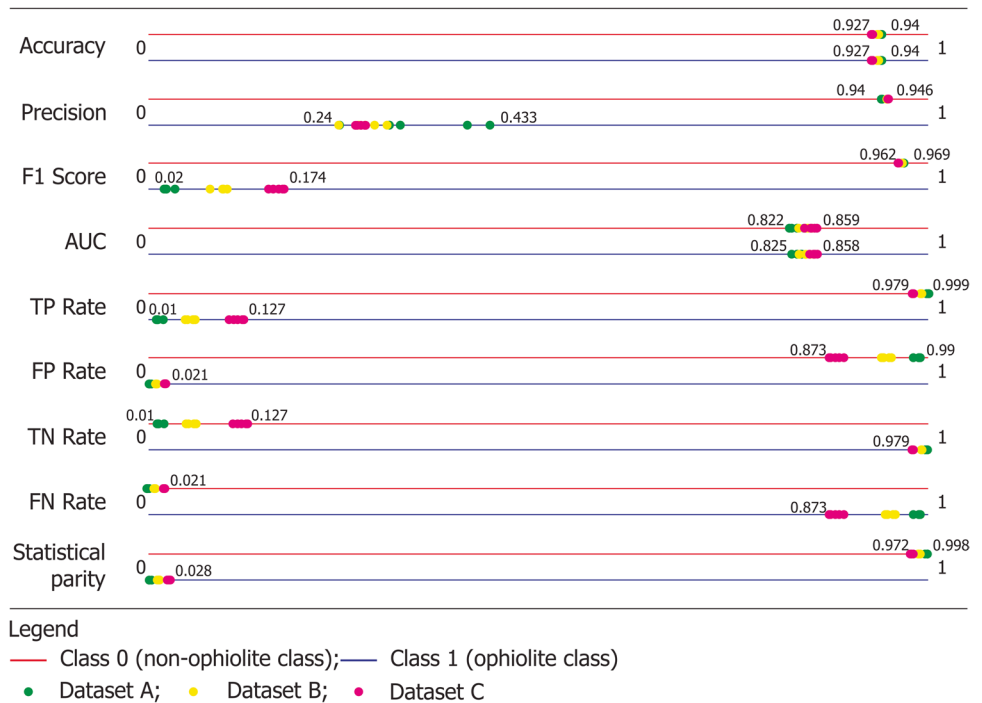
-non ophiolite class; **1**—ophiolite class;

Sum—total number of points per dataset

ophiolites. The RF modeling process encompassed five distinct models per dataset, each with varying numbers of trees that ranged from 50 to 250, with increments of 50. The RF modelling yielded both evaluation metrics and the predictive maps of ophiolites.

Figure 6 presents a set of ten evaluation metrics that have been chosen to assess the performance of the model across various training data sets and different numbers of trees in the RF model. Regarding accuracy, the models that employed the A dataset exhibit marginally better accuracy for both categories compared to RF models utilizing B or C datasets. However, the difference is negligible, with only a 1.3% variance between the model that demonstrated the highest accuracy (RF_A_100) and the model with the lowest accuracy (RF_C_100). The precision metric pertaining to the non-ophiolite class is analogous to the accuracy metric, as the disparity between the two is negligible. However, for the ophiolite class, the difference is considerably more obvious, with two distinct outliers exhibiting an overall higher precision. Specifically, the RF models that employ the A dataset with 100 and 250 trees demonstrate this trend. The F1 score is a crucial evaluation metric for assessing the overall performance of a model and is on par with the AUC (Area Under the Receiver Operating Characteristic Curve) metric. It displays a categorization of the values that are associated with different datasets for the ophiolite class. The F1 score exhibits the highest values for the balanced dataset (dataset C), reaching up to 0.174. Conversely, dataset A displays the lowest F1 score values, descending as low as 0.02 (RF_A_250). The AUC values demonstrate a distinction in values across various datasets,

Fig. 6 Selected class evaluation metrics for the RF modelling



with dataset C exhibiting higher values and dataset A exhibiting lower values. The true positive (TP) rate of the ophiolite class indicates that dataset C exhibits the highest rates, while conversely, the dataset C category displays the lowest rates of the true positive rate. The statistical parity, specifically, for dataset C, exhibits peak values of approximately 0.028. This indicates that when the number of data points decreases and the proportion of ophiolites to non-ophiolites approaches equilibrium, the trained model improves in its ability to predict the ophiolite class.

The results of the statistical analysis conducted on the evaluation metrics of the 15 RF models indicate that the most optimal dataset is the C dataset, which is characterized by an even balance of ophiolites to non-ophiolites. In general, the ophiolite class exhibits the highest F1 scores in the C dataset, while the F1 scores for the non-ophiolite class are relatively similar across all three datasets. The dataset denoted as C exhibits the highest AUC values, albeit a slight variance is discernible when compared to other input datasets. Based on the statistical parity metric, it can be observed that the C dataset exhibits the highest likelihood of predicting the ophiolite class. Furthermore, the TP rate for the ophiolite class is also the highest for the C dataset. Overall, within the C dataset, the RF_C_100 model, which employed 100 trees for modeling, was found to be the most optimal. The model exhibits a relatively good AUC of 0.853, with the ophiolite class achieving the highest F1 score of 0.172. Additionally, the model demonstrates a statistical parity value of 0.028, which is also the highest among all models.

In addition to the statistical examination of evaluation metrics, an equally significant analysis pertains to the visual analysis of classification maps, i.e., predictive maps. The predictive maps of the chosen models for all three datasets are illustrated in Fig. 7. The chosen RF models comprise of 50, 100, and 250 trees for every used input dataset.

The RF model, which employed the A dataset and consisted of 50 trees as depicted in Fig. 6, generated 935 predictions for the ophiolite class. Most of these predictions were observed in the NW region of the EVZ. Furthermore, a secondary group of ophiolites is predicted to be present in a limited region situated in the SE part of the EVZ. Conversely, the RF models comprising of 100 and 250 trees exhibited a significantly lower number of predictions, specifically 328 and 302, respectively. The small pocket located in the SE of the EVZ exhibited ophiolite class predictions by both RF_A_100 and RF_A_250. However, it is noteworthy that the A dataset models, overall, exhibit a significant under-prediction of the non-ophiolite class. This can be attributed to the imbalance of the training dataset employed. Furthermore, from a geological perspective, it can be observed that the A dataset models fail to depict the ophiolites' general NE to SW positioning within the EVZ, which is situated at the center of the region.

The initial visual comparison of RF models using dataset A and dataset B reveals that the RF models utilizing dataset B generated a significantly higher number of predictions for the ophiolite class. The RF models that employ a moderately balanced dataset, though not entirely balanced, yield ophiolite class predictions ranging from 2086 to 2391 instances.

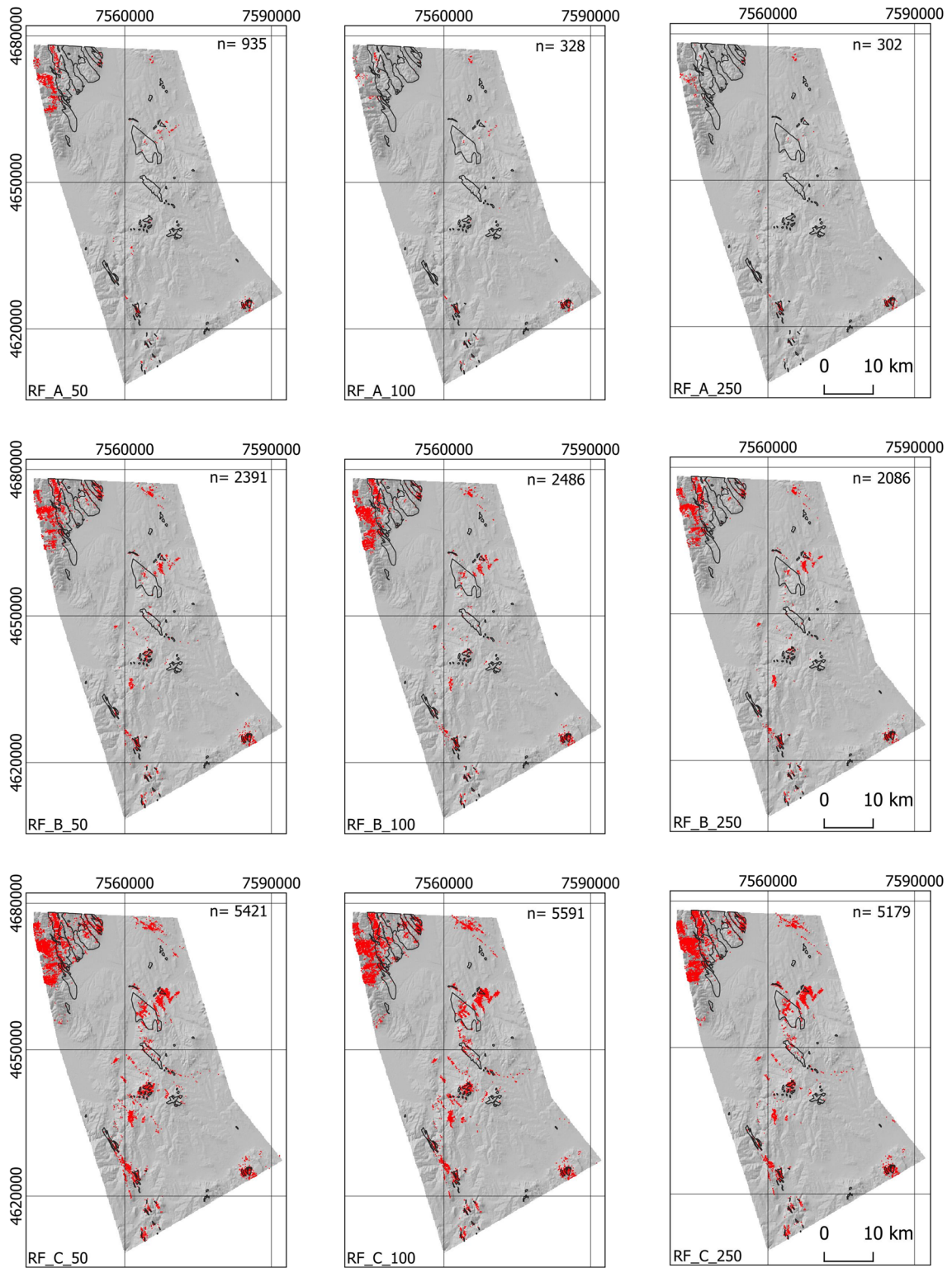


Fig. 7 Selected predictive maps for the RF models; n- number of ophiolite class instances; black outline- field mapped ophiolites; red markers- predicted ophiolites

This outcome is consistent with the statistical parity evaluation metric observed earlier. Most of the predictions are situated in the NW region of the EVZ, accompanied by supplementary predictions in the minor enclave in the SE portion of the EVZ. Further predictions pertaining to the ophiolite classification may be discerned at the center of the EVZ, exhibiting a directional orientation of NE-SW. The RF models that employed the B dataset exhibited better performance in comparison to the RF models that employed the A dataset, despite having approximately 50% less data. This observation was further corroborated by the evaluation metrics displayed previously.

Upon visual inspection, it can be inferred that the optimal input dataset among the three RF model groups is dataset C, as supported by the analysis of the evaluation metrics. The dataset denoted as C exhibits many ophiolite classifications in the NW sector of the EVZ. However, a noteworthy advantage of a fully balanced dataset is that the RF C models can accurately predict the geological positioning of ophiolites. The C dataset RF models provide the most accurate prediction of the overall NE-SW strike of ophiolites in the EVZ. Notably, the center of the EVZ exhibits the highest concentration of ophiolite class predictions across all three model groups. Furthermore, it was observed that the RF models trained on the C dataset exhibited the highest number of predictions for the ophiolite class, with a maximum of 5591 predictions achieved by the RF_C_100 model. It is noteworthy that despite having a significantly lower quantity of data, approximately 66% less than the A dataset and 30% less than the B dataset, the C dataset yielded the most accurate predictions.

The study conducted a visual and statistical analysis of evaluation metrics for three groups of RF models with varying numbers of trees. The results indicate that the RF model utilizing the completely balanced C dataset, despite having the least amount of data for model training compared to other input data groups, yields the best overall outcomes for predictive ophiolite mapping.

Feature importance analysis

The RF model, which was trained on the C dataset and consisted of 100 trees, exhibited better performance in terms of evaluation metrics and visual representation through prediction maps. All models employed the complete set of 13 features to generate the prediction map. Additionally, the RF modeling yielded feature importance outputs, which were determined by the MDNA and TINP, as depicted in Fig. 8.

In the RF_C_100 model, the top five features ranked by MDNA and TINP exhibit similarity. These features include the Bouguer anomaly map, distance to fault map, digital elevation model, total intensity of the Earth's magnetic field reduced to the pole, and BR3 map. Likewise, the remaining 14 models exhibited identical outcomes, with the top five most informative parameters consistently ranking in the same order. Additionally, all models shared a commonality in the two parameters that were ranked as the least informative, specifically the Landsat 7 ETM+ channel 4 and BR1 map. Regarding the RF_C_100 model, it is significant to mention that the BAM feature is by far the most informative feature for ophiolite prediction, as indicated by both the MDNA and TINP (Fig. 7).

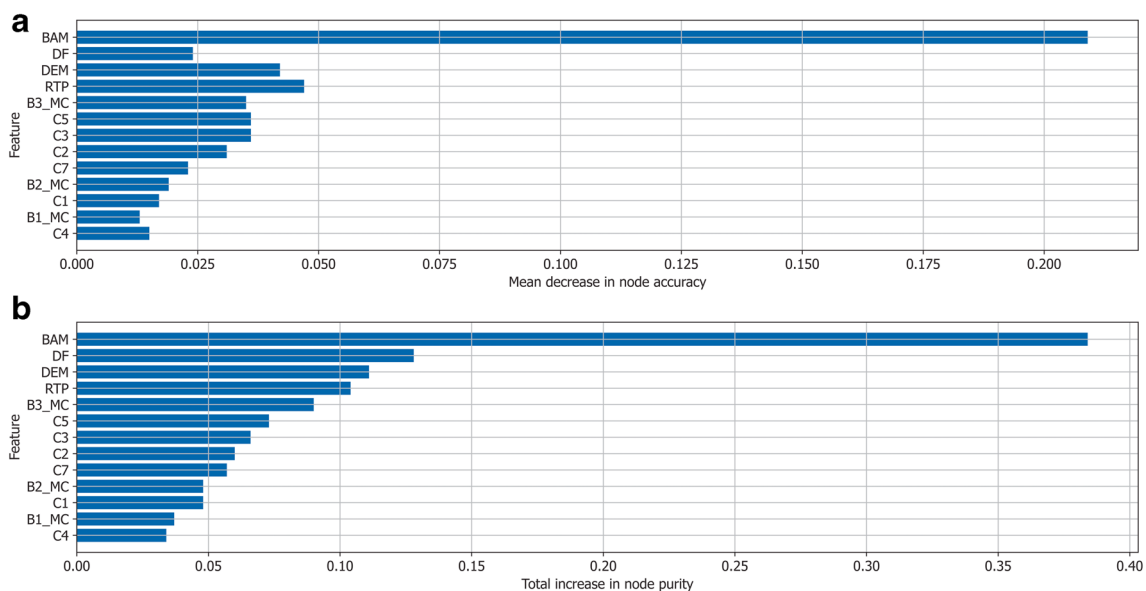


Fig. 8 Feature importance from RF_C_100 model; a) Mean decrease in node accuracy; b) total increase in node purity

Table 3 Comparison between RF_C_100 and RF_C_100_5MI

Model	RF_C_100		RF_C_100_5MI	
	0	1	0	1
Accuracy	0.927	0.927	0.922	0.922
Precision	0.946	0.274	0.944	0.182
F1 Score	0.962	0.174	0.959	0.118
AUC	0.852	0.854	0.829	0.831
TP Rate	0.979	0.127	0.975	0.087
FP Rate	0.873	0.021	0.913	0.025
TN Rate	0.127	0.979	0.087	0.975
FN Rate	0.021	0.873	0.025	0.913
Statistical Parity	0.972	0.028	0.971	0.029
Computational time[s]	210		93	

0- non ophiolite class; 1—ophiolite class; TP- True positive; FP- False positive; TN- True negative; FN- False negative.

To assess the prospective reduction in computational cost linked to RF modelling, the modelling procedure was replicated using the same input dataset (C) and tree count (100). However, unlike RF_C_100, which employed all 13 input features, only the five most informative features, as depicted in Fig. 7, were utilized. The model that was created (RF_C_100_5MI) was compared with the RF_C_100. Table 3 illustrates that the RF_C_100_5MI model shows lower overall performance, as evidenced by the F1 score for the ophiolite class. The discrepancy amounts to approximately 5%. Furthermore, it can be noted that there exists a disparity of approximately 2% in the AUC metric for the ophiolite class, with the RF_C_100 model exhibiting a slightly higher AUC value. The precision metric exhibits a significant contrast, with RF_C_100 demonstrating a 9% higher precision. In general, the evaluation metrics of the RF_C_100_5MI model are most compatible with the evaluation metrics exhibited by the models of the B dataset, except for the statistical parity parameter which is most compatible with the C dataset models. A noteworthy comparison can be drawn between the duration of training and testing for two models. Expectedly, the model that employs all 13 features, which is deemed the overall better model, requires approximately 55% more computational time than the model that employs the same dataset but only the five most informative features.

To fully quantify the performance of the RF_C_100_5MI model, a prediction map was generated using the same methodology as the preceding models (Fig. 9). Table 3 reveals that the RF_C_100_5MI model exhibits a statistical parity evaluation metric that is comparable to that of the RF_C_100 model. Furthermore, while the evaluation metrics of the RF_C_100_5MI model are more comparable with the B input dataset models, the prediction map and the number of ophiolite instances predicted are equivalent to those

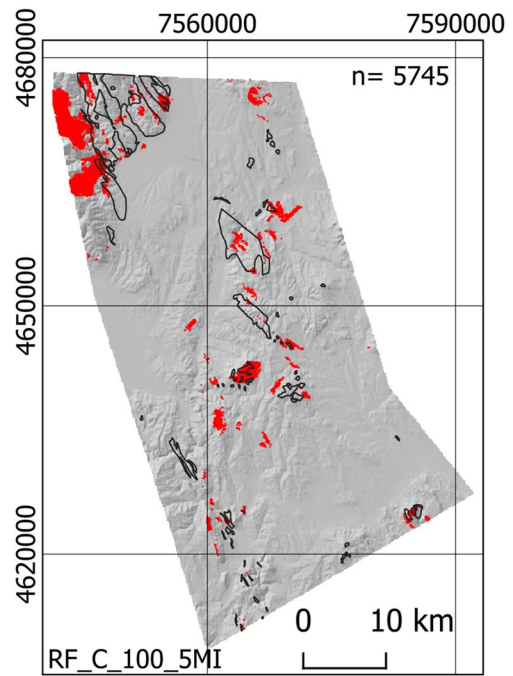


Fig. 9 RF model with only five most informative features (RF_C_100_5MI); n- number of ophiolite class instances; black outline- field mapped ophiolites; red markers- predicted ophiolites

of the C dataset models. In general, the RF_C_100_5MI model exhibits comparable prediction maps to those of the C dataset models. Notably, the NW region of the EVZ demonstrates many ophiolite class predictions, along with a minor pocket situated in the SE region of the EVZ. The ophiolites in the EVZ were accurately predicted to have a NE-SW strike. Moreover, the projected count of ophiolite occurrences is comparable to the models in the C dataset, with a predicted value of 5745 ophiolite instances.

In general, the RF_C_100_5MI model exhibits satisfactory evaluation metric statistics and provides an overall adequate geological (visual) prediction of ophiolites in the EVZ. Despite having 8 fewer features, this model offers the added benefit of a 55% reduction in computation time. In this instance, it may be beneficial to make a trade-off by selecting a model that is statistically inferior but yields satisfactory visual outcomes, while requiring 55% less computational time. This holds significant value as it displays that not an overabundance of features is needed for RF modelling, and effective modelling and prediction can be done with five most informative features and a balanced dataset.

K nearest neighbors

The KNN modeling was conducted utilizing solely the top five most informative features that were employed in the RF modeling. The RF_C_100_5MI model demonstrated

reasonable evaluation metric statistics and satisfactory visual predictions of ophiolites, while also offering the advantage of 55% less computational time. Figure 10 presents the evaluation metrics for the five KNN models that were constructed using K values ranging from 1 to 5, with increments of 1.

The F1 score for the ophiolite class indicates that the disparity between the highest-performing model and the lowest-performing model is approximately 1.7%. In contrast, the largest discrepancy in the AUC metric is approximately 8.8%. The difference between the AUC parameter of the optimal KNN model and the RF models is significant, with the KNN models exhibiting markedly inferior AUC values. It is noteworthy to mention that the KNN model exhibits a higher TP rate and statistical parity parameters. The KNN models exhibit a higher likelihood of predicting the ophiolite class in comparison to the RF models, with the former being approximately three times more likely to do so. Specifically, the maximum probability of ophiolite class prediction for the RF models is approximately 2.9%, while that for the KNN models is approximately 6.6%.

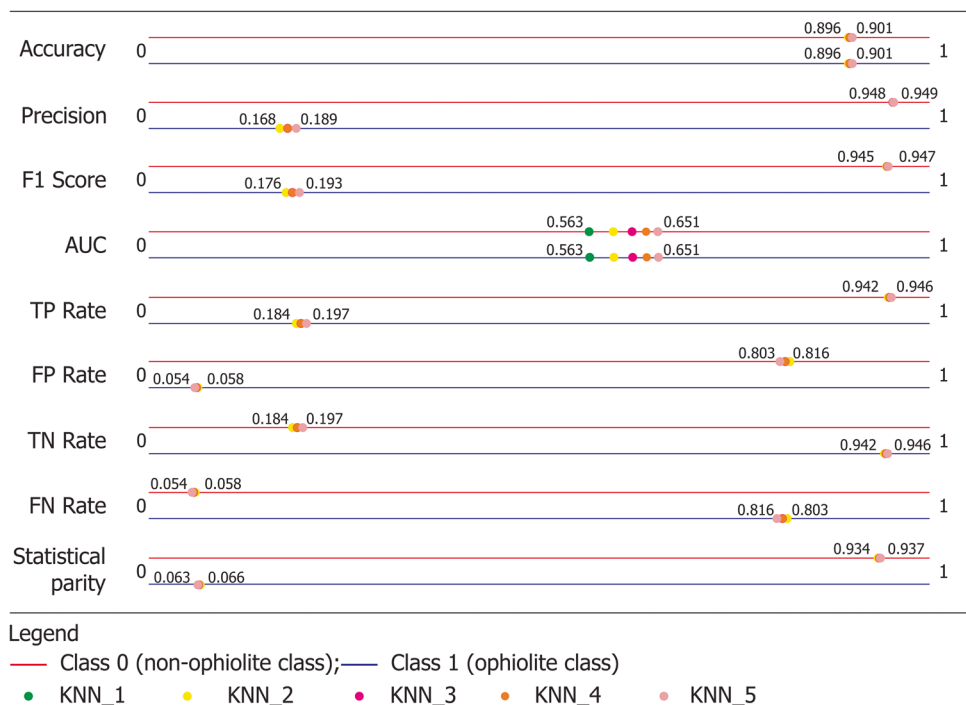
The prediction maps for the three KNN models, specifically the models utilizing 1, 2, and 5 nearest neighbors, are depicted in Fig. 11. One notable modification that can be implemented on the RF_C_100_5MI model pertains to the considerable surplus in the number of predicted ophiolite class instances, which surpasses ten thousand, analogous to the statistical parity metric. In addition to variations in the predicted count of ophiolite instances, the KNN models are comparable with the RF models. Most ophiolite class

instances predictions are situated in the NW portion of the EVZ. Additionally, the ophiolites' general orientation in the central region of the EVZ is NE-SW, a pattern that all the KNN models accurately predicted. Furthermore, it is possible to observe an area consisting of ophiolite instances in the NE region of the EVZ, which was present in the RF models but not as predominant as in the KNN models.

Based on the KNN modeling, evaluation metric, and analysis, it can be inferred that the KNN model with five nearest neighbors is the most optimal KNN model. Regarding the evaluation metrics, the ophiolite class exhibits the highest F1 score and AUC. However, the AUC value is notably lower than that of the corresponding RF model (RF_C_100_5MI). In general, the KNN models exhibited satisfactory visual and geological foundations for ophiolite prediction, utilizing only five most informative features. However, the predictions are comparatively inferior to those of the RF models.

A comparative analysis can be conducted between RF_C_100_5MI and KNN_C_5, both of which are considered the top-performing models, with RF_C_100 serving as the reference model. A false positive (FP) map can be utilized to draw a comparison, wherein all instances that a particular model has made a FP prediction for are represented (Fig. 12). The significance of the FP map lies in its ability to identify potential areas where ophiolites may be absent from mapping or situated beneath the subsurface and were not able to be mapped with traditional surface geological mapping techniques. The RF_C_100_5MI model exhibited a lower incidence of FP predictions as compared to the KNN model. This

Fig. 10 Selected class evaluation metrics for the KNN modelling



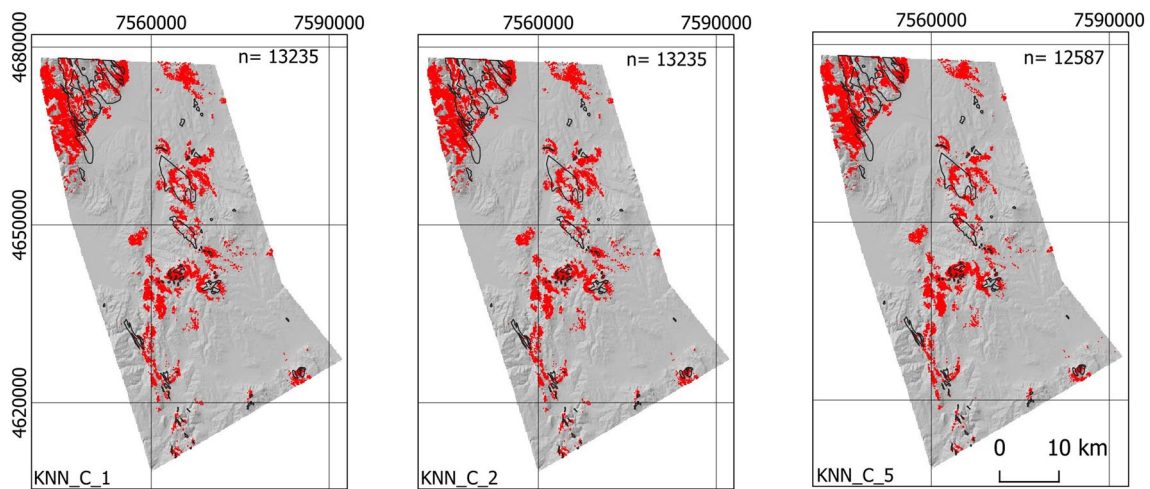


Fig. 11 Selected predictive maps for the KNN models; n- number of ophiolite class instances; black outline- field mapped ophiolites; red markers- predicted ophiolites

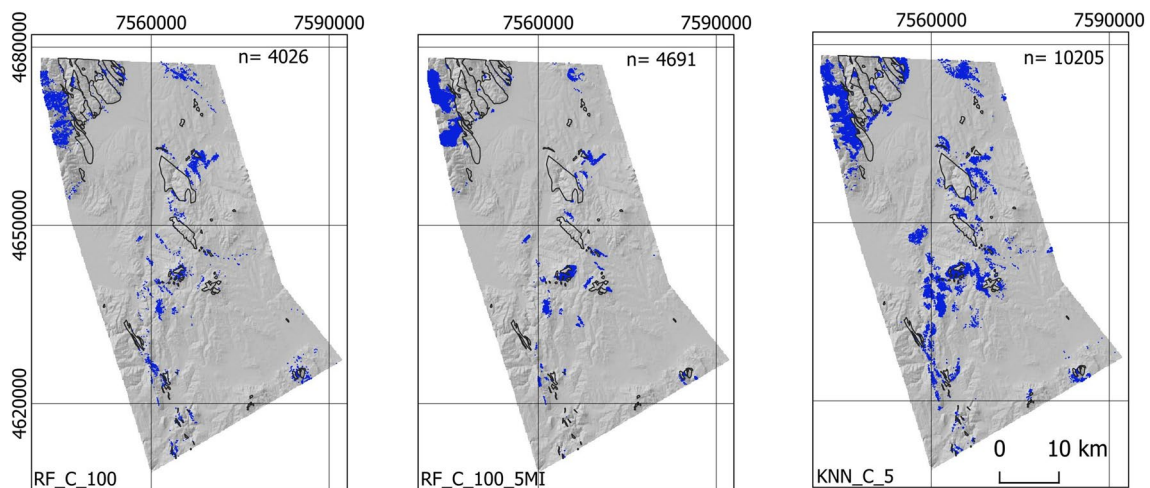


Fig. 12 False positive maps for the RF and KNN models; n- number of false positive instances; black outline- field mapped ophiolites; blue markers- false positive instances

outcome was anticipated, given the KNN model’s higher statistical parity metric and the greater number of predicted ophiolite instances overall. The RF_C_100 model exhibited a reduction of 600 FP instances in comparison to the RF model that employed 5 features. The three models exhibited most of the FP predictions in the proximity of pre-existing or mapped ophiolites. Most of them are in the NW part of the EVZ and in the central region of the EVZ, close to the identified ophiolites. When compared, it was observed that the KNN model exhibited a higher number of FP predictions in comparison to both the RF models. It is noteworthy that a specific subset of these FP predictions was situated along the NE-SW strike of the ophiolites through the EVZ.

Discussion

The evaluation of two RF models revealed that the model with a reduced number of features demonstrated lesser performance based on the evaluation metrics. However, it exhibited a notable reduction of approximately 55% in computational time in comparison to the other model. In this instance, the absolute disparity was a modest 137 s. However, when confronted with more extensive datasets, such as a broader region of interest or data of higher resolution, the decrease in computational time may prove to be considerably more noteworthy. The RF_C_100_5MI was selected as the optimal model. Regarding the visual

comparison of the two RF models, it can be observed that the RF_C_100_5MI model produced a prediction map that is satisfactory and closely resembles the one generated by the RF model that employed all 13 features. Furthermore, the KNN model exhibited poorer evaluation metrics statistics, particularly about the AUC parameter, which was not in line with those of the RF model, given the same set of features and input dataset. The KNN model exhibited visually satisfactory prediction maps. Regarding the most informative features, the results revealed intriguing patterns that require further investigation to fully comprehend the geological and geophysical implications. Feature importance is a valuable piece of information that necessitates additional research to establish a comprehensive link between its numerical output and the real nature of the phenomena.

When comparing the RF and KNN models using evaluation metric statistics, it is necessary to take into account various aspects. Although an effort was made to train and test the KNN model using only the five most informative features, as it operates more efficiently on lower dimensional data (i.e., with fewer features), it was ultimately exceeded by the RF model. However, in contrast to the KNN model, the RF model exhibits more robustness towards outliers, meaning it is less sensitive to them, and it also more effectively captures non-linear relationships within the data. In summary, the RF model exhibits better capability in capturing complex relationships within the data when compared to the KNN model; thus, the KNN model is more appropriate for simpler datasets. This suggests that future research efforts could be focused on more complex models rather than the RF model, as they may be better equipped to handle the utilized data. However, it is important to note that more complex models have a greater number of hyperparameters that require tuning, resulting in a more complicated method of model tuning. While this can be efficiently accomplished through the use of various grid search techniques, it will require greater computational time.

The FP prediction map holds great significance as it enables an investigation into potential ophiolite sites that were not identified through surface mapping methodologies or are situated beneath the surface. In the context of data analysis in areas with dense vegetation, such as the EVZ region, it is possible that the FP prediction map may provide valuable insights for future geological surveys, specifically in identifying potential ophiolite locations.

From a statistical standpoint, it can be observed that all the models generated exhibit a minimal number of predictions related to the total number of ophiolite instances. Despite being well-presented, the most proficient model failed to identify approximately 70% of ophiolite instances within the EVZ. Conversely, incorporating statistical rigor into models that utilize data collected from challenging conditions and

incorporating legacy geological data may not be the best course of action. The C dataset models demonstrated a reliable geological prediction map that could potentially serve as a basis for subsequent surveys or more comprehensive ophiolite exploration endeavors. The C dataset models accurately predicted the general occurrence of the ophiolite belt strike in the EVZ, despite their limited ability to predict most individual ophiolite instances. Despite the absence of the most statistically precise models, the visual analysis of the prediction maps demonstrated that the selected models are adequate for the given scale of mapping and the size of the study area.

As demonstrated in a number of studies (Cracknell and Reading 2014; Ge et al. 2020; Shayeganpour et al. 2021; Bachri et al. 2022), the RF method outperforms the KNN model, which was confirmed in this piece of research. There are multiple potential causes for this result; however, the most likely is that the KNN model may not be optimal when dealing with a large feature space, compounded by the likely nonlinear relationship between the features and the target variable. Furthermore, the RF algorithm emerged as a highly favorable initial option, as demonstrated in this study, which emphasized its straightforwardness in implementation (requiring only a few hyperparameters to be adjusted), resistance to overfitting, and overall computational simplicity.

Additionally, it is important to briefly discuss the features employed and their accessibility for researchers seeking to implement the same approach in their own survey region. The feature importance analysis revealed that the five previously mentioned features are sufficient for achieving a sufficiently accurate prediction of ophiolites. The features encompass data that may not be readily accessible to all researchers, such as the Bouguer anomaly map (which can be costly if a gravity survey was not conducted), the total intensity of the Earth's magnetic field (similar to the Bouguer anomaly map), the BR3 map (which necessitates at least two multispectral imagery bands), the distance to fault map (which requires extensive regional geologic mapping), and the digital elevation model (which relies on satellite data). Although ML methods necessitate a substantial amount of high-quality data, additional research can be conducted to provide best practices recommendations regarding the anticipated outcomes when utilizing all available data from various sources (such as this study) and offer insights into the expected outcomes when working with limited data and identify scenarios where attempting ML classification is advisable or not, as ML is a data-driven process.

Conclusion

Based on available geological data, satellite images, and geophysical measurements, this study focused on mapping the extremely complex geological and tectonic units under

conditions of dense vegetation. This approach was evaluated to determine if it provided any useful information. The survey was conducted in the central region of North Macedonia and encompasses a larger portion of the EVZ. The input data (features) for the RF algorithm consisted of bands of Landsat 7 ETM+ satellite images, a Total Intensity EMF Anomaly Map reduced to the pole, distance to fault, digital elevation model, and a Bouguer Anomaly Map, unlike other similar studies. The Bouguer Anomaly Map was chosen because it directly indicates the appearance of a body with a higher density, as was the case with gabbro and dolerite as integral components of ophiolite. Due to the 100 m working resolution, local geophysical variations were disregarded.

The RF algorithm was employed to compare input datasets containing varying degrees of ophiolite and non-ophiolite classes. The results of the comparison revealed that dataset C, which was characterized by a balanced 50–50 percent of ophiolites and non-ophiolites in the training dataset, yielded the most optimal outcomes.

According to the results of the feature importance analysis, the Bouguer gravity anomaly map, Total intensity EMF reduced to the pole, distance to fault map, digital elevation model, and BR3 map are the most significant features for predicting ophiolites. The statistical metrics of the model that incorporated all 13 features were better. However, the RF model that employed all features and the model that utilized only the five most informative features exhibited similar ophiolite distribution through the EVZ, as observed visually. Based on statistical analysis, the model that employed only the top five informative features is kin statistically for slightly imbalanced training datasets, such as dataset B. However, upon visual inspection, the models trained on dataset B and the RF model utilizing only the top five informative features and the balanced dataset C exhibit significant dissimilarities. One benefit of utilizing a reduced feature set consisting of only five of the most informative features is a reduction in computational time. Specifically, this approach results in a decrease of approximately 55% in computational time. Furthermore, it is crucial to emphasize the importance of employing a balanced training dataset when dealing with a target class imbalanced testing area. This is because models trained on an imbalanced dataset (dataset A) tend to exhibit significant overestimation of the non-ophiolite class. Implementing a balanced dataset that contains a reduced number of samples (approximately 66% fewer in this instance) yields more precise models. This indicates that in certain ML classification scenarios, a smaller amount of data can be advantageous.

Comparing the RF and KNN algorithms produced results consistent with the existing literature, the RF algorithm produced superior results based on the evaluation

metrics of both algorithms. The KNN model, utilizing dataset C as the input dataset, exhibited an FP rate that was nearly two and a half times that of the RF model, which also utilized the same input features and dataset. Furthermore, it was observed that the KNN model exhibited comparatively inferior AUC values, with the maximum value recorded at 0.651 (KNN_C_5). On the other hand, the RF model yielded AUC values of 0.831 (RF_C_100_5MI). Conversely, visual analysis indicated that the KNN algorithm accurately approximated the spatial distribution of the ophiolites within the test area of the EVZ.

In densely vegetated regions with complex geological and tectonic settings, a RF algorithm with appropriate features and a balanced dataset yields satisfactory results for predictive mapping of complex units, such as ophiolites, and may indicate the existence of these units even when they were not field mapped, especially when geophysical data are used as features.

Authors contribution This research project represents a collective effort in which each author played a crucial role in shaping its conception, methodology, and design. The initial idea was ignited by Dragana Đurić and Uroš Đurić, who laid the foundation for the study by developing the very first models and utilizing machine learning algorithms. Subsequently, Mileva Samardžić-Petrović took charge of the statistical analyses and delved into the initial exploration of machine learning algorithms. Building upon this groundwork, Filip Arnaut made significant contributions and novelty by enhancing the machine learning techniques, ultimately introducing the novel application of KNN, which led to the development of entirely new models and enriched the analysis. Moreover, the research was complemented by the invaluable expertise of Igor Peshevski, who conducted an in-depth literature review focusing on geological and tectonic aspects, contributing to a more comprehensive understanding of the context. The collaborative nature of this endeavor is exemplified in the writing process, wherein Filip Arnaut took the lead in crafting the manuscript while benefiting from substantial contributions and assistance from Dragana Đurić, Uroš Đurić, Mileva Samardžić-Petrović, and Igor Peshevski. Throughout this stage, the manuscript underwent internal review process, with each author providing valuable comments and feedback on earlier versions. The culmination of this collaborative effort is the final manuscript, which has been thoroughly reviewed, refined, and approved by all authors.

Funding The study was funded by the Ministry of Education, Science and Technological Development, Republic of Serbia under the "Agreement on the realization and financing of scientific research work in 2023": grant no. 451–03-47/2023–01/200126 (Faculty of Mining and Geology) and grant no. 451–03-68/2020–14/200092 (University of Belgrade, Faculty of Civil Engineering).

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflicts of interest/Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

- Akhavi MS, Webster TL, Raymond DA (2001) RADARSAT-1 Imagery and GIS Modeling for Mineral Exploration in Nova Scotia, Canada. *Geocarto Int* 16:57–64. <https://doi.org/10.1080/10106040108542183>
- Albert G, Ammar S (2021) Application of random forest classification and remotely sensed data in geological mapping on the Jebel Meloussi area (Tunisia). *Arab J Geosci* 14. <https://doi.org/10.1007/s12517-021-08509-x>
- Aliyu A, Adamu LM, Abdulmalik NF et al (2021) Application of remote sensing in lithological discrimination of precambrian basement rocks of Zungeru area, part of sheet 163 (Zungeru NW), North Central Nigeria. *FJS* 5:390–398. <https://doi.org/10.33003/fjs-2021-0503-729>
- Al-Rawashdeh S, Saleh B, Hamzah M (2006) The use of Remote Sensing Technology in geological Investigation and mineral Detection in El Azraq-Jordan. *Cybergeogeo*. <https://doi.org/10.4000/cybergeogeo.2856>
- Bachri I, Hakdaoui M, Raji M et al (2022) Identification of Lithology Using Sentinel-2A Through an Ensemble of Machine Learning Algorithms. *Int J Appl Geospat Res* 13:1–17. <https://doi.org/10.4018/ijagr.297524>
- Behnia P, Harris JR, Rainbird RH et al (2012) Remote predictive mapping of bedrock geology using image classification of Landsat and SPOT data, western Minto Inlier, Victoria Island, Northwest Territories, Canada. *Int J Remote Sens* 33:6876–6903. <https://doi.org/10.1080/01431161.2012.693219>
- Bilibajkić P, Mladenović M, Mujagić S, Rimac I (1979) Explanation for the gravity map of SFR Yugoslavia - Bouguer anomalies - 1:500 000. Federal Geological Institute, Belgrade
- Boev B, Cvetković V, Prelević D, Šarić K, Boev I (2018) East Vardar Ophiolites Revisited A Brief Synthesis Of Geology And Geochemical Data. *Contrib Sec Nat Math Biotech Sci, Macedonian Academy of Sciences and Arts* 39:51–68. <https://doi.org/10.20903/csnmbs.masa.2018.39.1.119>
- Bolt GH, Bruggenwert MGM (1976) Composition of the Soil. *Soil Chemistry: A Basic Elements* 1–12. [https://doi.org/10.1016/s0166-2481\(08\)70630-5](https://doi.org/10.1016/s0166-2481(08)70630-5)
- Breiman L (2001) Random Forests. *Mach Learn* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- Brockmann CE, Fernandez A, Ballon R, Claire I (Servicio Geologico de Bolivia, La Paz, Bolivia) (1977) Analysis of geological structures based on landsat-1 images. In: *Remote Sensing Applications for Mineral Exploration*. Down, Hutchinson and Ross, Stroudsburg, PA, USA, pp 292–317
- Carranza EJM, Laborte AG (2015a) Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Comput Geosci* 74:60–70. <https://doi.org/10.1016/j.cageo.2014.10.004>
- Carranza EJM, Laborte AG (2015b) Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm. *Ore Geol Rev* 71:777–787. <https://doi.org/10.1016/j.oregeorev.2014.08.010>
- Cassinis G (1930) Sur L'adoption D'une Formule Internationale Pour la Pesanteur Normale. *Bull Géod* 26:40–49. <https://doi.org/10.1007/bf03030025>
- Catani F, Lagomarsino D, Segoni S, Tofani V (2013) Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Nat Hazards Earth Syst Sci* 13:2815–2831. <https://doi.org/10.5194/nhess-13-2815-2013>
- Cover T, Hart P (1967) Nearest neighbor pattern classification. *IEEE Trans Inf Theory* 13:21–27. <https://doi.org/10.1109/tit.1967.1053964>
- Cracknell MJ, Reading AM (2014) Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput Geosci* 63:22–33. <https://doi.org/10.1016/j.cageo.2013.10.008>
- Cracknell JM (2014) *Machine Learning for Geological Mapping: Algorithms And Applications*. Dissertation, School of Physical Sciences (Earth Sciences), University of Tasmania
- Cutler DR, Edwards TC, Beard KH et al (2007) Random forests for classification in ecology. *Ecol* 88:2783–2792. <https://doi.org/10.1890/07-0539.1>
- Cvetkov V, Đurić D, Lesić V, Starčević M, Petković M, Petrović S (2016) Koenigsberger ratio and Total Magnetic Field Anomaly reduction to the pole for the area of Macedonia. *Geologica Macedonica* 4:429–534
- Dimitrijević MD (1978) *Geological mapping*. Publishing and informational student center, Belgrade (In Serbian)
- Dimitrijević MD (1997) *Geology of Yugoslavia*. Geological Institute – GEMINI, Belgrade, Spec. Publ., Monograph, 1–197
- Dumurdzanov N, Hristov S, Pavlovski B, Ivanova V (1981) Explanatory notes for sheets Vitolište and Kajmakčalan. General geological map (1:100,000) of the Socialist Federal Republic of Yugoslavia, Federal Geological Survey, Belgrade, p 61 (in Macedonian)
- Đurić D (2023) Overview on the Results of Potential Field Methods for the Spatial Position of East Vardar Ophiolites. In: Chitea F (ed) *Insights of Geosciences for Natural Hazards and Cultural Heritage*. Zenodo. <https://doi.org/10.5281/zenodo.8103476>
- Farrand WH (1997) Identification and mapping of ferric oxide and oxyhydroxide minerals in imaging spectrometer data of Summitville, Colorado, U.S.A., and the surrounding San Juan Mountains. *Int J Remote Sens* 18:1543–1552. <https://doi.org/10.1080/014311697218269>
- Fix E, Hodges JL (1951) Discriminatory Analysis. *Nonparametric Discrimination: Consistency Properties*. *International Statistical Review / Revue Internationale de Statistique* 57:238. <https://doi.org/10.2307/1403797>
- Foody GM, Mathur A (2004) A relative evaluation of multiclass image classification by support vector machines. *IEEE Trans Geosci* 42:1335–1343. <https://doi.org/10.1109/tgrs.2004.827257>
- Ge W, Cheng Q, Tang Y et al (2018) Lithological Classification Using Sentinel-2A Data in the Shibanzhong Ophiolite Complex in Inner Mongolia. *China Remote Sens* 10:638. <https://doi.org/10.3390/rs10040638>
- Ge G, Shi Z, Zhu Y et al (2020) Land use/cover classification in an arid desert-oasis mosaic landscape of China using remote sensed imagery: Performance assessment of four machine learning algorithms. *Glob Ecol Conserv* 22. <https://doi.org/10.1016/j.gecco.2020.e00971>
- Ge YZ, Zhang ZJ, Cheng QM, Wu GP (2022) Geological mapping of basalt using stream sediment geochemical data: Case study of covered areas in Jining, Inner Mongolia. *China J Geochem Explor* 232:106888. <https://doi.org/10.1016/j.gexplo.2021.106888>
- Ham J, Chen Y, Crawford MM, Ghosh J (2005) Investigation of the random forest framework for classification of hyperspectral data. *IEEE Trans Geosci* 43:492–501. <https://doi.org/10.1109/tgrs.2004.842481>
- Harris JR, Grunsky EC (2015) Predictive lithological mapping of Canada's North using Random Forest classification applied to geophysical and geochemical data. *Comput Geosci* 80:9–25. <https://doi.org/10.1016/j.cageo.2015.03.013>
- Harris JR, Rogge D, Hitchcock R et al (2005) Mapping lithology in Canada's Arctic: application of hyperspectral data using the minimum noise fraction transformation and matched filtering. *Can J Earth Sci* 42:2173–2193. <https://doi.org/10.1139/e05-064>
- Harris JR, Ford KL, Charbonneau BW (2009) Application of gamma-ray spectrometer data for lithological mapping in a cordilleran

- environment, Sekwi Region, NWT. *Can J Remote* 35:12–30. <https://doi.org/10.5589/m09-022>
- Harris JR, He JX, Rainbird R, Behnia P (2014) A Comparison of Different Remotely Sensed Data for Classifying Bedrock Types in Canada's Arctic: Application of the Robust Classification Method and Random Forests. *Geosci Can* 41:557. <https://doi.org/10.12789/geocanj.2014.41.062>
- Harris JR, Schetselaar EM, de Kemp E, St-Onge MR (2008) Case study 2. LANDSAT, magnetic and topographic data for regional lithological mapping, southeast Baffin Island. <https://doi.org/10.4095/226015>
- Hossain M, Sulaiman MN (2015). A Review on Evaluation Metrics for Data Classification Evaluations IJDKP. <https://doi.org/10.5121/ijdkp.2015.5201>
- Hristov S, Karajovanovik M, Stračkov M (1965) Basic geological map of SFRJ, sheet Kavadarci, M 1:100,000 (map & interpreter), Federal Geological Survey, Beograd, p 62 (in Macedonian)
- Hristov S, Karajovanovik M, Stračkov M (1973) Basic geologic map of Former Yugoslavia 1:100 000, explanatory booklet for sheet Kavadarci (in Macedonian)
- Huang C, Davis LS, Townshend JRG (2002) An assessment of support vector machines for land cover classification. *Int J Remote Sens* 23:725–749. <https://doi.org/10.1080/01431160110040323>
- Ivanovski T, Rakicević T (1966) Basic geologic map of Former Yugoslavia 1: 100 000, sheet of Gvegljija. Federal Geological Survey, Belgrade (In Macedonian)
- James G, Witten D, Hastie T, Tibshirani R (2013) *An Introduction to Statistical Learning: With Applications in R*. Springer, New York
- Joshi MV (2002) On evaluating performance of classifiers for rare classes. *IEEE Int Conf on Data Mining*, Maebashi City, Japan 641:644
- Kanevski M, Pozdnukhov A, Timonin V (2009) *Machine Learning for Spatial Environmental Data: Theory*. CRC Press, Boca Raton, USA, Applications and Software
- Karajovanovik M, Hadži-Mitrova S, (1975): Basic geologic map of Former Yugoslavia 1:100 000, explanatory booklet for the sheet of Titov Veles. Federal Geological Survey, Belgrade (In Macedonian)
- Karajovanovic M, Hristov S (1976) Explanatory note of the General Geological Map of Kumanovo 1:100,000 map sheet: Skopje, Federal Geological Survey Jugoslavia, p 58 (in Macedonian, with English summary)
- Kuhn S, Cracknell MJ, Reading AM (2018) Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. *Geophys* 83:B183–B193. <https://doi.org/10.1190/geo2017-0590.1>
- Kuncheva LI (2004) *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley & Sons Inc., Hoboken, New Jersey
- Leverington DW (2010) Discrimination of sedimentary lithologies using Hyperion and Landsat Thematic Mapper data: a case study at Melville Island, Canadian High Arctic. *Int J Remote Sens* 31:233–260. <https://doi.org/10.1080/01431160902882637>
- Leverington DW, Moon WM (2012) Landsat-TM-Based Discrimination of Lithological Units Associated with the Purtuniqu Ophiolite, Quebec, Canada. *Remote Sens* 4:1208–1231. <https://doi.org/10.3390/rs4051208>
- Longhi I, Sgavetti M, Chiari R, Mazzoli C (2001) Spectral analysis and classification of metamorphic rocks from laboratory reflectance spectra in the 0.4–2.5 μ m interval: A tool for hyperspectral data interpretation. *Int J Remote Sens* 22:3763–3782. <https://doi.org/10.1080/01431160010006980>
- Lorenz H (2004) Integration of Corona and Landsat Thematic Mapper data for bedrock geological studies in the high Arctic. *Int J Remote Sens* 25:5143–5162. <https://doi.org/10.1080/01431160410001705097>
- Moores EM (1982) Origin and emplacement of ophiolites. *Rev Geophys* 20:735–760. <https://doi.org/10.1029/RG020i004p00735>
- Neville RA, Lévesque J, Staenz K et al (2003) Spectral unmixing of hyperspectral imagery for mineral exploration: comparison of results from SFSI and AVIRIS. *Can J Remote Sens* 29:99–110. <https://doi.org/10.5589/m02-085>
- Novak ID, Soulakellis N (2000) Identifying geomorphic features using LANDSAT-5/TM data processing techniques on Lesvos, Greece. *Geomorphology* 34:101–109. [https://doi.org/10.1016/s0169-555x\(00\)00003-9](https://doi.org/10.1016/s0169-555x(00)00003-9)
- Pendžerkovski J, Rakićević T, Ivanovski T, Gjuzelkovski D (1963) Geological map and guide book for sheet Kožuf (K 34–105), Basic Geological Map of SFRJ 1:100,000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 47
- Petrović D (2015) Spatial position of the East Vardar ophiolite: geophysical-geological model and implications on its geodynamic evolution. University of Belgrade (In Serbian with English abstract), Faculty of Mining and Geology
- Petrović D, Cvetkov V, Vasiljević I, Cvetković V (2015) A new geophysical model of the Serbian part of the East Vardar ophiolite: Implications for its geodynamic evolution. *J Geodyn* 90:1–13. <https://doi.org/10.1016/j.jog.2015.07.003>
- Puissant A, Rougier S, Stumpf A (2014) Object-oriented mapping of urban trees using Random Forest classifiers. *Int J Appl Earth Obs Geoinf* 26:235–245. <https://doi.org/10.1016/j.jag.2013.07.002>
- Rakičević T, Stojanov R, Arsovski M (1965) Geological map and guide book for Sheet Prilep (K 34–92), Basic Geological Map of SFRJ 1:100,000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 65
- Rakičević T, Dumurdžanov N, Petkovski P (1969) Geological map and guide book for Sheet Štip (K 34–81), Basic Geological Map of SFRJ 1:100,000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 70
- Rakičević T, Penderovski J, Kovačević M (1973) Geological map and guide book for sheet Strumica (K 34–94), Basic Geological Map of SFRJ 1:100000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 69
- Resimić-Šarić K, Cvetković V, Balogh K, Koroneos A (2006) Main characteristics of ophiolitic complexes within the eastern branch of the Vardar Zone Composite Terrane in Serbia. In *International Symposium on the Mesozoic ophiolite belts of the northern part of the Balkan Peninsula*, Belgrade (Serbia) and Banja Luka (Bosnia and Herzegovina), May 31st–June 6th, 2006, Abstracts. Faculty of Mining and Geology, University of Belgrade, Belgrade, pp 112–115
- Robertson A, Karamata S, Šarić K (2009) Overview of ophiolites and related units in the Late Palaeozoic–Early Cenozoic magmatic and tectonic development of Tethys in the northern part of the Balkan region. *Lithos* 108:1–36. <https://doi.org/10.1016/j.lithos.2008.09.007>
- Rodriguez-Galiano VF, Chica-Olmo M, Chica-Rivas M (2014) Predictive modelling of gold potential with the integration of multi-source information based on random forest: a case study on the Rodalquilar area, Southern Spain. *Int J Geogr Inf Sci* 28:1336–1354. <https://doi.org/10.1080/13658816.2014.885527>
- Schefer S, Fügenschuh B, Schmid SM, Egli D, Ustaszewski K (2007) Tectonic evolution of the suture zone between Dinarides and Carpatho-Balkan: Field evidence from the Kopaonik Region, Southern Serbia. In: *Geophysical Research Abstracts*, Vol. 9, 03891. European Geosciences Union
- Schetselaar EM, Ryan JJ (2008) A Remote Predictive Mapping Case Study of the Boothia Mainland Area, Nunavut, Canada. In: Harris JR (ed) *Remote Predictive Mapping: An Aid for Northern Mapping*. Geological Survey of Canada, Open File 5643, pp 261–281. <https://doi.org/10.4095/226028>

- Shayeganpour S, Tangestani MH, Gorsevski PV (2021) Machine learning and multi-sensor data fusion for mapping lithology: A case study of Kowli-kosh area, SW Iran. *Adv Space Res* 68:3992–4015. <https://doi.org/10.1016/j.asr.2021.08.003>
- Sposito G (1989) *The Chemistry of Soils*. Oxford University Press
- Steinbach M, Tan PN (2009) kNN: k- Nearest Neighbors. In: *The top ten algorithms in data mining*, 1st edn. Chapman and Hall/CRC
- Waske B, Braun M (2009) Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS J Photogramm Remote Sens* 64:450–457. <https://doi.org/10.1016/j.isprsjprs.2009.01.003>
- Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM (2016) Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides* 13:839–856. <https://doi.org/10.1007/s10346-015-0614-1>
- Zelić M, Agostini S, Marroni M, Pandolfi L, Tonarini S (2010) Geological and geochemical features of the Kopaonik intrusive complex (Vardar zone, Serbia). *Ofoliti* 35:33–47
- Zuo R, Carranza EJM (2023) Machine Learning-Based Mapping for Mineral Exploration. *Math Geosci* 55:891–895. <https://doi.org/10.1007/s11004-023-10097-3>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Article

Ionospheric Response on Solar Flares through Machine Learning Modeling

Filip Arnaut ^{*}, Aleksandra Kolarski , Vladimir A. Srećković  and Zoran Mijić 

Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; aleksandra.kolarski@ipb.ac.rs (A.K.); vlada@ipb.ac.rs (V.A.S.); zoran.mijic@ipb.ac.rs (Z.M.)

* Correspondence: filip.arnaut@ipb.ac.rs

Abstract: Following solar flares (SF), the abrupt increase in X-radiation and EUV emission generates additional ionization and higher absorption of, e.g., electromagnetic waves in the sunlit hemisphere of the Earth's ionosphere. The modeling of the ionosphere under solar flares are motivated by new observations with spacecrafts, satellites, and ground-based measurements. The estimation of modeling parameters for the ionospheric D-region during SF events poses a significant challenge, typically requiring a trial-and-error approach. This research presents a machine learning (ML) methodology for modeling the sharpness (β) and reflection height (H') during SF events occurred from 2008 to 2017. The research methodology was divided into two separate approaches: an instance-based approach, which involved obtaining SF parameters during the peak SF, and a time-series approach, which involved analyzing time-series data during SFs. The findings of the study revealed that the model for the instance-based approach exhibited mean absolute percentage error (MAPE) values of 9.1% for the β parameter and 2.45% for the H' parameter. The findings from the time-series approach indicated that the model exhibited lower error rates compared to the instance-based approach. However, it was observed that the model demonstrated an increase in β residuals as the predicted β increased, whereas the opposite trend was observed for the H' parameter. The main goal of the research is to develop an easy-to-use method that provides ionospheric parameters utilizing ML, which can be refined with additional and novel data as well as other techniques for data pre-processing and other algorithms. The proposed method and the utilized workflow and datasets are available at GitHub.



Citation: Arnaut, F.; Kolarski, A.; Srećković, V.A.; Mijić, Z. Ionospheric Response on Solar Flares through Machine Learning Modeling. *Universe* **2023**, *9*, 474. <https://doi.org/10.3390/universe9110474>

Academic Editor: Vladislav Demyanov

Received: 5 October 2023
Revised: 2 November 2023
Accepted: 3 November 2023
Published: 7 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: regression; kernel density estimation; solar flares; ionospheric data; ionospheric parameters modeling

1. Introduction

The low ionosphere, located between approx. 50 and 90 km above the Earth's surface [1], varies in ionization with the solar flow [2] and displays the effects of solar flares (SF) [3–5]. Crucial for its modeling, the ionosphere parameters are directly/indirectly measured with sounding rockets, which offer a single measuring point [6] and are associated with high operational costs [7–9]. Consequently, the utilization of the very low frequency (VLF) approach is commonly employed [10,11]. Usually, both the VLF signal's amplitude and phase display the impacts of SF occurrences that follow X-ray flux, in most cases often seen as increased values during such events, but with patterns highly influenced with VLF trace geometry [11–14]. The D-region of the ionosphere is defined by two parameters: sharpness (β) in km^{-1} and reflection height (H') in km [15], also known as Wait's parameters or waveguide ionospheric parameters. The electron density (ED) may be computed for both undisturbed and disturbed ionospheric conditions using an equation displayed by Wait and Spies [16]. Typically, the estimation of the ionospheric parameters is performed by the utilization of the Long Wavelength Propagation Capability (LWPC) software LWPC. Computer Programs for Assessment of Long-Wavelength Radio

Communications, V2.1. Available online: <https://github.com/space-physics/LWPC> (accessed on 2 February 2023). [17], which is a challenging task due to the complex nature of ionospheric modeling during disturbances, especially related to SF events of extreme intensity [18]. Figure 1 gives an example of one such energetic event, i.e., X17.2 SF that occurred on 28 October 2003 with peak soft X-ray Irradiance at 11:10UT, third in strength within the entire 23rd solar cycle. The response of the mid-latitude lower ionosphere over the Balkan region to this energetic SF, as recorded by BEL VLF station in Belgrade (Serbia), is given in Figure 2, together with the modeling results of the obtained electron density estimated by the use of a numerical procedure developed by Srećković et al. [19], relying on approximate equations for obtaining ionospheric parameters directly from measured X-ray data. In many cases of intense SFs, obtaining lower ionospheric response based on VLF signal technology indirectly from amplitude and phase perturbations of monitored VLF signals through classical approach using LWPC software can be problematic due to many factors, in first line including model limitations/restrictions and factors related to the geometry of Great Circle Paths of VLF signals. So, developing new approaches for retrieving necessary data for estimation of electron density profiles during such extreme events is of great importance.

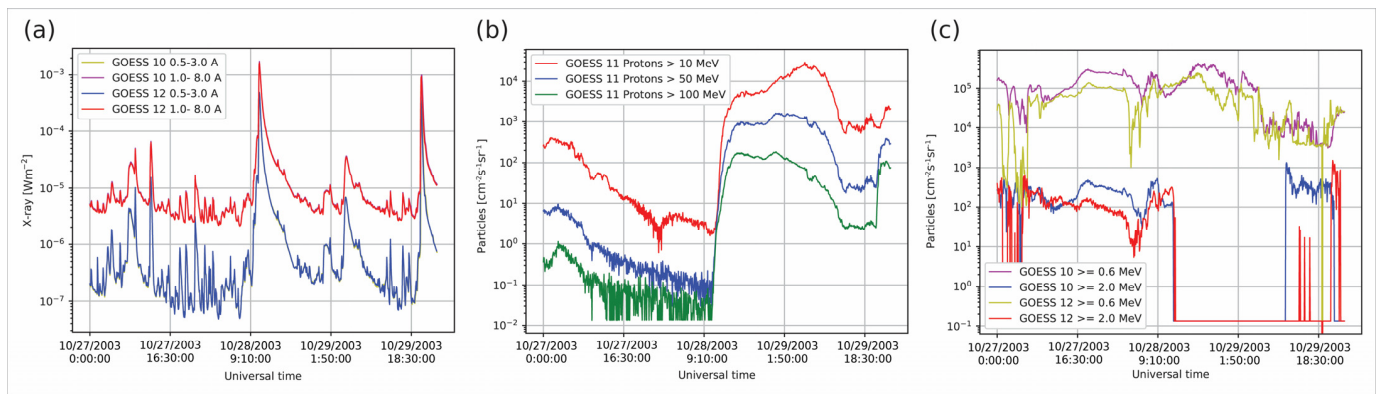


Figure 1. Measured data from GOES satellites for period 27–29 October 2003: variation in X-ray flux (a), proton flux (b), and electron flux (c) (data available at <https://www.ncei.noaa.gov/data/goes-space-environment-monitor/access/avg>, accessed: 2 February 2023).

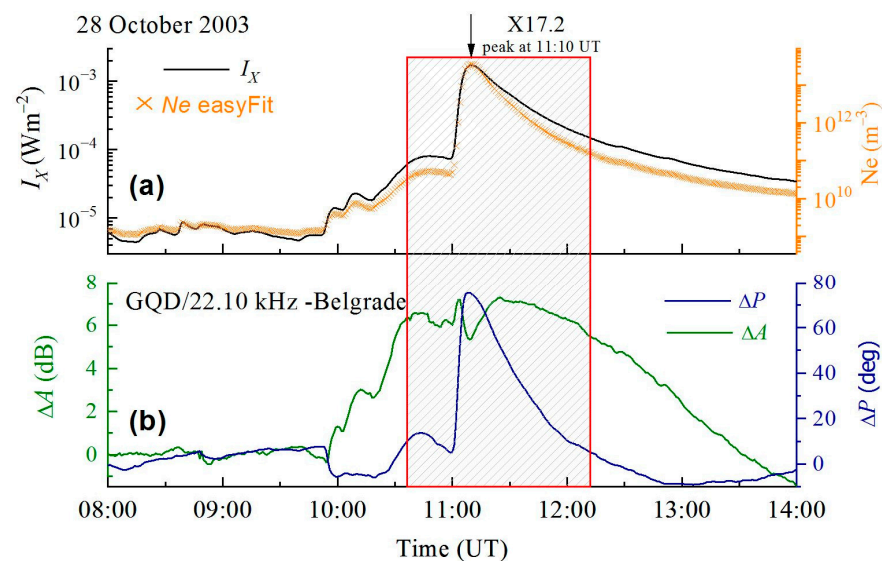


Figure 2. (a) Variation in X-ray flux from GOES-15 satellite (solid black line) and corresponding electron density (red crosses) at a reference height of 74 km versus universal time on 28 October 2003

during an extreme event of X17.2 class SF, obtained through the numerical procedure presented in Srećković et al. [20], (b) GQD signal amplitude and phase perturbations induced by X17.2 class SF, as recorded in Belgrade (Serbia).

Previous research has examined the estimation of ionospheric parameters by employing artificial neural networks (ANNs) and utilizing synthetic data created using the LWPC software [21]. Furthermore, Alpatov et al. [22] demonstrated that the inherent characteristics of the ionosphere necessitate the utilization of statistical and probabilistic approaches for effective ionospheric modeling. Gross and Cohen [21] developed an ANN target function that was designed to accept a single time-step of VLF amplitude and phase as the input and generate the waveguide parameters as the output. Subsequent research revealed that the model underwent expansion to encompass nighttime conditions [23]. In relation to the use of ANNs, the tuning of several hyperparameters, such as the number of epochs, learning rate, number of hidden layers, and number of nodes per hidden layer, presents challenges in the process of fine-tuning and renders them more vulnerable to overfitting. In this study, the primary focus will not be on ANNs. Instead, attention will be directed on comparatively simpler yet still-efficient ML models.

The research objectives have been divided into two distinct components. The initial step involves assessing the feasibility of deploying an ML model for the accurate prediction of β and H' parameters. The primary objective can be categorized into two distinct methods: the instance-based approach and the time-series approach. In the instance-based approach, a specific set of features, such as X-ray irradiance, SF class, amplitude changes (ΔA), and phase shifts (ΔP), are utilized to effectively provide input to a ML regression model to determine the peak SF effect on waveguide parameters. On the other hand, the time-series approach involves training the model using a limited number of days' worth of X-ray irradiance and statistical data derived from X-ray irradiance. Afterwards, the model is tested on subsequent days to forecast waveguide parameters, i.e., to continuously determine waveguide parameters as a time-series. The secondary aim of this study is to determine the need for utilizing oversampling techniques with a post hoc analysis of oversampling. The advantages of employing an ML technique for the estimation of ionospheric parameters are evident in the (computationally) time consuming process involved in the LWPC modeling of those parameters. Four frequently employed ML algorithms are employed, namely, Random Forest (RF), Decision Tree (DT), K-nearest neighbors (KNN), and XGBoost (XGB). The datasets used in this study, along with certain components of the workflow, are available for access on GitHub (see Supplementary Material).

As ionospheric research focuses on radio signal propagation and influences on the great diversity of space-borne and ground-based technological systems [24–27], the overall aim of this study is to provide an alternative approach allowing the wider scientific community to obtain crucial ionospheric parameters being affected by SFs, promptly and efficiently. As existing and standard methods to determine ionospheric parameters are based on a trial-and-error process, e.g., the method presented in [17], there is a need to develop other methods to determine the β and H' like easyFit [19] and FlareED [20]. Our aim is to develop a user-friendly method that can be refined with new data, other pre-processing techniques, and algorithms, etc. In future research, if the ML method provides satisfactory accuracy and stability, we can compare the ML method with the aforementioned methods (e.g., under different SF classes, etc.).

2. Materials and Methods

The research methodology was divided into two parts. The first part involved an instance-based approach, where instances of SF peaks and their associated features (such as ΔP , ΔA , X-ray irradiance, etc.) were provided. Based on these features, waveguide parameters were to be determined using ML techniques. The second methodology employed in this study involved the utilization of a time-series approach. Specifically, a time-series dataset of X-ray irradiance was provided, and various statistical features were computed

based on the X-ray irradiance values. The determination of waveguide parameters was conducted using an ML model, based on the aforementioned features.

2.1. Instance-Based Approach

The data analysis initiated with the utilization of the original dataset, which contained 212 unique observations conducted throughout the course of various SF’s. The initial dataset was utilized to generate synthetic data points in subsequent stages of the workflow. However, the validation data, consisting of an extra 45 data points, was excluded from the original dataset prior to the pre-processing phase (Figure 3). The initial dataset included observations pertaining to the X-ray irradiance, the difference between the amplitude and phase of the VLF signal, and the ionospheric parameters namely. The initial data cleaning process involved randomizing the data to obtain a representative sample and transforming the X-ray irradiance to assure accurate Kernel Density Estimation (KDE) without any errors, and subsequent de-transformation was performed before the ML modeling.

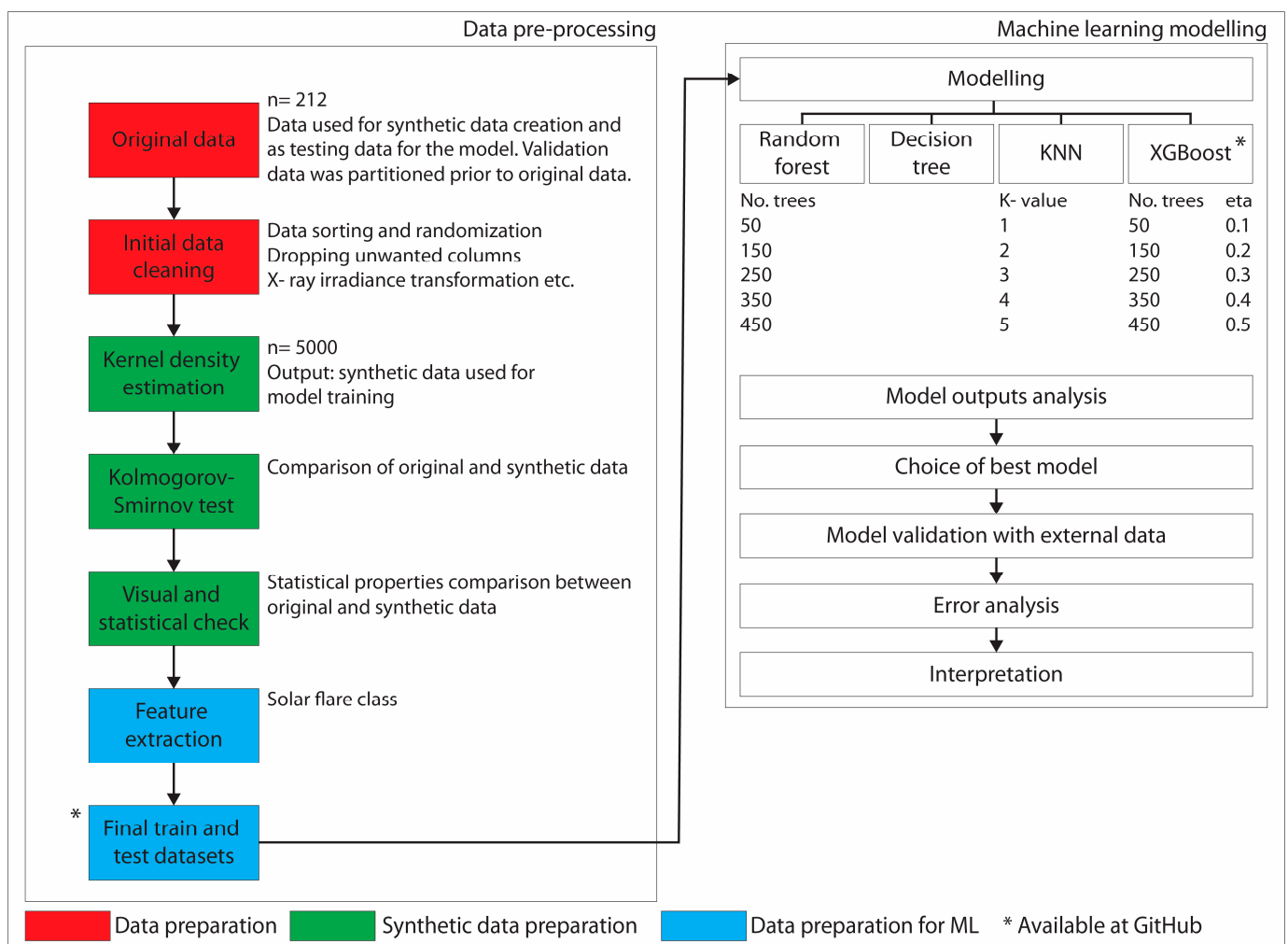


Figure 3. Data preparation, synthetic data preparation, and machine learning modeling workflow.

The KDE serves as the initial stage in the synthetic data preparation procedure, aiming to generate a more extensive collection of samples (in this case, 5000) to facilitate the training of the ML model. The KDE method is utilized to estimate the underlying distribution of the original data. Subsequently, multiple samples are generated from this estimated distribution, therefore producing an increased number of samples. In order to verify the accuracy of the KDE’s estimation and sampling from the original data distribution, it is necessary to incorporate two more stages. This involves employing the Kolmogorov–

Smirnov test (KST), a non-parametric statistical test used to assess whether two datasets follow the same distributions [28]. The interpretation of the findings of the KST involves the evaluation of the p -value. If the p -value exceeds the predetermined significance level of 0.05, it is not possible to reject the null hypothesis, i.e., the two datasets originate from the same distribution, and vice versa [29]. The second stage validating the KDE involves doing a visual and statistical examination. This entails visualizing the distributions of both the original and synthetic data for all features and target variables, as well as calculating and comparing various statistical metrics such as the mean, median, mode, skewness, and kurtosis, etc.

If the results of the preceding statistical analysis are deemed to be valid, then the initial stage in data preparation for ML involves feature extraction. Feature extraction, also known as feature discovery, is a procedure in ML that involves adding to a dataset additional feature(s). This is performed with the objective of increasing the number of features available, hence potentially enhancing the performance of the ML model. In this particular instance, the sole addition made was the inclusion of the SF class, wherein solely the alphabetical component of the SF class (B, C, M, or X) was retained and afterwards converted into a numerical representation (e.g., 1, 2, 3, or 4).

The synthetic data for training and the original data for testing were used as inputs to the ML modeling process. Four distinct algorithms were employed in this study, namely, Random Forest (RF) [30], Decision Tree (DT) [31], K-nearest neighbors (KNN) [32,33], and XGBoost (XGB) [34]. With the exception of the DT algorithm, each of the algorithms applied in this study required the tuning of at least one hyperparameter (RF and KNN) or two hyperparameters (XGB). The hyperparameter that determines the performance of the RF model is the number of trees. Additionally, the number of trees is also a hyperparameter for the XGB algorithm. On the other hand, for the KNN algorithm, the hyperparameter is the value of K , i.e., the number of nearest neighbors. The XGB method has an extra hyperparameter known as the learning rate (LR). The hyperparameter values used for the algorithms employed are presented in Figure 3.

The outputs generated by all models were employed in order to choose the most optimal model, as determined by the mean and maximum percentage errors computed on the testing dataset. Following the selection of the optimal model based on the in-distribution training and testing, model validation was performed using the validation data, and, subsequently, an error analysis was carried out.

One noteworthy aspect of this study is the inclusion of an extra target variable. Unlike conventional regression or classification models that include n features and a single target variable, this research incorporates two output variables, namely, ionospheric parameters (β and H'), as seen in Figure 4. The utilized methods incorporate SF class, ΔP , ΔA , and X-ray irradiance as their respective features, whereas the output parameters are waveguide parameters.

The evaluation metrics employed in this study to distinguish between models, specifically to determine which model performs better, are the mean absolute percentage error (MAPE) and the maximum absolute percentage error observed over all occurrences of the testing or validation datasets. Furthermore, the model's outputs may be further analyzed by displaying absolute errors, visual representations of predicted and true observations, etc.

2.2. Time-Series Approach

The time-series based approach necessitated minimal data pre-processing, specifically in terms of data cleaning and feature extraction, with the utilization of statistical features. The time-series based approach employed a methodology that closely resembled the instance-based approach. In this approach, the modeling process utilized the best model and (hyper)parameters determined from the instance-based approach. The dissemination of results exhibited similarities to the instance-based approach.

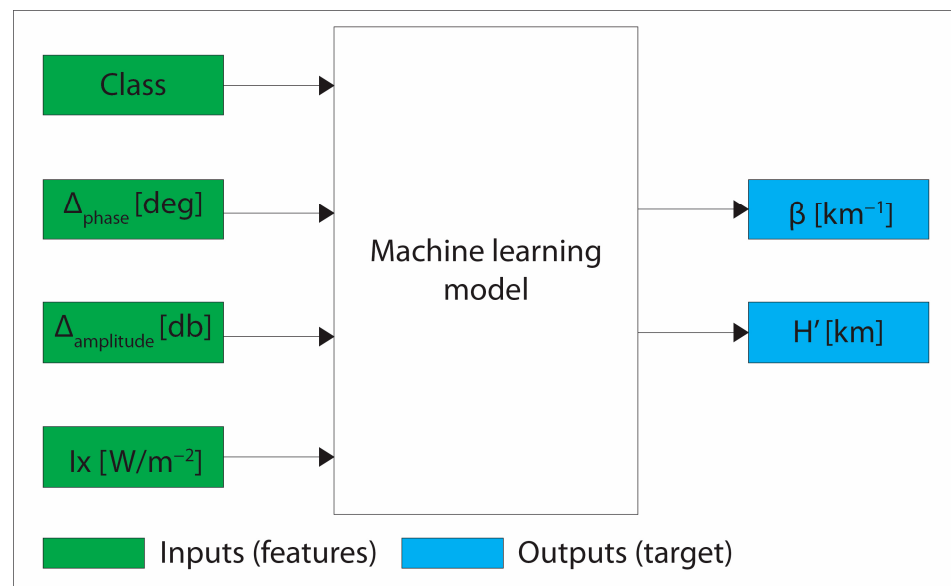


Figure 4. Diagram of features and targets for multi-output machine learning modeling.

3. Results and Discussion

3.1. Instance-Based Approach

3.1.1. Data Pre-Processing

The original dataset comprises 212 data points representing SF events from 2008 to 2017. Each data point includes the recorded time of measurement, transmitter information, X-ray irradiance, ΔA and ΔP , ionospheric parameters, and the corresponding computed electron density. The validation dataset, initially removed from the original dataset and subsequently excluded from any subsequent analysis, was exclusively employed for model validation. It comprises 45 data points collected between 2004 and 2017, with a notable concentration of data points occurring during the periods of 2004–2006 and 2014–2017.

The comparison of SF intensities between the initial dataset used for ML model testing and the validation dataset reveals the distribution of each class of SF as a percentage. In the initial dataset, a small proportion of SFs belong to the X-class, specifically accounting for 1.89% (four occurrences). Conversely, in the validation dataset, this number significantly increases to 20% (nine instances). Similar observations can be made about the M-class SFs in the datasets. Approximately 18% (38 occurrences) of the original dataset consist of M-class SFs, but the validation dataset exhibits an almost doubled percentage of 35.56% (16 instances). The proportion of C-class SFs is larger in the original dataset (79%, 169 instances) compared to the validation dataset (42%, 19 instances). Ultimately, the B-class SFs in both datasets exhibit a similar level of outcomes, with a prevalence of 0.47% in the original dataset and 2.22% in the validation dataset. It is worth noting that both datasets contain only a single occurrence of B-class SFs.

The allocation of classes within the original dataset, specifically the testing and validation subsets, was conducted in a deliberate manner to ensure that the model undergoes initial testing with a distribution that is known and expected to yield higher performance. This approach allows for the selection of the most optimal model based on the original testing dataset. The aforementioned best model will then undergo further testing, using an unfamiliar distribution, i.e., the validation dataset. As a result, the evaluation metrics obtained from the validation dataset will accurately reflect the model's predictive performance.

The output of the KDE is a synthetic dataset that is drawn from the identical distribution as the original dataset, however, with an increased number of data points, namely, 5000 in this instance. The verification of the KDE was conducted using the KST. The results of the KST test demonstrated that the original dataset and the synthetic dataset exhibited similar distributions for all five parameters, namely, X-ray irradiance, ΔA , ΔP , β , and H' . Further

examination was conducted by visually inspecting and comparing the two distributions (Figure 5).

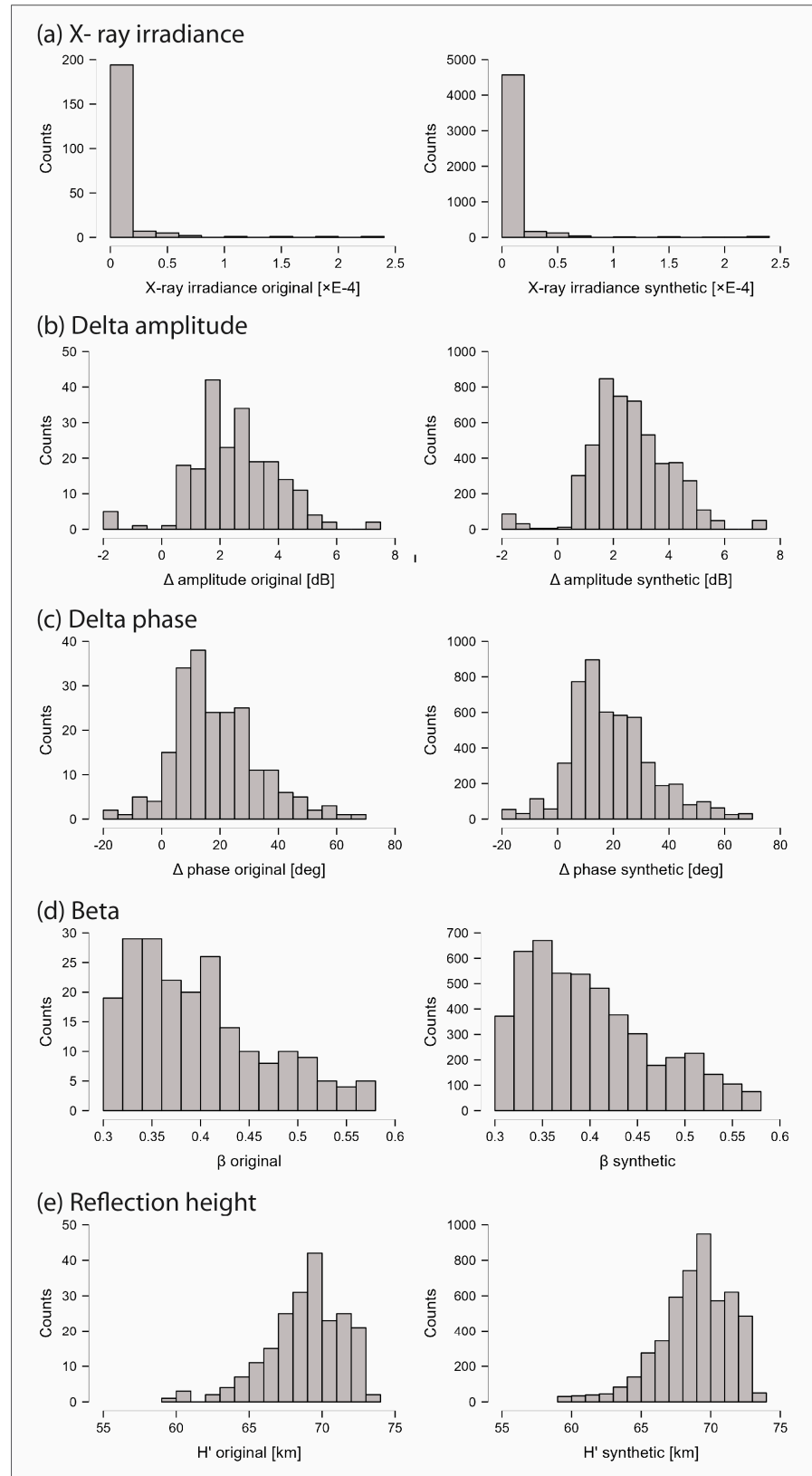


Figure 5. Comparison of original and synthetic data distributions for all features and targets for machine learning modeling.

Based on the analysis of Figure 5, it is apparent that all synthetic distributions exhibit a high degree of resemblance to the original distributions. The X-ray irradiance distribution in the original dataset had a pronounced skewness, with a tail extending towards positive values. As a result, the synthetic data also display a similar characteristic, with a severely skewed distribution towards positive values. This can also be expressed by using the skewness and kurtosis parameters. In the case of the original data, these parameters assume values of 6.178 and 42.932, respectively. Conversely, for the synthetic data distribution, the corresponding values are 6.22 and 42.622. These values, along with the KST, the visualization of the distribution, further confirm the efficacy of the KDE and the validity of the generated samples. A parallel analysis may be conducted on the target variables, such as the β parameter. In the original distribution, the mean value is 0.401, which is consistent with the value seen in the synthetic distribution. The median values for the original and synthetic distributions exhibit a little disparity, namely, 0.383 and 0.386, respectively. The H' parameter exhibits comparable mean and median values between the original and synthetic distributions, which highlights that the synthetic data have been generated correctly.

Following the successful application of data oversampling techniques, the training dataset, i.e., the oversampled dataset, was prepared. Additionally, the testing dataset, which consists of the original, non-oversampled, measured data, and the validation dataset, comprising out-of-sample, measured data that was not utilized in the data pre-processing phase, were also finalized. These datasets were then employed for the purpose of ML modeling.

3.1.2. Initial Phase of Machine Learning Modeling

The first iteration of ML modeling involved the utilization of four distinct algorithms (RF, DT, KNN, and XGB). In all, a set of 16 models were created and evaluated. The evaluation metrics (MAPE) for both target variables is presented in Figure 6, with the upper panel representing the β parameter, and the lower panel representing the H' parameter. Based on Figure 6, it is apparent that the algorithms DT and KNN, throughout all iterations, may be disregarded since they exhibit greater MAPE values compared to RF and XGB. On the other hand, both RF models and XGB models exhibit relatively comparable MAPE values. The RF models have constant MAPE values for both target variables. Consequently, additional analysis may be conducted to identify the most optimal RF model for exclusion. The model with 250 trees was selected as the optimal RF model across all iterations, based on the maximal percentage error observed for both target variables (11.5% for β and 1.1% for H'). In the initial round of modeling, the XGB model with 150 trees was identified as the most optimal choice. This particular model exhibited MAPE values of about 0.8 and 0.029% for the respective target variables. It is important to emphasize that the MAPE values should be interpreted cautiously due to the fact that the model was evaluated using the original dataset (in-sample model evaluation). When applying the model to new data (out-of-sample validation data), it is likely that the MAPE values will be higher. However, the purpose of this modeling is to identify the best overall model.

The XGB method was employed for the second stage of modeling because of the inclusion of an extra hyperparameter, specifically the LR. The XGB additional modeling was conducted using a predetermined number of trees (150) and a varying LR value ranging from 0.1 to 0.5, with increments of 0.1. A comparison was conducted between the optimal XGB model obtained in the initial phase of modeling, characterized by a learning rate of 0.3, and the model derived from the additional phase of XGB modeling, featuring a learning rate of 0.2. The MAPE values indicate that both models exhibit comparable error rates. Specifically, the LR = 0.3 model has MAPE values of 0.8 and 0.029 for β and H' , whereas the LR = 0.2 model shows MAPE values of 0.71 and 0.033 for β and H' , respectively. In contrast, there is a disparity in the maximal percentage error values between the LR = 0.3 model, which exhibits a maximal percentage value of 4.6% for the β parameter, and the LR = 0.2 model, which demonstrates a maximal percentage value of 3.29% for the same parameter. The XGB model with 150 trees and a LR value of 0.2 was selected as the optimal model, both within the XGB models and in comparison to the other models. A comparison was

conducted between the optimal XGB model and the optimal RF model, revealing disparities in both the MAPE and the maximal percentage error. Specifically, the RF model exhibited maximal percentage errors of 11% for the β value, whereas the XGB model demonstrated values of 3%. This distinction was crucial in finding the most optimal model overall.

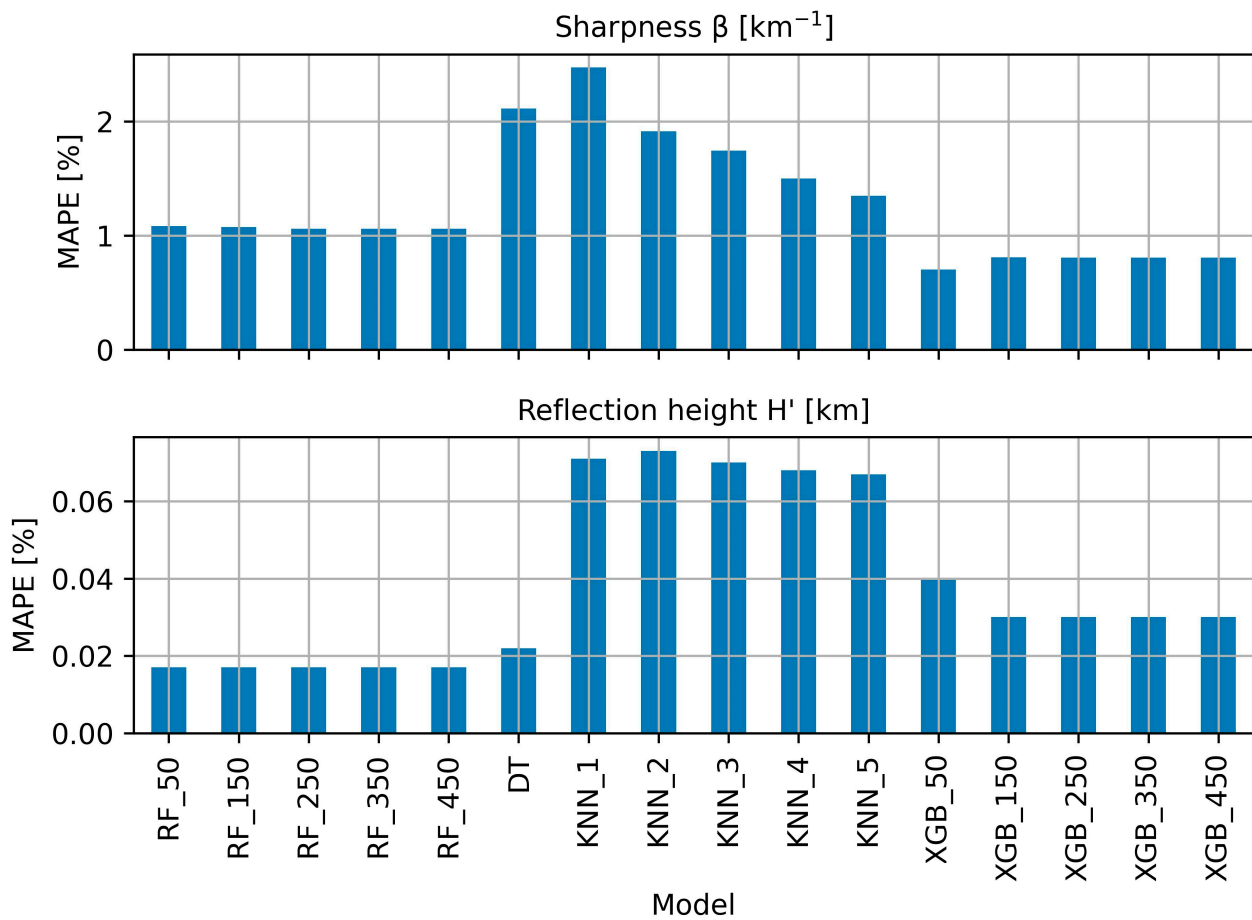


Figure 6. Mean absolute percentage errors for sharpness (**upper** panel) and reflection height (**lower** panel) for the initial phase of modeling; MAPE—Mean absolute percentage error; RF—Random Forest; DT—Decision tree; KNN—K-nearest neighbors; XGB—XGBoost; Adjacent number to model names (50–450 or 1–5) is the hyperparameter for the given model.

3.1.3. Model Validation

The process of model validation involved the utilization of an out-of-sample dataset that was initially omitted from the original dataset at the beginning of the analysis. MAPE values for both β and H' exhibit greater values compared to the training dataset, as anticipated, with values of 9.1% and 2.45%, respectively. The maximum percentage error values for both parameters are around 38.8% and 12.2%, respectively. Figure 7 presents the error distribution for the parameters, denoted as Figure 7a,b, as well as the error distribution for the averaged error (Figure 7c), which is calculated as the average of both percentage errors.

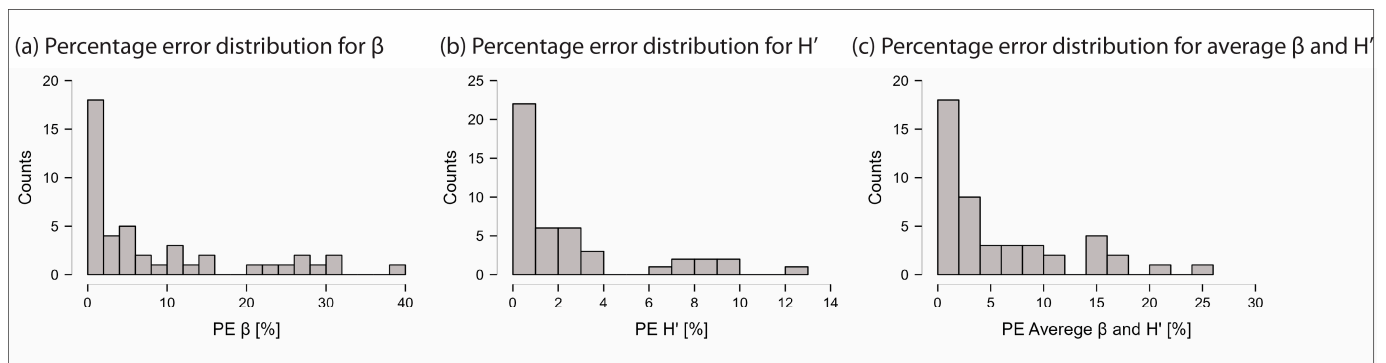


Figure 7. Model validation mean absolute percentage error distributions. (a) Sharpness; (b) Reflection height; (c) Averaged mean absolute percentage error for sharpness and reflection height; PE—Percentage error.

It is worth mentioning that the distributions for β and H' exhibit a pronounced skew towards higher values, indicating that a majority of the values correspond to lower error rates. Approximately 66% of the data points pertaining to the β parameter exhibit a percentage error below 10%, whereas approximately 55% of the data points exhibit a percentage error below 5%. In contrast, it can be observed that around 97% of the data points exhibit an error rate of less than 10% for the H' parameter, whereas 82% of the data points have error rates below 5%. This suggests that the MAPE is significantly impacted by a small number of high percentage errors. This observation is supported by the fact that there are nine occurrences where the percentage error values for β exceed 20%, and three cases where they exceed 30%. In contrast, the H' parameter exhibits a single outlier, namely, the outlier characterized by a percentage inaccuracy of over 12%. This indicates that the model can produce relatively satisfactory β and H' values, with the exception of a few significantly exaggerated errors, especially in the β parameter case. It is important to acknowledge the absolute range of errors generated by this method (MAE). The MAE for the β value is around 0.039 km^{-1} , whereas, for the H' value, the MAE is around 1.64 km.

Further model validation may be conducted by visually analyzing the predicted and real data for both parameters, as seen in Figure 8. Linear fit line was used to construct a line of best fit across the predicted and observed data. The coefficient of determination (CD) for the β parameter was about 0.67, whereas, for the H' parameter, the fit was greater at 0.8. The CD had values that were reasonably satisfactory. Significant outlier data points are evident in both cases. For instance, a notable outlier for the β parameter is observed (upper panel on Figure 8, green rectangle), where the predicted value was about 0.38 km^{-1} , whereas the real value was approximately 0.53 km^{-1} , resulting in an error rate of almost 28%. The observed data point can be associated with a C4.8 SF event, characterized by a very modest ΔA of around 0.06 dB, whereas the ΔP measured around 26 degrees. In the given case, the H' parameter exhibited a minimal percentage error of 0.82%, with a true value of 70 km and a predicted value of 70.6 km. Another noteworthy example is observed in the lower section of Figure 8 (green rectangle), where a C9.6 SF yielded a ΔA measurement of 5.13 dB and a ΔP measurement of 50.04 degrees. The predicted value was 61.4 km, whereas the actual value was 70 km. In a prior case, there was an incorrect prediction of the β parameter, but the H' parameter was predicted with a significantly low error rate. In the current circumstance, we observe a similar scenario where the β parameter was predicted with a minimal error rate of 1.2% (0.57 km^{-1} reality against 0.563 km^{-1} predicted).

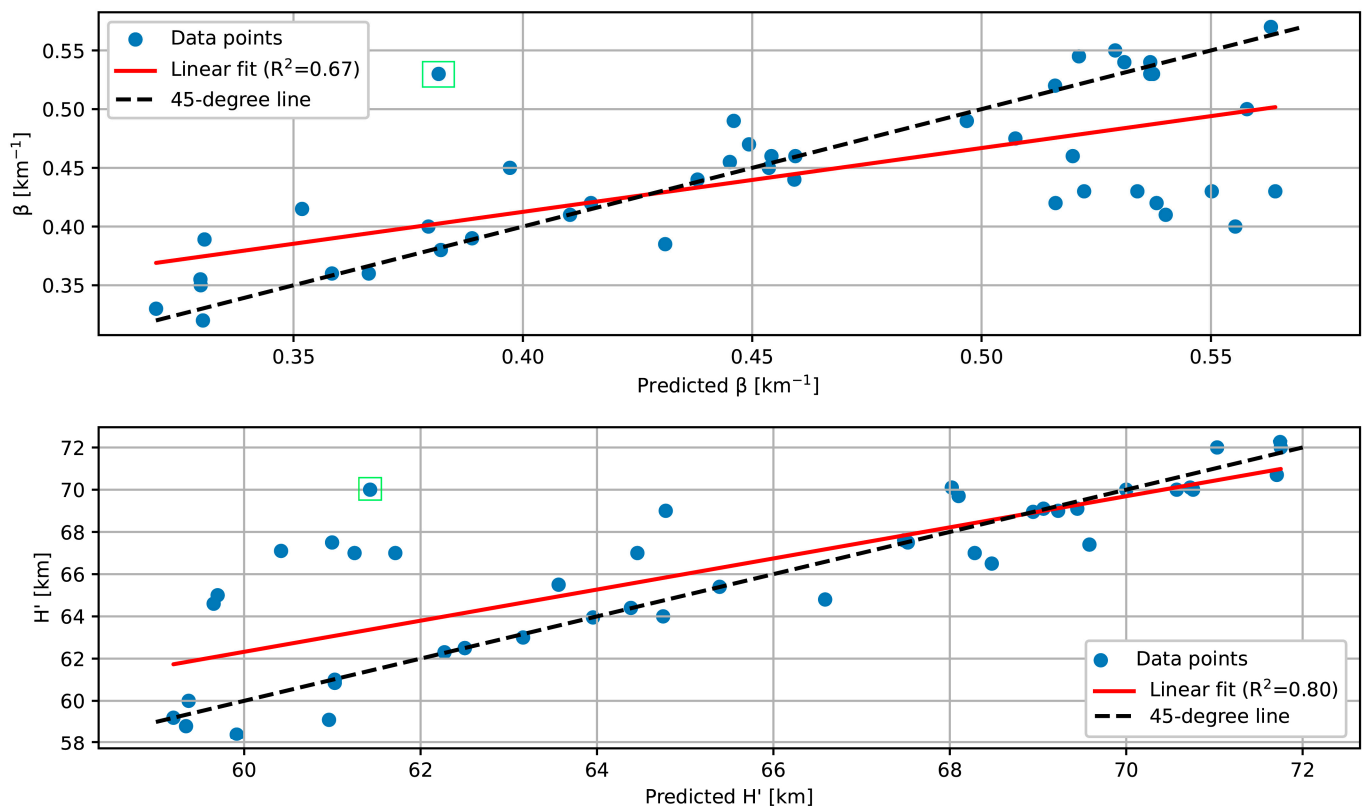


Figure 8. Predicted and observed value for sharpness (**upper** panel) and reflection height (**lower** panel) for the instance-based approach.

The inclusion of a residuals plot in conjunction with the predicted β and H' values provide valuable insights into the predictive performance of the model (Figure 9). Specifically, the upper panel of Figure 9 which represents the residuals plot of the β parameter, does not exhibit any noticeable pattern, as well as the residual distribution is a normal distribution. Most of the residuals are concentrated within the range of -0.05 to 0.05 km^{-1} , with the largest outlier being the previously mentioned discrepancy of 0.14 km^{-1} , which is associated with a C4.8 SF. In contrast, the residuals plot for the H' parameter exhibits a discernible pattern that aligns with a decrease in residuals as the predicted H' parameter increases. The lower panel of Figure 9 reveals that the predicted value of 62 km and beyond exhibit residuals within the range of -2 to 2 km. However, before these values, the residuals are greater, ranging from -4 to 6 km. This suggests that the model's predictions of the H' parameter over 62 km (with the exception of two cases) are associated with reduced error rates compared to its predictions of H' values below 62 km. This observation is further substantiated by examining the average percentage error rates for predictions made below 62 km and those made beyond 62 km. The MAPE for predictions above 62 km is around 1.23%, but predictions below 62 km exhibit a greater MAPE value at 5.15%.

The XGB model exhibited a minor bias in its predictions of H' parameters above a certain threshold value of 62 km. It is important to acknowledge, for future studies employing a similar methodology, that there exists a possible association between the SF class and the predicted H' parameter. Specifically, among all the predictions in the validation dataset that are below 62 km, 8 out of 9 X-class solar flares are shown to have predictive values of H' below 62 km. The observation can be understood as the model exhibiting elevated error rates when predicting ionospheric parameters on X-class SFs. This observation is supported by the fact that the validation dataset contains a significantly higher proportion of X-class SFs (20%) compared to the testing dataset (1.89%). The primary objective of the validation dataset was to serve as an out-of-sample test for the model, in which the original distribution of the SF classes is not present. In this regard, the validation

dataset proved to be effective in achieving its intended purpose. The possible bias shown in this research might be mitigated by increasing the number of original samples and refraining from employing an oversampling approach.

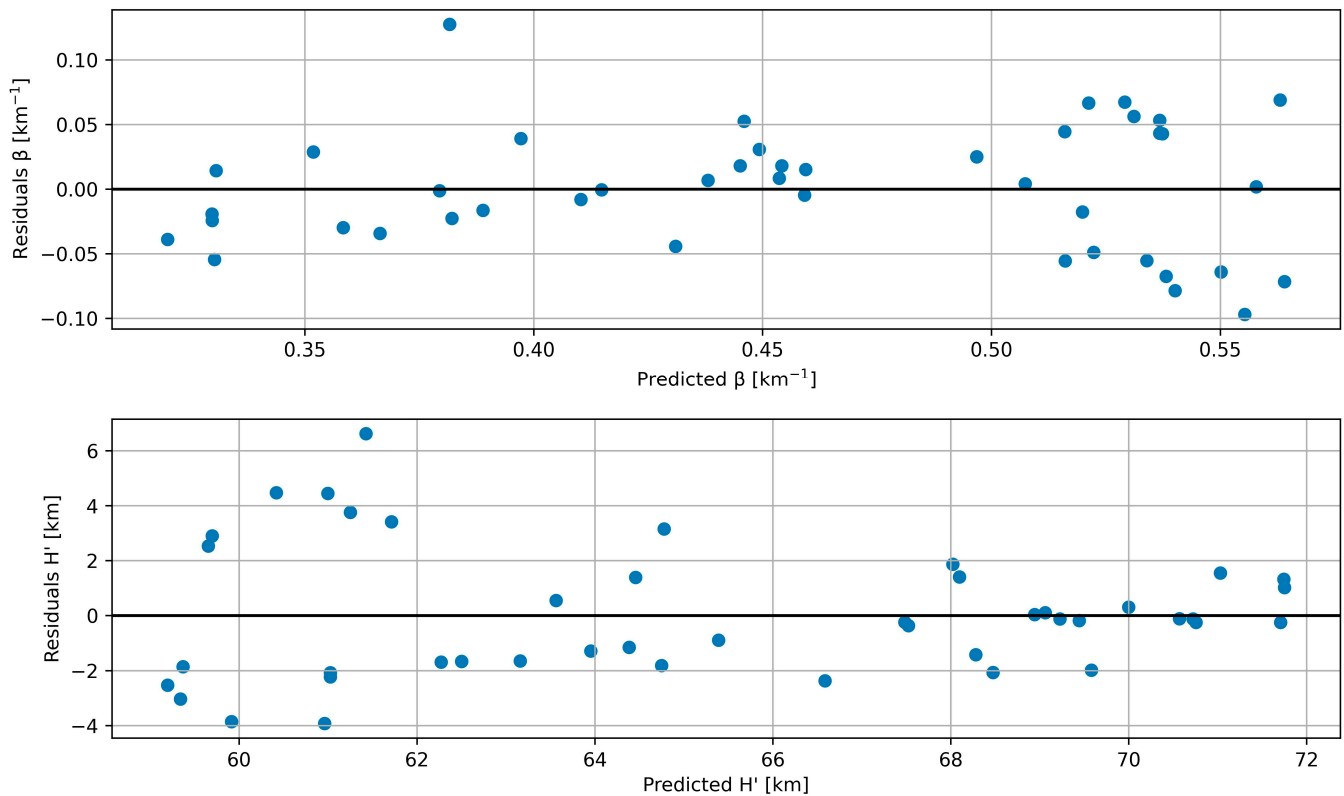


Figure 9. Residual and predicted plot for sharpness (**upper** panel) and reflection height (**lower** panel) for the instance-based approach.

A potential alternative approach, while hardly employed, is the balanced distribution of SF class features, ensuring an equal representation of X-, M-, and C-class SFs. The utilization of class balancing approaches in ML regression is not commonly employed, as these techniques are often utilized in ML classification tasks and are applied to the target variable. However, it may be worthwhile to investigate this approach further in future research, as the outcomes have the potential to alleviate potential model bias.

3.1.4. Post Hoc Analysis of the Sample Size

In ML tasks, the determination of the optimal number of samples is a challenging problem that often requires the researcher to engage in iterative experimentation. When granted permission, the collection of additional samples becomes advantageous until a specific threshold is reached, beyond which enlarging the sample size does not result in a reduction in the associated model error rate. In this study, we choose to augment the original dataset by oversampling through the utilization of KDE. This approach allows us to preserve the original data distribution while simultaneously expanding the sample size. The efficacy of the KDE approach was confirmed, and the oversampling technique yielded positive results. However, it would be advantageous for future research to conduct a concise assessment of the necessity for oversampling.

The test was conducted using both the original dataset and a synthetic dataset that shared the same distribution as the original. The datasets located between the original dataset and the synthetic dataset were acquired using the Random Undersampling (RUS) technique [35]. Afterwards, the KST was conducted to assess if these datasets exhibited the same distribution as the original dataset. The intermediate datasets were generated

to represent various proportions of the synthetic dataset (e.g., 10%, 20%, etc.). The XGB method was employed for the modeling process, employing 150 trees and a LR value of 0.2. This LR value and the number of trees was previously established to be the optimal choice for the modeling process. The findings of the post hoc examination of the sample size are presented in Table 1.

Table 1. Post hoc analysis results for the sample size; MAPE—Mean absolute percentage error; PE—Percentage error; KST—Kolmogorov–Smirnov test; RUS—Random undersampling; T—True (passed the KST); NA—Not applicable.

Percentage of Synthetic Data	β (km ⁻¹)		H' (km)		Note	KST
	MAPE (%)	Max PE (%)	MAPE (%)	Max PE (%)		
4.24	8.46	39.58	2.28	9.70	Original dataset	NA
10	9.24	36.45	2.36	11.38	RUS	T
20	9.61	38.57	2.45	11.52	RUS	T
30	9.94	46.13	2.52	12.17	RUS	T
40	10.39	49.20	2.44	12.31	RUS	T
50	9.13	35.78	2.48	11.91	RUS	T
60	10.07	49.05	2.54	12.01	RUS	T
70	9.28	40.48	2.31	11.82	RUS	T
80	9.04	47.01	2.64	12.61	RUS	T
90	9.47	43.95	2.35	11.08	RUS	T
100	9.08	38.83	2.46	12.25	Full synthetic data	T
Minimum	8.46	35.78	2.28	9.70		
Maximum	10.39	49.20	2.64	12.61		
Range	1.94	13.42	0.36	2.91		

The post hoc examination of the sample size reveals that all intermediate samples, which constitute a portion of the synthetic sample, successfully passed the KST, thus they maintained the same distribution as the original and synthetic datasets. Both the β and H' exhibit a MAPE range that is less than 2%, namely, 1.94% for β and 0.36% for H' . This suggests that expanding the sample size with the KDE did not result in any improvements for the model in terms of reducing the error rate. However, it is important to note that the highest percentage error for the β parameter does not exhibit the same trend, as it spans a range of 13.42%. Furthermore, the correlation analysis revealed that there was no significant association seen between the rise in the dataset and the decrease in the evaluation metric for both parameters. The observed maximum correlation coefficient was 0.49, indicating the relationship between the percentage of synthetic data and the maximal percentage error of the H' parameter. However, this correlation coefficient does not reach a significant level, i.e., suggesting a lack of strong correlation intensity. Therefore, it may be inferred that the augmentation of the sample size, while preserving the initial distribution, did not confer any notable benefits for the ML model.

The examination of sample size in a post hoc analysis is a crucial aspect for future research, particularly when considering the utilization of oversampling techniques on smaller datasets, such as this one. This specific example demonstrates that even with a sample size of 212 data points from the original dataset, it is possible to attain error rates comparable to those observed with synthetic, oversampled data.

3.2. Time-Series Approach

3.2.1. Data Pre-Processing

The training data consisted of observations from 6 September 2017 to 9 September 2017, whereas the model testing was conducted using data from 10 September measured on a 1 min interval. The utilization of statistical attributes as features for modeling time-series ionospheric VLF data have been demonstrated in the study conducted by Arnaut et al. [36]. Similar statistical features were developed for the present research. Specifically,

the characteristics encompassed rolling mean, median, and standard deviation statistics for different window sizes (5, 20, and 60 min). Additionally, it includes the first and second differential of the data and the percentage change between adjacent data points as well as lagged values in 1–5 min intervals. The identical model, specifically XGBoost with 150 trees and a LR of 0.2, was employed in a manner consistent with the instance-based approach.

3.2.2. Machine Learning Modeling for the Time-Series Approach

Figure 10 illustrates the training and testing datasets. Specifically, the training dataset showcases a prominent SF of X9.4 class that transpired on 6 September 2017. The training dataset contains a variety of smaller SF, which contributes to its suitability for training the model. In contrast, the testing dataset also exhibits a prominent SF of X8.2 class, which transpired on 10 September 2017. The objective of time-series-based modeling for β and H' can be described as the determination of waveguide parameters during an extreme SF event, given the availability of a training dataset containing waveguide parameters for training the model. Further support for the utilization of such modeling can be derived from the observed correlation between X-ray irradiance data and both β and H' . Specifically, by employing the Spearman correlation coefficient, the correlations between X-ray irradiance and both β and H' are found to be 0.804 and -0.876 , respectively. The observed values of the correlation coefficient indicate a significant association between the waveguide parameters and X-ray irradiance, thereby suggesting their appropriateness for utilization in ML modeling.

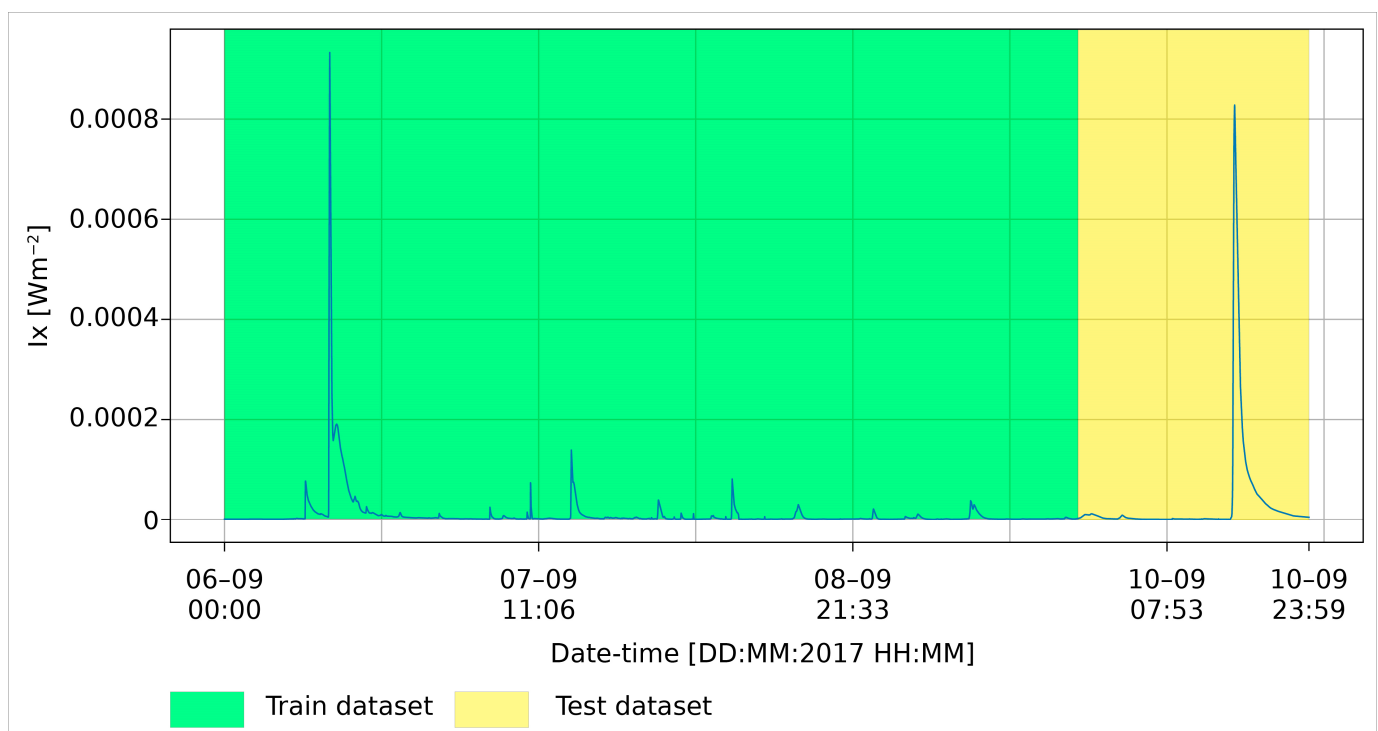


Figure 10. The training and testing dataset utilized for ML modeling of β and H' .

The outcomes of the modeling demonstrated significantly better results compared to those achieved through the instance-based approach. The MAPE value for the β parameter was 0.1%. Additionally, the maximum percentage error observed was 2.1%, corresponding to a maximum absolute error of 0.01 km^{-1} . In contrast, the findings pertaining to the H' parameter reveal a MAPE of 0.04% and a maximum percentage error of 1.2%. The aforementioned values exhibit a correlation with MAE values of 0.02 km and 0.74 km, respectively. Additional analysis was conducted in the time-series based approach, similar to the instance-based approach, to validate the predicted values against the actual values

(Figure 11). The evaluation metrics were validated by comparing the predicted and actual values, with both variables yielding an R^2 score close to 1 (i.e., 100%), which signified a highly accurate regression model. In a similar vein, the linear regression analysis conducted on the data substantiated the alignment with the 45-degree line, thereby providing additional evidence of a highly accurate correspondence between the predicted and observed waveguide parameters.

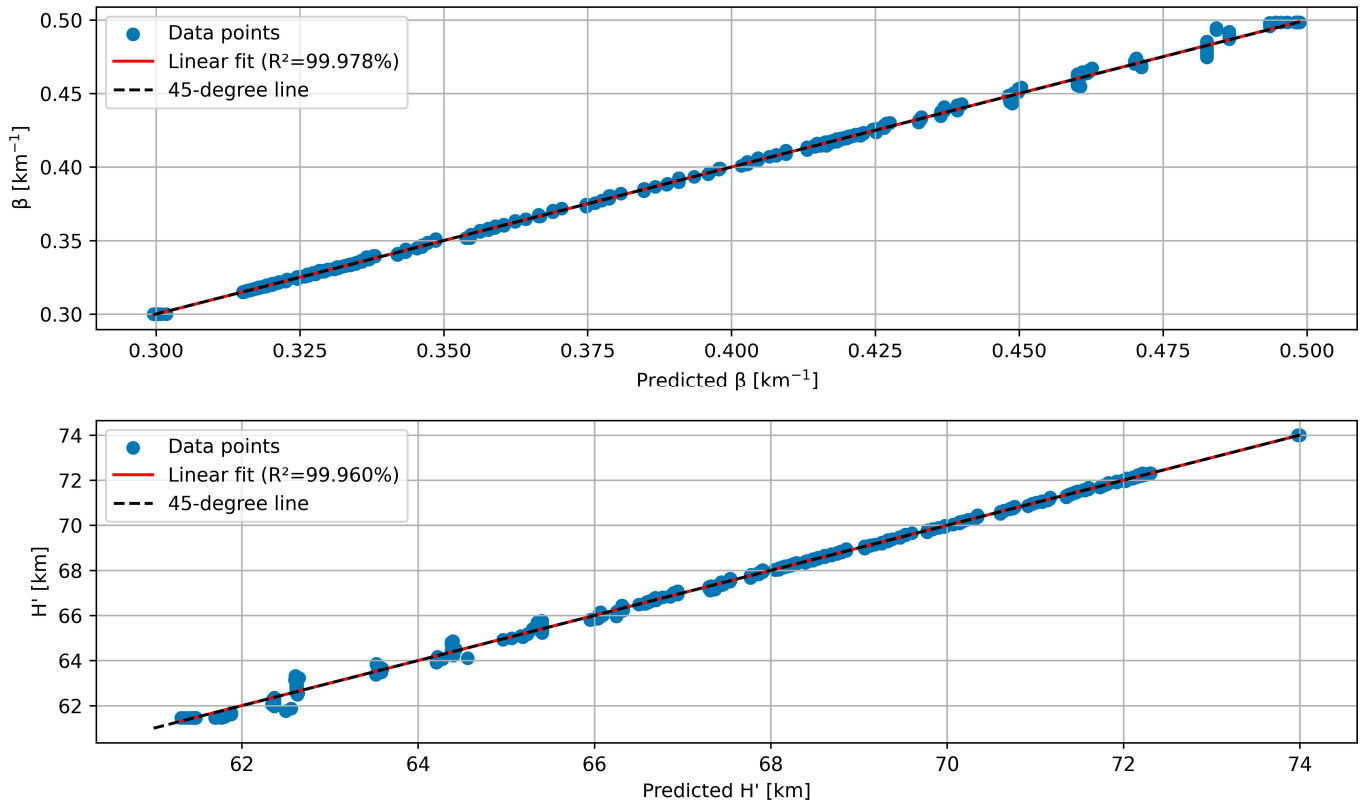


Figure 11. Predicted and observed value for sharpness (**upper** panel) and reflection height (**lower** panel) for the time-series approach.

Akin to the previous example involving the instance-based approach, an analysis of residuals was conducted for the time-series approach, as depicted in Figure 12. In contrast to the previous instance of residual analysis, the residuals observed in the time-series approach exhibited pronounced patterns. Specifically, in the upper panel of Figure 12, it is evident that the residuals increases as the predicted β also increase. Additionally, there is a negative relationship between the predicted H' and the corresponding residuals, indicating that as the predicted H' increases, the residuals decrease. The β predictions at a rate of 0.45 km^{-1} exhibit a notable increase in residual values compared to the previous range of $+0.005$ to -0.005 km^{-1} . Similarly, for H' values exceeding the predicted threshold of 66 km, the residuals fall within the range of $+0.25$ to -0.25 km .

The analysis of feature importance involved the ranking of features based on their informedness. Consistent with expectations, the X-ray irradiance feature exhibited the highest importance, accounting for approximately 89% of the overall feature importance. Following this, the rolling median statistic with a 5 min window demonstrated a significance of 8.38%. The collective contribution of additional features amounts to approximately 2.76% of the informedness for the model, suggesting that these features can be readily modified or eliminated to align with the researcher’s requirements. One potential challenge faced by researchers is the exclusion of the initial 60 data points from the analysis due to the implementation of a rolling window statistic with a duration of 60 min. However, it is worth noting that from the information given by the feature importance analysis, these

features can be disregarded without compromising the model's effectiveness, thus enabling the development of an equally proficient model.

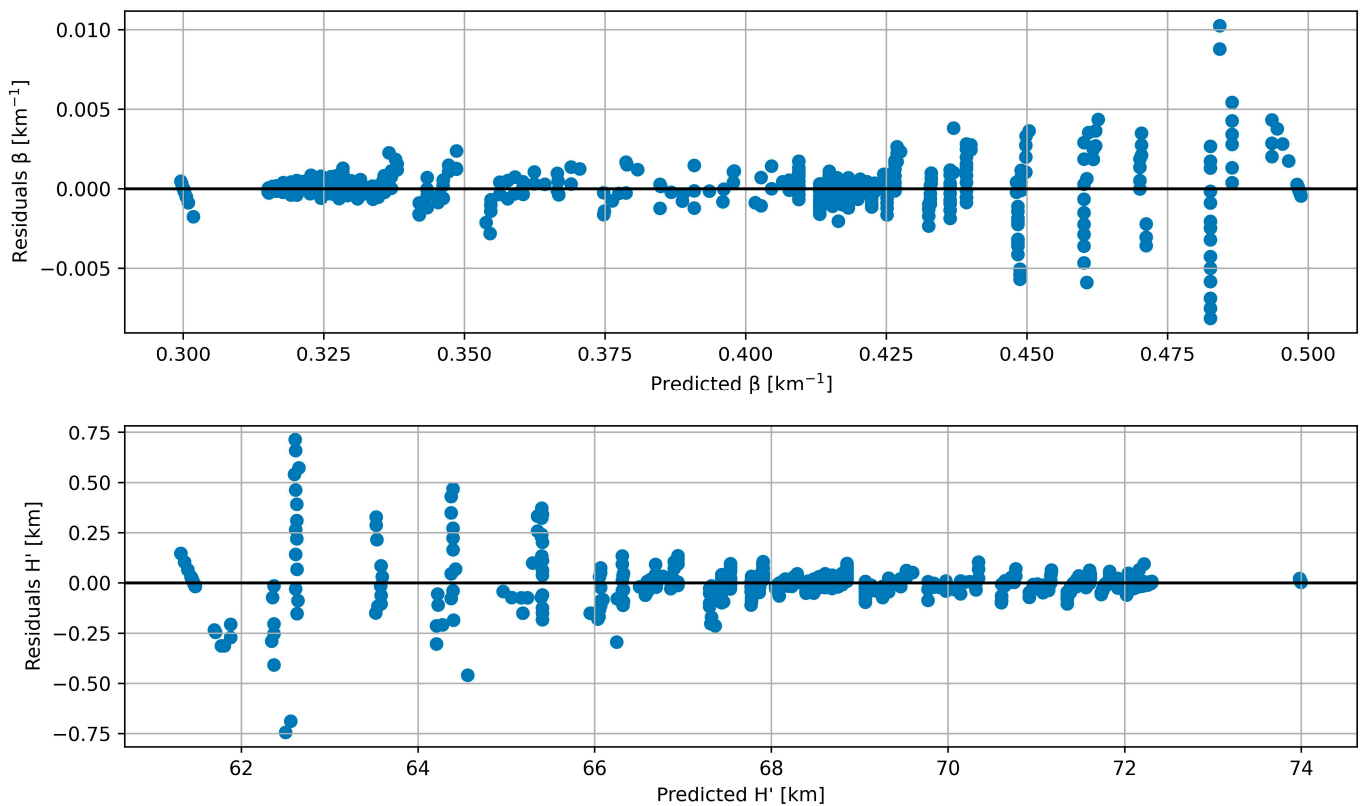


Figure 12. Residual and predicted plot for sharpness (**upper** panel) and reflection height (**lower** panel) for the time-series approach.

4. Conclusions

The employment of ML regression techniques on the provided SF data yielded significant insights into the feasibility of modeling ionospheric parameters without relying on the complicated and difficult to use software, e.g., LWPC software. One possible advantage of this approach, provided that the models are optimized, and the error rates are minimized, is the possibility of automatically determining ionospheric parameters in real-time or, at the very least, near-real time. Also, another potential benefit of such a model is the determination of the ED from the ionospheric parameters. The primary findings of this research can be summarized as:

- The utilization of synthetic data estimated using the KDE technique yielded datasets that were deemed adequate for ML modeling, as they closely adhered to the distribution of the original dataset. Further investigation is required to validate the outcomes of this study. Subsequent research should involve a more extensive dataset and, if feasible, refrain from relying on synthetic data, instead opting for a greater number of original samples. In relation to the present study, the utilization of synthetic data proved to be adequate, as the primary aim of this research was to determine the feasibility of employing ML regression techniques for the estimation of ionospheric parameters.
- The RF and XGB algorithms demonstrated adequate performance; however, the KNN and DT algorithms exhibited greater error rates compared to the aforementioned techniques. Subsequent investigations ought to integrate and prioritize the utilization of ANNs due to their benefits; however, they do necessitate careful hyperparameter tuning in order not overfit the model. Regarding XGB, it is worth noting that it possesses an additional hyperparameter compared to RF. This additional hyperparameter

allows for finer adjustments to the model, perhaps leading to improved predictions. Nevertheless, both RF and XGB are highly recommended as primary methodologies for investigating concepts that have not been completely explored.

- The residual analysis conducted in this study revealed that the final model had a possible minor bias towards predicting H' values greater than 62 km, with a reduced error rate compared to predictions below 62 km.
- The results obtained from the time-series based approach exhibited a higher level of favorability compared to the instance-based approach, as indicated by the lower error rates. The model exhibited a potential bias in both the β and H' parameters. Specifically, the β parameter demonstrated an increasing error rate as the predicted value increased, whereas the H' parameter showed a decreasing error rate as the predicted value increased. Future research should consider placing more emphasis on a time-series based approach. This approach has shown the ability to efficiently present precise values of waveguide parameters over an extended period of time. Additionally, it has been observed that the features of this approach can be customized to meet the specific requirements of the researcher. Notably, it has been found that only two features contribute significantly to the informativeness of the model.
- Standard methods for determining ionospheric parameters are tedious and time-consuming, necessitating the development of other methods for determining such parameters. As to our knowledge, the literature and freely available methods for providing ionospheric parameters utilizing ML are not widely realized. Future comparison of the displayed ML method can be performed with methods such as easyFit and FlareED, where all the techniques can be tested and mutually compared under different SF classes and ionospheric perturbations.

The primary objective of this study is to employ an alternative approach for estimating low ionospheric parameters under the influence of SF events that enable easy modeling of this medium. An advantage of this method is the potential ability to streamline the process and obtain results in real-time or near-real time, as well as the potential to obtain parameters for the calculation of ED. However, further investigation is required to refine the methodology, investigate alternative algorithms, and explore additional pre-processing techniques. The wider statistical determination of the capabilities of the model for all SF classes can be enabled with additional data. As expected, the majority of the data fell within the C or M class, and future research is needed with more B- (barely detectable except in conditions of solar minimum) and X-class SFs. The research demonstrates the promise of the approach; nonetheless, additional comprehensive research is required to ensure its readiness for production.

Supplementary Materials: The training and testing data as well as parts of the workflow used in this study are available online at: https://github.com/arnautF/IR_SF_ML, accessed on 3 November 2023.

Author Contributions: Conceptualization, F.A. and A.K.; writing—original draft preparation, F.A. and A.K.; writing—review and editing F.A., A.K., V.A.S. and Z.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

Data Availability Statement: In this study, publicly accessible datasets were examined. These data are accessible: https://hesperia.gsfc.nasa.gov/goes/goes_event_listings/, accessed on 9 April 2023, and <https://www.ncei.noaa.gov/data/goes-space-environment-monitor/access/avg/>, accessed on 7 May 2023.

Acknowledgments: The article is based upon work from COST Action CA22162—A transdisciplinary network to bridge climate science and impacts on society (FutureMed). Authors thank D. Šulić for fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fedrizzi, M.; de Paula, E.R.; Kantor, I.J.; Langley, R.B.; Santos, M.C. Mapping the low-latitude ionosphere with GPS. *GPS WORLD* **2002**, *13*, 41–47.
2. Ahmedov, B.J.; Mirzaev, B.S.; Mamatov, F.M.; Khodzhaev, D.A.; Julliev, M.K. Integrating of gis and gps for ionospheric perturbations in d-And f-layers using vlf receiver. *InterCarto InterGIS* **2020**, *26*, 547–560. [[CrossRef](#)]
3. Kumar, S.I.; Kumar, A.; Menk, F.W.; Maurya, A.K.; Singh, R.; Veenadhari, B. Response of the low-latitude D region ionosphere to extreme space weather event of 14–16 December 2006. *J. Geophys. Res. Space Phys.* **2015**, *120*, 788–799. [[CrossRef](#)]
4. Mitra, A. The D-region of the ionosphere. *Endeavour* **1978**, *2*, 12–21. [[CrossRef](#)]
5. Ohya, H.; Nishino, M.; Murayama, Y.; Igarashi, K.; Saito, A. Using tweek atmospheric to measure the response of the low-middle latitude D-region ionosphere to a magnetic storm. *J. Atmos. Sol.-Terr. Phys.* **2006**, *68*, 697–709. [[CrossRef](#)]
6. Reddybattula, K.D.; Panda, S.K.; Sharma, S.K.; Singh, A.K.; Kurnala, K.; Haritha, C.S.; Wuyyuru, S. Anomaly effects of 6–10 September 2017 solar flares on ionospheric total electron content over Saudi Arabian low latitudes. *Acta Astronaut.* **2020**, *177*, 332–340. [[CrossRef](#)]
7. Ishisaka, K.; Okada, T.; Hawkins, J.; Murakami, S.; Miyake, T.; Murayama, Y.; Nagano, I.; Matsumoto, H. Investigation of electron density profile in the lower ionosphere by SRP-4 rocket experiment. *Earth Planets Space* **2005**, *57*, 879–884. [[CrossRef](#)]
8. Quan, L.; Cai, B.; Hu, X.; Xu, Q.; Li, L. Study of ionospheric D region changes during solar flares using MF radar measurements. *Adv. Space Res.* **2021**, *67*, 715–721. [[CrossRef](#)]
9. Richardson, D.; Cohen, M. Exploring the Feasibility of a Unified D-region Ionosphere Model. In Proceedings of the AGU Fall Meeting Abstracts, New Orleans, LA, USA, 13–17 December 2021; p. AE35B-1920.
10. Silber, I.; Price, C. On the Use of VLF Narrowband Measurements to Study the Lower Ionosphere and the Mesosphere–Lower Thermosphere. *Surv. Geophys.* **2017**, *38*, 407–441. [[CrossRef](#)]
11. Kolarski, A.; Veselinović, N.; Srećković, V.A.; Mijić, Z.; Savić, M.; Dragić, A. Impacts of Extreme Space Weather Events on September 6th, 2017 on Ionosphere and Primary Cosmic Rays. *Remote Sens.* **2023**, *15*, 1403. [[CrossRef](#)]
12. Grubor, D.; Šulić, D.M.; Žigman, V. Classification of X-ray solar flares regarding their effects on the lower ionosphere electron density profile. *Ann. Geophys.* **2008**, *26*, 1731–1740. [[CrossRef](#)]
13. Kolarski, A.; Grubor, D. Sensing the Earth’s low ionosphere during solar flares using VLF signals and goes solar X-ray data. *Adv. Space Res.* **2014**, *53*, 1595–1602. [[CrossRef](#)]
14. Kolarski, A.; Grubor, D. Comparative Analysis of VLF Signal Variation along Trajectory Induced by X-ray Solar Flares. *J. Astrophys. Astron.* **2015**, *36*, 565–579. [[CrossRef](#)]
15. Thomson, N.R.; Clilverd, M.A.; McRae, W.M. Nighttime ionospheric D region parameters from VLF phase and amplitude. *J. Geophys. Res. Space Phys.* **2007**, *112*, A07304. [[CrossRef](#)]
16. Wait, J.R.; Spies, K.P. *Characteristics of the Earth-Ionosphere Waveguide for VLF Radio Waves*; US Department of Commerce, National Bureau of Standards: Washington, DC, USA, 1964; Volume 300.
17. Ferguson, J. *Computer Programs for Assessment of Long-Wavelength Radio Communications, Version 2.0: User’s Guide and Source Files*; TD-3030, Space and Naval Warfare Systems Center: San Diego, CA, USA, 1998.
18. Bekker, S.Z.; Ryakhovskiy, I.A.; Korsunskaya, J.A. Modeling of the Lower Ionosphere During Solar X-Ray Flares of Different Classes. *J. Geophys. Res. Space Phys.* **2021**, *126*, e2020JA028767. [[CrossRef](#)]
19. Srećković, V.A.; Šulić, D.M.; Vujčić, V.; Mijić, Z.R.; Ignjatović, L.M. Novel Modelling Approach for Obtaining the Parameters of Low Ionosphere under Extreme Radiation in X-Spectral Range. *Appl. Sci.* **2021**, *11*, 11574. [[CrossRef](#)]
20. Srećković, V.A.; Šulić, D.M.; Ignjatović, L.; Vujčić, V. Low Ionosphere under Influence of Strong Solar Radiation: Diagnostics and Modeling. *Appl. Sci.* **2021**, *11*, 7194. [[CrossRef](#)]
21. Gross, N.C.; Cohen, M.B. VLF Remote Sensing of the D Region Ionosphere Using Neural Networks. *J. Geophys. Res. Space Phys.* **2020**, *125*, e2019JA027135. [[CrossRef](#)]
22. Alpatov, V.V.; Bekker, S.Z.; Kozlov, S.I.; Lyakhov, A.N.; Yakim, V.V.; Yakubovskiy, S.V. Analyzing existing applied models of the ionosphere to calculate radio wave propagation and a possibility of their use for radar-tracking systems. II. Domestic models. *Sol.-Terr. Phys.* **2020**, *6*, 60–66.
23. Richardson, D.K.; Cohen, M.B. Seasonal Variation of the D-Region Ionosphere: Very Low Frequency (VLF) and Machine Learning Models. *J. Geophys. Res. (Space Phys.)* **2021**, *126*, e29689. [[CrossRef](#)]
24. Berdermann, J.; Kriegel, M.; Banyś, D.; Heymann, F.; Hoque, M.M.; Wilken, V.; Borries, C.; Heßelbarth, A.; Jakowski, N. Ionospheric Response to the X9.3 Flare on 6 September 2017 and Its Implication for Navigation Services over Europe. *Space Weather* **2018**, *16*, 1604–1615. [[CrossRef](#)]
25. de Paula, V.; Segarra, A.; Altadill, D.; Curto, J.J.; Blanch, E. Detection of Solar Flares from the Analysis of Signal-to-Noise Ratio Recorded by Digisonde at Mid-Latitudes. *Remote Sens.* **2022**, *14*, 1898. [[CrossRef](#)]
26. Reddybattula, K.D.; Nelapudi, L.S.; Moses, M.; Devanaboyina, V.R.; Ali, M.A.; Jamjareegulgarn, P.; Panda, S.K. Ionospheric TEC Forecasting over an Indian Low Latitude Location Using Long Short-Term Memory (LSTM) Deep Learning Network. *Universe* **2022**, *8*, 562. [[CrossRef](#)]
27. Yasyukevich, Y.; Astafyeva, E.; Padokhin, A.; Ivanova, V.; Syrovatskii, S.; Podlesnyi, A. The 6 September 2017 X-Class Solar Flares and Their Impacts on the Ionosphere, GNSS, and HF Radio Wave Propagation. *Space Weather* **2018**, *16*, 1013–1027. [[CrossRef](#)] [[PubMed](#)]

28. Berger, V.W.; Zhou, Y. Kolmogorov–smirnov test: Overview. In *Wiley StatsRef: Statistics Reference Online*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014. [[CrossRef](#)]
29. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)] [[PubMed](#)]
30. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
31. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees. *Biometrics* **1984**, *40*, 874.
32. Fix, E.; Hodges, J.L. Discriminatory analysis. Nonparametric discrimination: Consistency properties. *Int. Stat. Rev./Rev. Int. Stat.* **1989**, *57*, 238–247. [[CrossRef](#)]
33. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
34. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
35. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor. Newsl.* **2004**, *6*, 20–29. [[CrossRef](#)]
36. Arnaut, F.; Kolarski, A.; Srečković, V.A. Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. *Universe* **2023**, *9*, 436. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Improving Air Quality Data Reliability through Bi-Directional Univariate Imputation with the Random Forest Algorithm

Filip Arnaut^{1,*}, Vladimir Đurđević², Aleksandra Kolarski¹, Vladimir A. Srećković¹ and Sreten Jevremović³

¹ Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11000 Belgrade, Serbia; kolarski@ipb.ac.rs (A.K.); vlada@ipb.ac.rs (V.A.S.)

² Faculty of Physics, University of Belgrade, Cara Dušana 13, 11000 Belgrade, Serbia; vdj@ff.bg.ac.rs

³ Scientific Society “Isaac Newton”, Volgina 7, 11160 Belgrade, Serbia; jevremovic749@gmail.com

* Correspondence: filip.arnaut@ipb.ac.rs

Abstract: Forecasting the future levels of air pollution provides valuable information that holds importance for the general public, vulnerable populations, and policymakers. High-quality data are essential for precise and reliable forecasts and investigations of air pollution. Missing observations arise when the sensors utilized for assessing air quality parameters experience malfunctions, which result in erroneous measurements or gaps in the dataset and hinder the data quality. This research paper presents a novel approach for imputing missing values in air quality data in a univariate approach. The algorithm employs the random forest (RF) algorithm to impute missing observations in a bi-directional (forward and reverse in time) manner for air quality (particulate matter less than 2.5 μm ($\text{PM}_{2.5}$)) data from the Republic of Serbia. The algorithm was evaluated against simple methods, such as the mean and median imputation methods, for missing observations over durations of 24, 48, and 72 h. The results indicate that our algorithm yielded comparable error rates to the median imputation method for all periods when imputing the $\text{PM}_{2.5}$ data. Ultimately, the algorithm's higher computational complexity proved itself as not justified considering the minimal error decrease it achieved compared with the simpler methods. However, for future improvement, additional research is needed, such as utilizing low-code machine learning libraries and time-series forecasting techniques.

Keywords: data imputation; air quality; $\text{PM}_{2.5}$; air pollution; missing observations; machine learning



Citation: Arnaut, F.; Đurđević, V.; Kolarski, A.; Srećković, V.A.; Jevremović, S. Improving Air Quality Data Reliability through Bi-Directional Univariate Imputation with the Random Forest Algorithm. *Sustainability* **2024**, *16*, 7629. <https://doi.org/10.3390/su16177629>

Academic Editors: José Carlos Magalhães Pires and Álvaro Gómez-Losada

Received: 9 August 2024

Revised: 26 August 2024

Accepted: 1 September 2024

Published: 3 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Air quality (AQ) is a continuous and significant factor in public health since it is associated with increased mortality [1,2], cardiovascular diseases [2], increased susceptibility to allergens [3], and respiratory illnesses [4–6]. Industrialization, urbanization, and population growth can all contribute to an increase in the concentration of harmful particulate matter in the air [7–9]. Among all the measured AQ parameters, the presence of particulate matter with a diameter less than 2.5 μm ($\text{PM}_{2.5}$) is especially problematic because it can be inhaled deeply into the body through the respiratory system [10,11]. The origins of these fine particles can be attributed to a byproduct of fossil fuels burning due to the use of vehicles and the utilization of materials for home heating [12]. In highly populated regions that experience significant traffic density, having accurate information about the AQ is crucial for the proper functioning and well-being of the community.

In 2022, diseases of the circulatory system were the primary cause of death in the Republic of Serbia (RS), where they constituted 47.3% of all recorded deaths. An additional 6% of deaths were ascribed to respiratory system diseases [13]. Both of these affected groups are susceptible to heightened air pollution. Furthermore, the annual death toll due to the consequences of air pollution amounts to approximately 6 to 7 million individuals worldwide [14,15].

The adverse effects of heightened air pollution on human health are extensively documented and widely recognized. There are numerous solutions available to reduce air pollution, but the responsibility for implementing them lies with decision makers. One effective approach to address air pollution is to predict the upcoming levels of air pollution. This offers helpful insights into future air pollution levels, which can be of great significance to the general public, vulnerable populations, and policymakers. There is an abundance of literature available on the topic of predicting air quality that ranges from traditional time-series (TS) techniques to more advanced machine learning (ML) and deep learning methods.

Classical TS forecasting methods primarily utilize the auto-regressive integrated moving average (ARIMA) model [16,17] or its variant that includes the seasonal component, which is known as SARIMA [18]. Another relatively novel classical TS forecasting method is Facebook's Prophet [18–20]. Both methods are traditional TS forecasting techniques that can be effectively used to predict AQ values. Additionally, ML-based methods are also extensively employed in AQ forecasting. Various ML models can be employed for this purpose, including extreme gradient boosting (XGBoost) [21], support vector regression (SVR) [21–23], random forest [22], and recurrent neural networks and long short-term memories [22,24]. The primary differentiation between TS forecasting methods and ML methods lies in the data requirements of each approach. Specifically, TS forecasting methods can be utilized in a univariate manner, where future values are predicted based on the past values of the modelled variable. However, ML-based methods typically require the use of additional parameters, such as meteorological, traffic, and other data, to model the target variable. In certain instances, when there is a lack of available data, this can pose a significant problem.

However, in order to achieve accurate and reliable predictions of air pollution, it is crucial to prioritize the quality of data. This applies to both traditional TS and ML methods, as data quality is the key determinant of the accuracy of forecasts.

Missing observations (MOs) significantly contribute to the poor quality of AQ data. MOs occur when the sensors used to measure AQ parameters malfunction, which leads to either erroneous measurements that need to be removed by the researcher or gaps in the dataset. Figure 1 showcases 10 chosen AQ measurement stations in the RS from the Serbian Environmental Protection Agency (SEPA) network, which provides a concise representation of the seriousness of MOs over a 4-year span from 2018 to 2021. Figure 1 clearly illustrates the presence of MOs that extend over a significant duration, which presents a challenge for researchers in terms of ensuring data quality. In Figure 1, the average percentage of MOs for the AQ stations was 15.3%, with the lowest value being 3.9% (station b, Figure 1) and the highest value being 30.9% (station j, Figure 1). The possible explanation of the MOs in the SEPA network can be attributed to technical errors at the measurement sites, which occur randomly and are unrelated to the actual measured values, i.e., the data itself.

Figure 1 depicts the extent of MOs in AQ data that researchers must contend with. As a result, over the years, numerous AQ imputation techniques have been created, suggested, and assessed [15,25–27]. The mean-before-after method, which is a variation of a simple imputation method (mean imputation method), is displayed in [28]. This method takes into account two data points before and after the gap. Similar to this, bi-directional long short-term memory, which is a neural network architecture, evaluates data both before and after the MO location [29]. A comparison of other advanced methods, as well as deep learning methods, was done to analyze PM_{2.5} data from Peru [30]. In [31], it is displayed that Kalman smoothing on structural time series was the most effective method for data imputation using air quality data from Sydney. Furthermore, the MICE method demonstrated its advantages in the imputation of missing AQ data [32]. A method known as missforest, which is associated with the random forest algorithm, was employed to successfully impute missing AQ values for data from Kuwait [33].

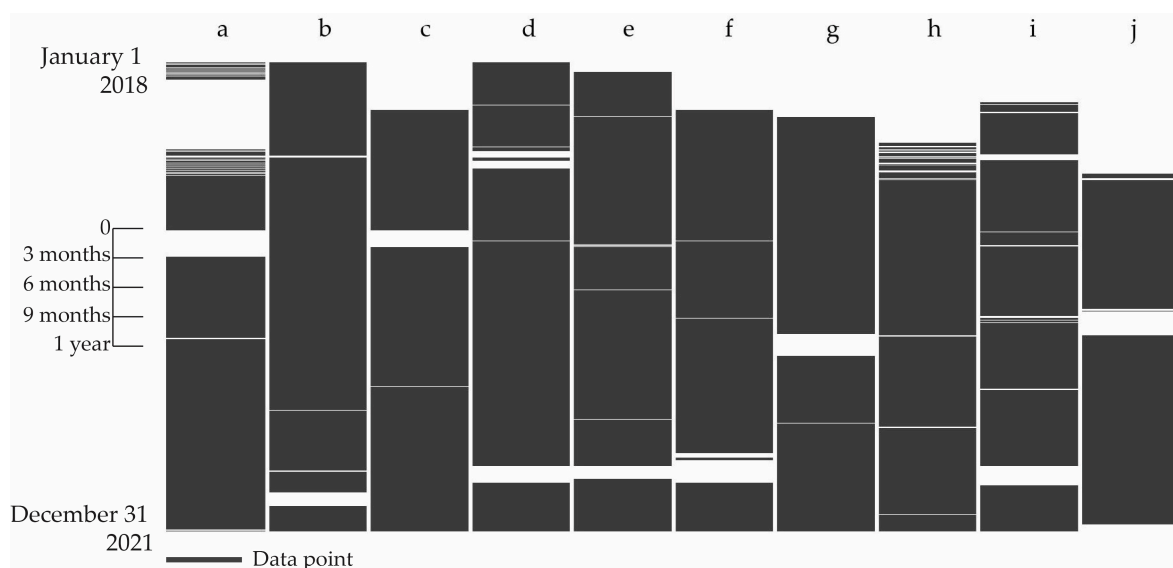


Figure 1. Missing $PM_{2.5}$ observations at selected air quality measuring stations from 1 January 2018 until 31 December 2021; a—Novi Sad Rumenačka St.; b—Belgrade Old Town; c—Belgrade New Belgrade; d—Belgrade Mostar; e—Smederevo Center; f—Obrenovac Center; g—Valjevo; h—Bor City Park; i—Kosjerić; j—Niš “Sveti Sava” Elementary School.

The goal of this paper was to present a newly developed algorithm for imputing missing values in univariate AQ data. The decision to employ a strictly univariate approach was made due to the recognition that additional data for ML algorithms may not always be easily accessible. Therefore, having a robust, accurate, and reliable univariate method is advantageous. The algorithm is designed to determine the location of MOs, divide the data into “before” and “after” segments, and utilize the ML model to predict and fill in MOs in both forward and reverse directions in time.

2. Methods and Data

2.1. Algorithm Workflow

The $PM_{2.5}$ data, which includes missing observations, were used as the input time series for the algorithm (Figure 2). The algorithm first identifies the gaps in the data, i.e., the MOs’ locations and sizes. Then, the input TS data are divided into two separate TSs: TS1 and TS2. TS1 is used as the TS from which the algorithm will make a forecast in the future, as is typically done in forecasting problems. On the other hand, TS2 is employed to make a forecast backward in time. To enable backward forecasting, TS2 is inverted.

The algorithm proceeds by applying a data transformation for both TS1 and TS2 (Figure 2). The data transformation aims to reduce the range of data used for forecasting in order to simplify the data-modeling process (usually to make the data more stationary). The data transformation utilized in this research was the log transformation.

Feature discovery or feature generation is a crucial initial step in ML modeling. As previously stated, the imposed constraints of this research are that only $PM_{2.5}$ data were used, without incorporating any additional data, such as meteorological conditions, traffic patterns, or other air quality parameters. Consequently, the available features that could be employed were solely the $PM_{2.5}$ data. The features used in this research consisted of lagged features, i.e., look-back data, which referred to the number of data points analyzed prior to a given data point, as well as rolling window statistics, such as the mean, median, and standard deviation. The new parameters introduced into the algorithm in this step were the number of lagged features and the window size for the rolling window statistics.

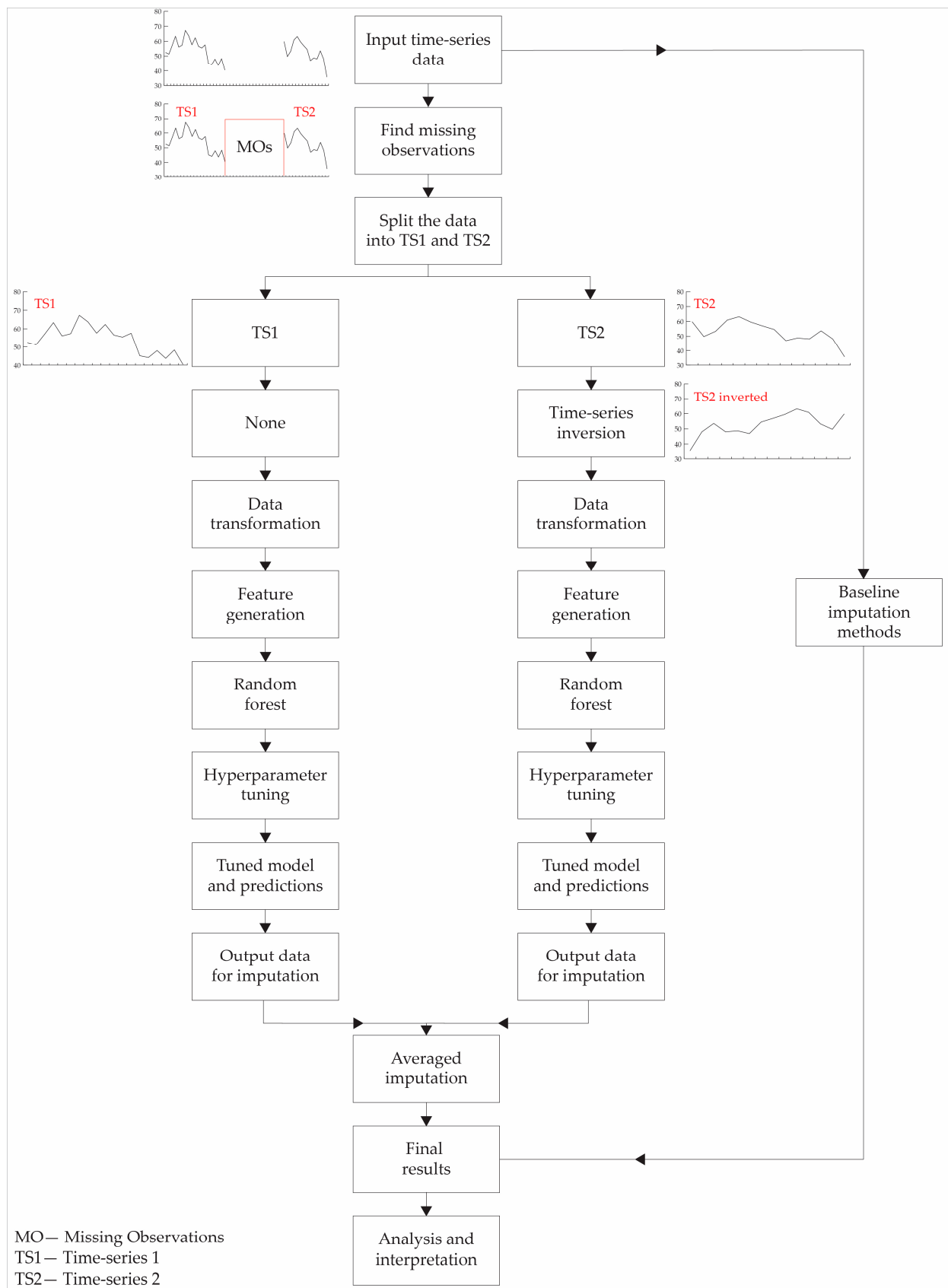


Figure 2. Simplified air quality data imputation workflow employed in this study.

Once the features are generated, the data are passed to the random forest (RF) regression model. The RF model was introduced by Breiman in 2001 [34] and has since become one of the most widely used ML algorithms. The versatility of the RF algorithm in the Earth sciences can be seen in its wide utilization across many fields, such as near-Earth physics [35–38], lithological prediction [39,40], mineral prospectivity [41–43], and land classification [44]. The RF model is a tree-based model that can be seen as a progression and enhancement of decision trees (DTs). DTs use a measure of homogeneity, such as the Gini index or entropy for classification tasks or the mean square error for regression tasks, to split a group of instances, where the resulting group has pure (more homogenous) target variables (the majority belonging to one class or similar range of values). This allows them to make predictions, either for classification problems (class prediction) or regression problems (number prediction). Upon receiving a new instance, the DT model assigns it to a suitable group based on the features of the given instance, and subsequently assigns the target class of the group to the new instance. DTs are simple models that were subsequently improved and surpassed by the RF model. DTs have several drawbacks, including a tendency to overfit and limited predictive power when dealing with larger and more complex datasets. However, one of their benefits is their interpretability. The RF model is a collection of DTs (ensemble model) that employs the bootstrap method with replacement to generate multiple bootstrap samples from the original dataset, which enables training more DTs. The ultimate prediction is determined through voting in classification problems or averaging in regression problems. The RF model has the number of trees as a hyperparameter, which determines the number of DTs used to train the model. The RF model offers advantages in terms of its lower susceptibility to overfitting by averaging the predictions of multiple DTs. Due to the aforementioned reasons and the ease of implementation in modern software packages and libraries, the RF model is considered an excellent initial choice when developing novel research areas.

The RF model was trained on 80% of the data, while the remaining 20% was kept aside for testing and generating values for imputation. This research employed the random search hyperparameter optimization technique, which involved testing 50 models within the range of 10 to 1000 trees. The optimized model was subsequently employed to produce values for imputation. As the model generated a new value, it was added to the feature list as an additional entry in the test set. The model then proceeded to generate another value. This process was repeated iteratively until all the data were inputted.

The identical procedure of model creation, training, tuning, and so forth that was carried out for TS1 was replicated for TS2. The TS1 and TS2 imputed values were used to calculate an average imputed value from the model, which resulted in a total of three values: forward, backward, and averaged imputation.

Alongside this process, baseline imputation methods were applied, including simple methods like mean and median imputations, where the MOs were imputed with the mean or median values of the dataset.

The error quantification for this research was done in a straightforward manner by using the absolute error (AE) and its aggregate: the mean absolute error (MAE). This research utilized multiple AQ monitoring stations, and because of that, the mean, median, minimum, and maximum MAE for each imputation method are displayed throughout this research paper.

2.2. Dataset Preparation

This research utilized AQ data obtained from the Serbian Environmental Protection Agency (SEPA) network. The SEPA network employs Grimm EDM 180 PM monitors [45,46] throughout the RS, which are fully automated PM monitors that operate on the principle of light scattering, i.e., an optical particle counter, for PM detection. As the AQ data collected by the SEPA network are sent to the European Environmental Agency (EEA) database, all PM measurements adhered to the standardization requirements established by the European Union and the EEA. The initial dataset comprised AQ measurements from 2018

to 2021 for different parameters, including $PM_{2.5}$, PM_{10} , SO_2 , CO , and NO , across different locations in the RS. The initial datasets were employed to identify the longest uninterrupted sequences of recorded $PM_{2.5}$ data. Figure 3 displays a total of nine stations that were considered suitable for the current study. The durations of the consecutive data ranged from 4715 h, which was equivalent to approximately 196 consecutive measured days in Belgrade New Belgrade, to 8735 h, which was equivalent to 364 consecutive measured days in Belgrade Mostar and Old Town.

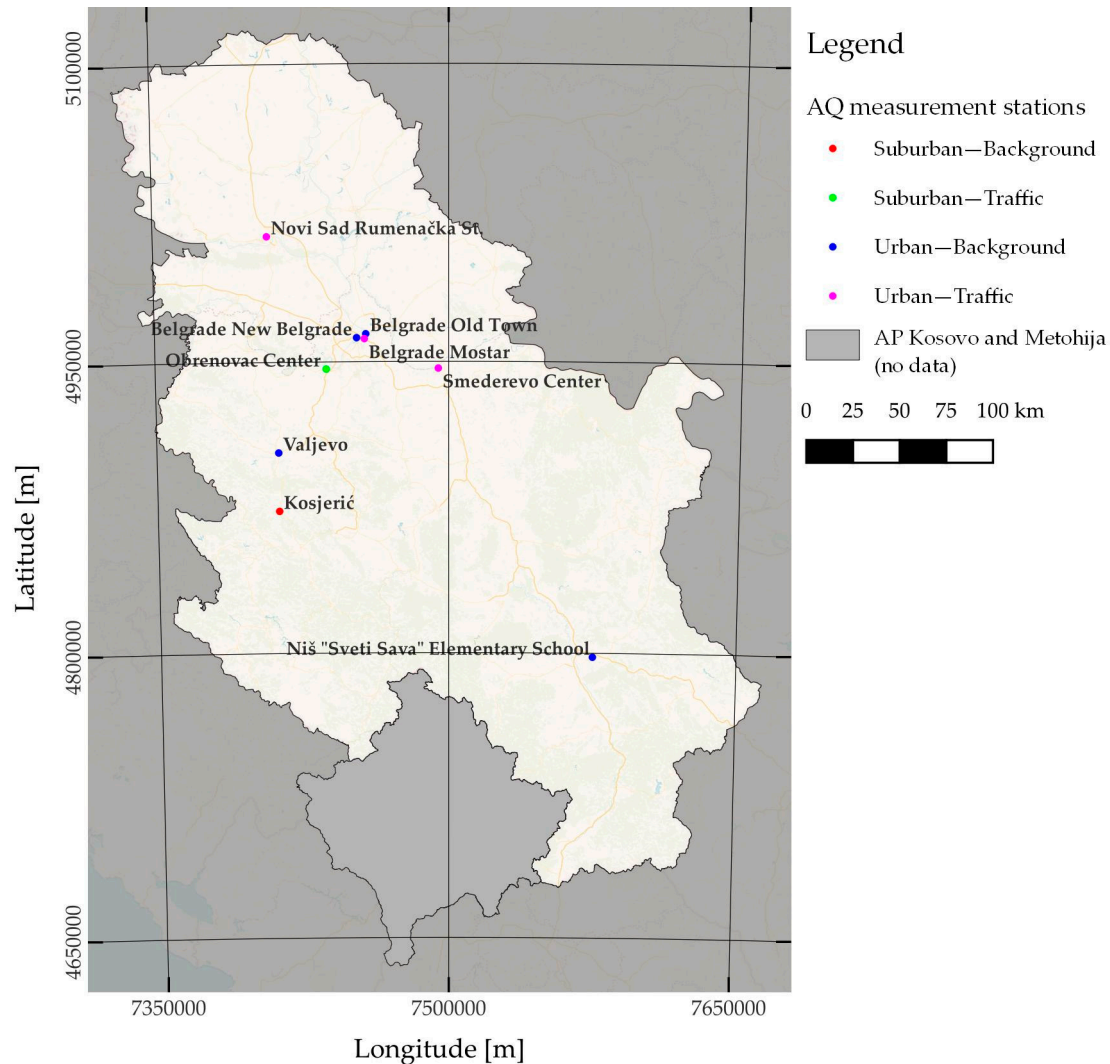


Figure 3. Distribution of air quality monitoring stations whose data were utilized in this research across the Republic of Serbia.

The chosen $PM_{2.5}$ measurement stations were distributed throughout the RS, with three located in the capital and largest city of Belgrade (Mostar, New Belgrade, and Old Town). The other two largest cities in the RS, Novi Sad and Niš, each had one selected $PM_{2.5}$ measurement station. The remaining stations were spread across various localities in the RS, including Obrenovac, Valjevo, and Smederevo. In terms of the area and station classifications, the three largest cities (Belgrade, Novi Sad, and Niš) had stations situated in urban areas of the city. These stations assessed either traffic-related or background air pollution levels. The Mostar AQ station in Belgrade was situated near a prominent traffic junction that experienced heavy traffic congestion during peak hours throughout the day. Additional AQ stations were situated in suburban or urban environments and assessed either background or traffic pollutants.

Once the suitable AQ stations were determined, MOs of 24, 48, and 72 h were introduced into the TS. Random locations for the missing observations were chosen for each AQ station, which resulted in varying lengths of TS1 and TS2 in order to test the algorithm and precisely gauge its effectiveness. To quantify the error rates of the aforementioned imputation methods, the original, i.e., measured, data, which were located where the MOs were introduced, were separated and preserved for subsequent use when the models imputed the data.

A decision was made to select relatively small values of MOs (24, 48, and 72 h) and compare our algorithm with simple imputation methods, such as mean and median imputations. Additionally, the algorithm was provided with the maximum amount of available data in the given circumstances from the SEPA network. In order to provide the algorithm with optimal conditions for success, we specifically generated small MO values that yielded more data for model training and a narrower range of data that needed to be imputed. However, the decision to use mean and median imputations was made to provide the proposed algorithm with relatively simple yet widely utilized methods for comparison. Mean and median imputations, while straightforward, are suitable for data imputation in certain cases. However, they are not recommended for long sequences of data, particularly those with expected periodic patterns or variability [47,48], such as AQ data.

Furthermore, selecting the longest measured sequences of data from the SEPA network enabled the model to be in optimal conditions for demonstrating an enhancement over simple imputation methods. In other words, our algorithm was provided with the most favorable conditions in terms of the quantity of data, as well as the methods to which it was compared.

It is important to mention that certain AQ monitoring stations that were considered to be of interest (e.g., displaying high error rates and having a high variability in the data or other) were subjected to testing with alternative imputation techniques. These techniques included classic RF, which incorporates additional features, such as meteorological data and time parameters (time of day, day of week, etc.), and incorporates a classical instance-based approach and modern imputation techniques, like the iterative imputer [49] with the dataset consisting of data up to the MO location. Due to the greater number of stations and variations in the conditions and lengths of the MOs, not all stations underwent testing with alternative techniques and different features. Only a limited number of stations were selected for this purpose and are displayed in the results section.

3. Results

3.1. One Day (24 h) of Consecutive Missing Observations

Figure 4 displays the MAE distribution for all five imputation methods used on a dataset with nine AQ measurement stations with 24 consecutive MOs. The distribution revealed the presence of a significant outlier in all the imputation methods. Specifically, for the forward imputation method (Figure 4a), the MAE for the given AQ station was approximately $147 \mu\text{g}/\text{m}^3$, while for the other methods, it ranged from $80 \mu\text{g}/\text{m}^3$ to around 100 to $120 \mu\text{g}/\text{m}^3$. This outlier skewed the average MAE calculated for all AQ stations using a given imputation method, as was evident when comparing the mean and median MAE values. When comparing the imputation methods based on the median MAE value, two methods stood out with similar minimal values: forward imputation with $7.55 \mu\text{g}/\text{m}^3$ (Figure 4a) and median imputation with $7.87 \mu\text{g}/\text{m}^3$ (Figure 4e).

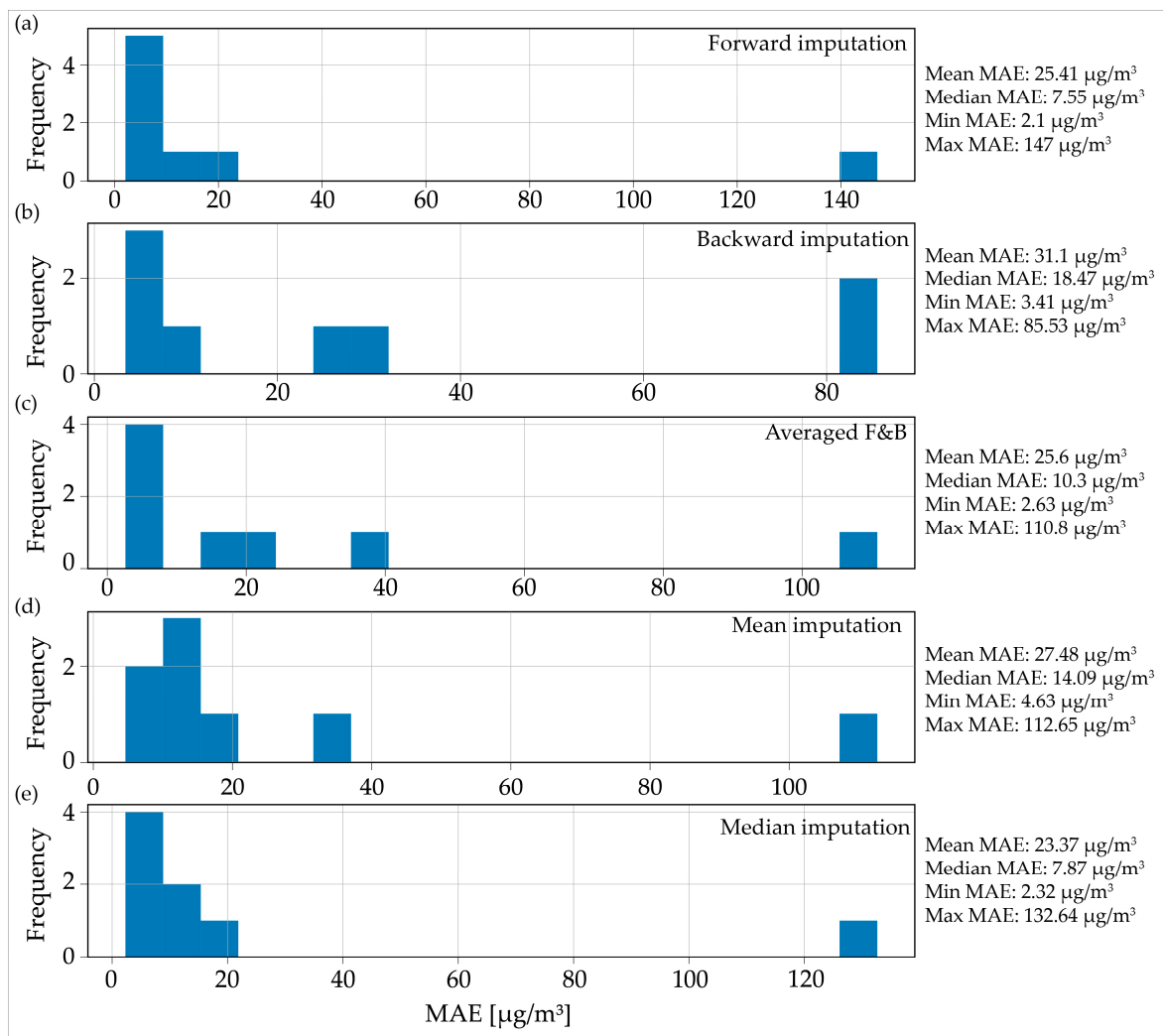


Figure 4. Distribution of the mean absolute error for the imputed data for one day (24 h) consecutive missing observations: (a) forward imputation; (b) backward imputation; (c) averaged imputation; (d) mean imputation; (e) median imputation.

In the distributions presented in Figure 4, only eight AQ stations were included out of the total of nine that were analyzed. Upon investigating the reason for the omission of one station from the distribution, it was found that the algorithm failed to converge on one specific example, which resulted in the absence of any results for that AQ station.

Regarding the severe outlier, all imputation methods in Figure 4 indicated that the dataset originated from the city of Valjevo. This dataset covered the period from 20 June 2018 to 31 March 2019, thus totaling 6803 h or 283 consecutive measuring days. After analyzing the Valjevo dataset, it was evident that there was a significant increase in the $\text{PM}_{2.5}$ parameter values that started from the month of October (Figure 5). The locations of MOs are depicted in Figure 5 by the green dashed line, which were situated in the highly fluctuating section of the increased $\text{PM}_{2.5}$ values. In the locations of the 24 h MOs, the $\text{PM}_{2.5}$ values varied between approximately $80 \mu\text{g}/\text{m}^3$ and $250 \mu\text{g}/\text{m}^3$, thus exhibiting distinct peaks and lows. Due to the wide range of $\text{PM}_{2.5}$ concentrations and the significant variability in these concentrations, none of the tested models were successful in accurately imputing the data with a relatively low margin of error. In addition to the higher error rate that resulted from the fluctuating levels of $\text{PM}_{2.5}$, one possible explanation is that the decrease in temperature, which started in October 2018, led the residents of Valjevo to rely on wood and other combustible materials for home heating. This, in turn, contributed to the rise in ambient pollution levels. The Pearson's and Spearman's correlation coefficients were

computed to determine the relationship between temperature and $PM_{2.5}$ concentrations for the given period. The results indicate a moderate, negative correlation, meaning that as the ambient temperature decreased, the $PM_{2.5}$ concentration increased and vice versa. These findings align with similar results obtained from [6].

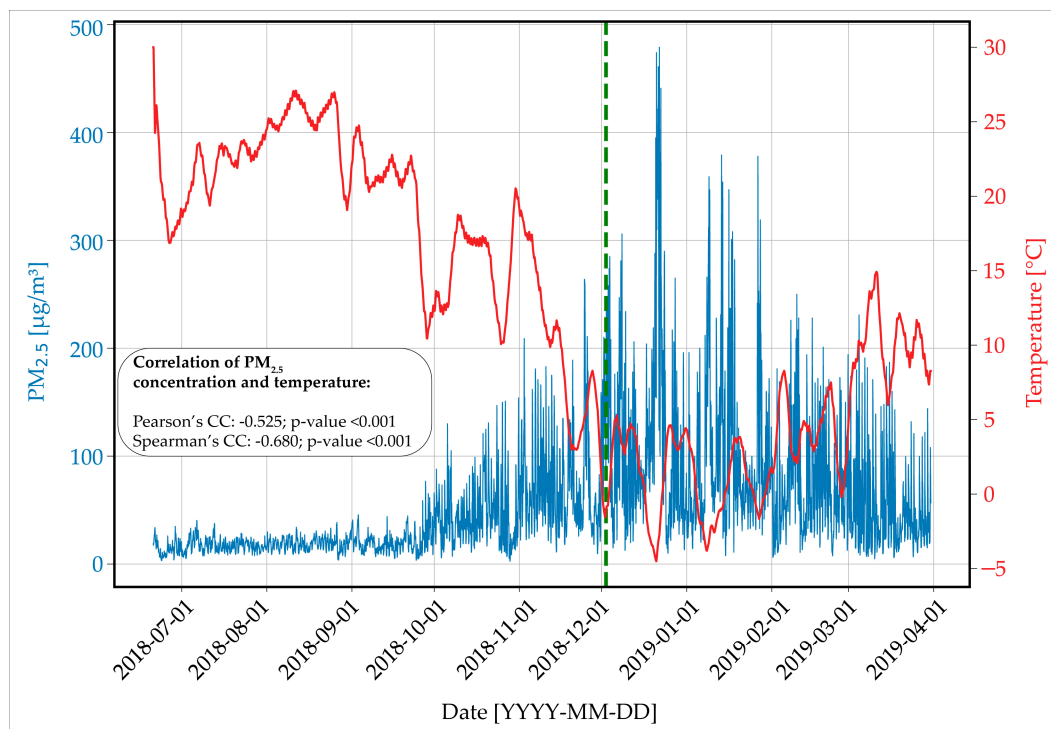


Figure 5. $PM_{2.5}$ concentrations from 20 June 2018 to 31 March 2019 in the city of Valjevo, Republic of Serbia; blue line— $PM_{2.5}$ concentrations; red line—temperature (measurement station in Belgrade, Republic of Serbia); green line—locations of missing observations.

Furthermore, two additional models, namely, the classic RF and iterative imputer, were tested on the data to evaluate their suitability for this extreme example. The two models used incorporated meteorological variables, such as the temperature, dew point, relative humidity, wind direction, wind speed, and visibility, as well as time-related variables, including time of day, day of week, and day or night classification. Both models surpassed the performance of the five previously used models; however, they still exhibited significant MAE values of $79.7 \mu\text{g}/\text{m}^3$ and $90 \mu\text{g}/\text{m}^3$ for the 24 h MO period.

3.2. Two Days (48 h) of Consecutive Missing Observations

The median imputation method yielded the lowest mean MAE value across all AQ measurement stations in the case of 48 h of MOs (Figure 6e). This was followed by the mean imputation (Figure 6d) method and the forward imputation method (Figure 6a). The application of backward imputation in this particular scenario resulted in the poorest overall performance, with an average value of $32.19 \mu\text{g}/\text{m}^3$. When examining the median MAE values, the situation varied, and the most effective method overall was forward imputation, followed by median imputation.

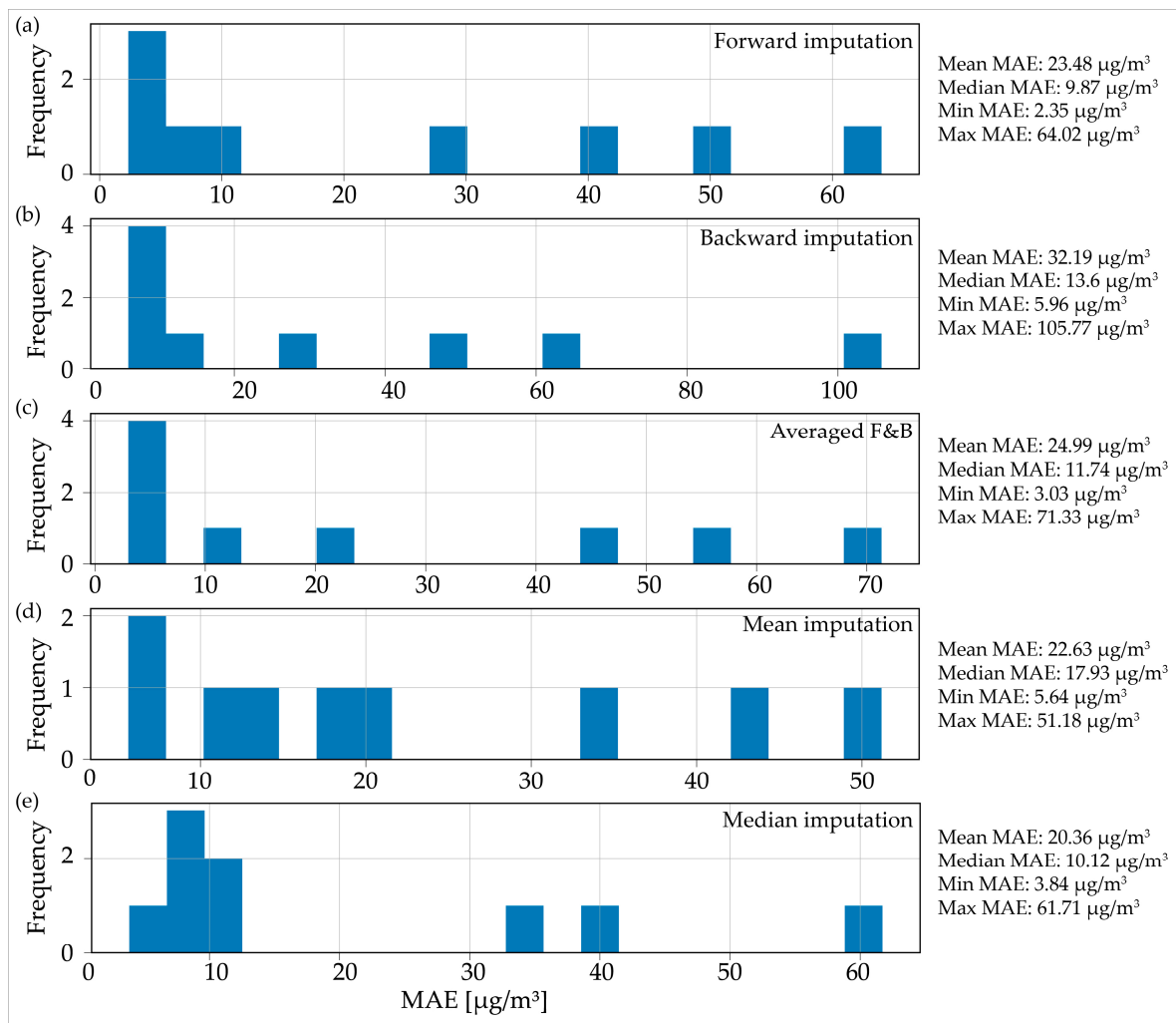


Figure 6. Distributions of the mean absolute error for the imputed data for two days (48 h) of consecutive missing observations: (a) forward imputation; (b) backward imputation; (c) averaged imputation; (d) mean imputation; (e) median imputation.

Choosing the most efficient method for the 2-day MO case came down to deciding between forward and median imputation, as the mean and median MAE values were quite similar for both methods. However, it was crucial to consider both the computational complexity and time of the proposed method. The average duration for the proposed algorithm to generate the forward and backward data imputations was 13.21 min. However, the time split between the forward and backward imputations was not always evenly divided at 50–50%. In contrast, the median and mean imputation methods were instantaneous on modern computers. Consequently, the proposed method in this research was less effective when compared with straightforward methods for the 48 h MOs scenario.

Figure 7 provides an illustration of the imputations made by different models for the Niš “Sveti Sava” elementary school example. Figure 7a displays the 48 h MO period, along with the actual $\text{PM}_{2.5}$ values and imputations from different models. The proposed algorithm’s forward forecast yielded the most favorable results in terms of minimal MAE values, with a value of $2.34 \mu\text{g}/\text{m}^3$. This was closely followed by the average forecast from our algorithm ($3.3 \mu\text{g}/\text{m}^3$), and then by the RF classic ($4.16 \mu\text{g}/\text{m}^3$) and iterative imputer ($4.17 \mu\text{g}/\text{m}^3$). When examining Figure 7b, it shows the complete $\text{PM}_{2.5}$ data for the specified location, along with the temperature for the given time period and the locations of the MOs. This situation is comparable to the previous example. During the winter, the decrease in temperature led to an increase in air pollution due to the heightened usage of wood

and other materials for home heating. Due to the positioning of the MOs, the backward forecast was trained using values that were higher than those used in the forward forecast. As a result, the backward forecast predicted larger values than the forward forecast and performed as the third worst model. The two models with the lowest performance were the mean imputation method, with a value of $21 \mu\text{g}/\text{m}^3$, and the median imputation method, with a value of $8.2 \mu\text{g}/\text{m}^3$. This information is noteworthy because both methods made use of the entire dataset rather than just the portion of the dataset before the locations of the MOs. Mean values are highly susceptible to the influence of outlier values, which results in the mean imputation method producing more inaccurate estimations. In contrast, median values are not affected by outliers, which leads to more accurate estimations. When a dataset has significant variations in its data values, it is more efficient to use the median instead of the mean for imputation [50].

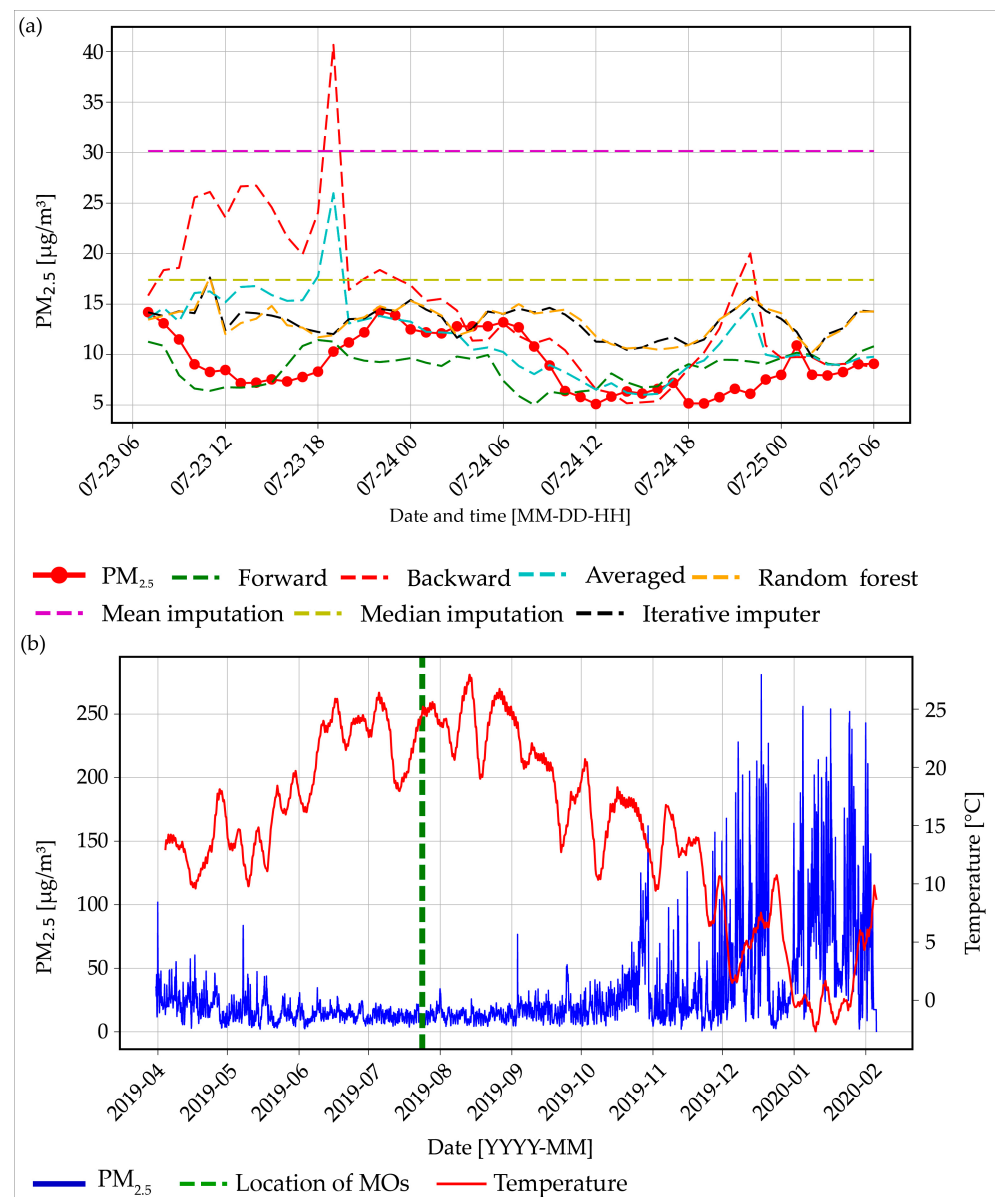


Figure 7. (a) The 48 h missing observations period and imputed values from different models for the Niš “Sveti Sava” Elementary School example and (b) PM_{2.5} values for the Niš “Sveti Sava” Elementary School example with temperature values for the period.

3.3. Three Days (72 h) of Consecutive Missing Observations

During the three days or 72 consecutive hours of MOs, the situation was analogous to the previous two examples. Figure 8 compares the forward and median imputation methods (Figure 8a,e), which were the two best overall methods. As previously mentioned, while there was a comparison between the two methods in terms of MAE values, the computational complexity and time for our method was not justified by the reduction in error estimates. In fact, the error values between a simple and instantaneous method and our method were quite similar. Furthermore, it is worth mentioning that similar to the previous example, the median imputation method yielded more accurate estimates of PM_{2.5} MOs compared with the mean imputation method, which showed that when having highly fluctuating data, a better choice is the median imputation method.

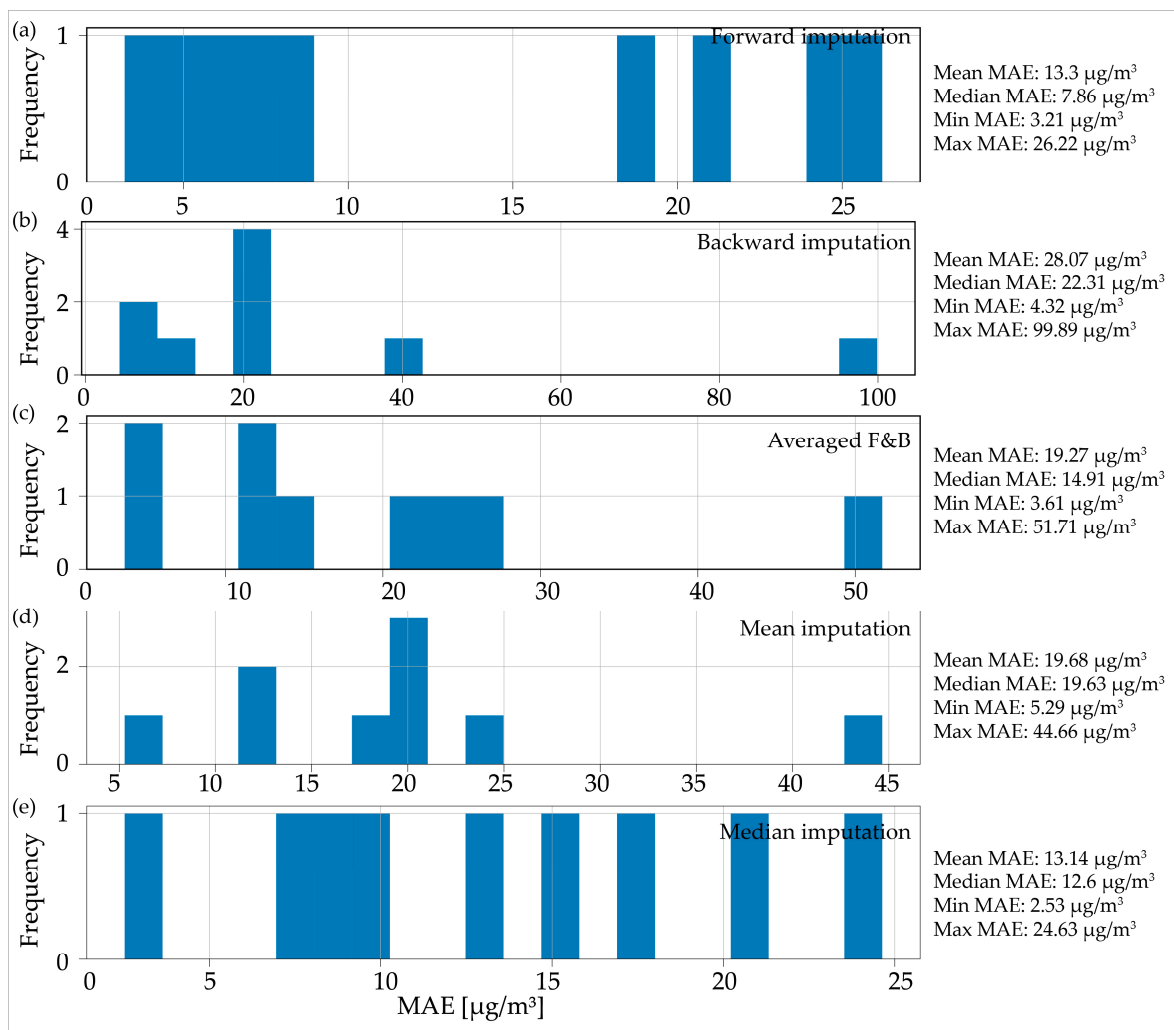


Figure 8. Distribution of the mean absolute error for the imputed data for three days (72 h) of consecutive missing observations: (a) forward imputation; (b) backward imputation; (c) averaged imputation; (d) mean imputation; (e) median imputation.

Given the previous results and the use of the iterative imputer and classic RF up to the locations of the MOs, a different approach was taken for the example that involved three days of consecutive MOs. From the previous two examples, it was observed that there was a correlation between the temperature and the rise in PM_{2.5} concentrations. The correlation between the temperature and PM_{2.5} may be indirect due to the heating season in the RS, which begins on October 15 and ends on April 15 nationwide. This seasonal pattern could be an important feature for the model. In this example, in addition to the

usual meteorological features (temperature, relative humidity, wind speed, visibility) and time features (time of day, day of week), a binary feature was included to indicate the beginning and end of the heating season in the RS. In addition, the RF model was provided with the complete dataset both before and after the locations of the MOs. The example used was Belgrade, Old Town, which included data from 31 March 2019 to 29 March 2020, thus totaling approximately 8702 consecutive measured hours (equivalent to 362 consecutive measured days). The MAE for the model that did not include the heating season feature was $11.9 \mu\text{g}/\text{m}^3$. However, when the heating season feature was included in the model, the MAE decreased to $7.8 \mu\text{g}/\text{m}^3$. In addition, the feature importance analysis determined that the heating season feature was the second most informative, after visibility and before wind speed, temperature, and relative humidity. Similarly, in the same example, the forward imputation yielded an MAE value of $7.8 \mu\text{g}/\text{m}^3$, while the backward imputation resulted in a significantly higher value of $99 \mu\text{g}/\text{m}^3$. The MAE values for the mean and median imputation models were $19.6 \mu\text{g}/\text{m}^3$ and $9.4 \mu\text{g}/\text{m}^3$, respectively.

4. Discussion

4.1. Benefits and Limitations

Considering the results presented earlier, it is necessary to address the limitations of the suggested method. While our proposed method showed favorable results in certain cases, it did not outperform simpler methods, like the mean and median imputation, overall.

Our method incurred significantly greater computational complexity and time compared with the other methods employed in this research study. The computational time for the classical, instance-based RF and iterative imputer methods, which are effective and modern techniques for data imputation, was only a few seconds. Regarding the mean and median imputation methods, as previously mentioned, the computational time was very fast, i.e., instantaneous on modern computers. Regarding our proposed method, the average computational time (on a standard home computer) for the entire algorithm was approximately 13 min. However, it is important to note that there was an unequal time distribution between the forward and backward imputations, with a 50–50% split not always being observed. Also, the computational time mentioned was achieved by allowing the algorithm to fully leverage the maximum processing capabilities. Therefore, for configurations with lower capacities, the computational time could be greater.

It is worth noting that the task of locating extended, uninterrupted sequences of measured data can be challenging. In this research paper, we selected nine SEPA AQ measurement stations that had a minimum of 4715 continuous measured data points and a maximum of 8735 continuous measured hours. This research paper presented the ideal situation for our proposed method; however, in reality, the data were characterized by a high frequency of low-duration MOs (low MO duration but high occurrence frequency) and a low frequency of high-duration MOs (high MO duration but relatively low occurrence frequency).

One more constraint of this research is that we employed the RF model to produce imputed values. The RF model possesses hyperparameters, such as the number of trees. However, our approach introduced new parameters that should be taken into account, thereby increasing the complexity of our method. The number of features in this research can be regarded as a parameter, as it is connected to lagged values and rolling window statistics. The researcher needs to determine the optimal number of lagged values to utilize, as well as the size of the rolling window. In theory, our method can be adjusted for each specific case, but this inevitably leads to a higher level of computational complexity.

Our method has a drawback in which the model generates values to sequentially fill in missing data. These generated values are then transferred to the feature set one-by-one, where they are used as data points to generate new values. This approach has a weakness in which it is trained using measured data. Introducing a synthetic data point in the feature list poses a challenge, unless that data point has a correct, i.e., expected value. In other words, the errors in our model accumulate at a faster rate compared with standard forecasting

problems because each new generated data point is contingent upon the “correctness” of the previous generated data point.

Finally, it is important to acknowledge that creating a one-size-fits-all method for data imputation on AQ data is highly challenging. Another factor to consider is the increased difficulty in making the method exclusively univariate due to the presence of numerous location-specific factors (as observed with the heating period feature in this research) and the influence of multiple different factors on AQ values (traffic patterns for a given location, weather conditions, holiday season, etc.). An option is to individually address each dataset that requires data imputation by employing a variety of methods specifically tailored for the specific data at hand.

4.2. Future Perspectives

In addition to the limitations displayed earlier, there are a few noteworthy points that need to be highlighted, and further research can be expanded upon. Based on our research, we consider the proposed method of providing imputations or forecasts of MOs using two directions to be an interesting approach that deserves further investigation. In order to conduct future research, it is possible to explore two alternative methods. One possible approach is to use a low-code ML library like PyCaret [51] instead of the currently used RF method. PyCaret not only offers the ability to train, test, validate, and tune the model but also the capability to do so quickly, thereby reducing the computational time and complexity of our approach. In addition, a comparison can be made between the forward and backward imputations to determine whether divergence occurs. If it does, only one of the imputations should be used. In addition to this, another comparable approach that can be employed is the utilization of standard TS forecasting techniques, such as ARIMA or Facebook’s Prophet [52]. These methods are robust TS forecasting techniques that incorporate automated workflows for tuning parameters, specifically in the case of the ARIMA technique. Furthermore, it is worth considering a hybrid approach that combines PyCaret, which is a standard ML framework, with TS forecasting techniques. This combination has the possible potential to further improve the proposed algorithm, which we plan to conduct in the future.

Furthermore, an interesting parameter for future research in data imputation methods is the planetary boundary layer (PBL), which is a parameter that is not measured but modeled. The PBL exhibits a negative correlation with $PM_{2.5}$ values [53], which makes this valuable information to include in future models. Additionally, the PBL is thinner during nighttime and in colder months. This is consistent with our observation that as the ambient temperature decreased, there was an increase in ambient AQ pollution. Future research in the field is expected to explore various perspectives in utilizing non-inherently measured data. This was because sensors often require maintenance periods, experience downtime, and may encounter malfunctions.

Finally, it is important to acknowledge for future research that AQ should be considered not only as a variable that changes over time, but also as a variable that changes both in space and time, as air pollution moves across different locations and periods, as displayed in [54,55]. Regarding data imputation methods, if the stations are in close proximity to each other, the MOs of one station can be assumed to be the same as the values of the other station in the event of a malfunction. Furthermore, studying the spatial and temporal fluctuations of air pollution is crucial, as it can provide insights into the movement of pollutants and the factors that impact their dispersion.

5. Conclusions

The matter of AQ is a crucial concern for public health that necessitates attention in vulnerable regions. Data quality is a critical aspect that includes the quantity of missing data. One possible approach to enhance data quality in AQ measurements is through the use of imputation methods.

This research paper proposes a novel method for imputing missing values in AQ data that was tested on PM_{2.5} data from the Republic of Serbia. The method relies solely on univariate data, which means that only the data of the parameter itself was used for imputation. This decision was made due to the limited availability or unreliability of additional data in certain situations. Hence, it is beneficial to devise a technique for estimating or imputing data for a particular AQ parameter solely using its own data.

The algorithm we proposed yielded comparable results to the median imputation methods during the specified period of MOs (24, 48, and 72 consecutive hours). The computational complexity of our algorithm was not justified by the marginal error reduction it achieved, as the reduction was neither substantial nor consistent across all AQ stations.

Our research analysis revealed a strong correlation between the decreased ambient temperature throughout the year and a significant increase in PM_{2.5} concentrations. This was interpreted as individuals utilizing wood and other flammable substances for residential heating, thereby elevating the surrounding levels of PM_{2.5} and other particles.

When the data gathered from this research was incorporated into an ML model as a binary feature called heating season, the model exhibited improved MAE values compared with when the feature was not included. The feature importance ranking placed the heating season feature as the second most informative, just below visibility and above wind speed.

The limitations of this research are thoroughly discussed and can be summarized as follows: computational complexity/time, the lack of long uninterrupted streams of data, data quality issues, and the challenge of developing a solely univariate method.

However, there are future research opportunities that will be further developed. Further research is required to explore the application of time-series forecasting techniques, potentially in conjunction with low-code ML libraries, to reduce the computational complexity and time.

Author Contributions: Conceptualization, F.A.; methodology, F.A., A.K., V.A.S. and S.J.; software, F.A.; formal analysis, F.A., A.K., V.A.S. and S.J.; writing—F.A., A.K., V.A.S. and S.J.; writing—review and editing; F.A., V.D., A.K., V.A.S. and S.J.; supervision, V.D., A.K. and V.A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the project “UniBelgrade: Climate attribution SRB 23/24 (Mentoring programme for young researchers to adopt advanced knowledge in climate research and to effectively communicate their results)”, grant number: G-2311-67522, funded by the European Climate Foundation, as well as by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to express gratitude to the Serbian Environmental Protection Agency (SEPA), OpenStreetMaps, and Iowa State University—Iowa Environmental Mesonet for providing open data that was utilized in this research paper. And we would also like to express our gratitude to the anonymous reviewers for their thoughtful and meaningful comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dockery, D.W.; Schwartz, J.; Spengler, J.D. Air pollution and daily mortality: Associations with particulates and acid aerosols. *Environ. Res.* **1992**, *59*, 362–373. [[CrossRef](#)]
2. Araujo, J.A. Particulate air pollution, systemic oxidative stress, inflammation, and atherosclerosis. *Air Qual. Atmos. Health* **2011**, *4*, 79–93. [[CrossRef](#)]
3. Bernstein, J.A.; Alexis, N.; Barnes, C.; Bernstein, I.L.; Nel, A.; Peden, D.; Diaz-Sanchez, D.; Tarlo, S.M.; Williams, P.B. Health effects of air pollution. *J. Allergy Clin. Immunol.* **2004**, *114*, 1116–1123. [[CrossRef](#)] [[PubMed](#)]

4. Libasin, Z.; Ul-Saufie, A.Z.; Ahmat, H.; Shaziayani, W.N. Single and multiple imputation method to replace missing values in air pollution datasets: A review. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Seoul, Republic of Korea, 23–24 July 2020; p. 012002.
5. Rakholia, R.; Le, Q.; Vu, K.; Ho, B.Q.; Carbajo, R.S. AI-based air quality PM_{2.5} forecasting models for developing countries: A case study of Ho Chi Minh City, Vietnam. *Urban Clim.* **2022**, *46*, 101315. [[CrossRef](#)]
6. Arnaut, F.; Cvetkov, V.; Đurić, D.; Samardžić-Petrović, M. Short-term forecasting of PM₁₀ and PM_{2.5} concentrations with Facebook’s Prophet Model at the Belgrade-Zelena brdo. *Geofizika* **2023**, *40*, 162–177. [[CrossRef](#)]
7. Harishkumar, K.; Yogesh, K.; Gad, I. Forecasting air pollution particulate matter (PM_{2.5}) using machine learning regression models. *Procedia Comput. Sci.* **2020**, *171*, 2057–2066.
8. Wardana, I.N.K.; Gardner, J.W.; Fahmy, S.A. Estimation of missing air pollutant data using a spatiotemporal convolutional autoencoder. *Neural Comput. Appl.* **2022**, *34*, 16129–16154. [[CrossRef](#)]
9. Zhang, Z.; Zhang, S. Modeling air quality PM_{2.5} forecasting using deep sparse attention-based transformer networks. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 13535–13550. [[CrossRef](#)]
10. Rahman, E.A.; Hamzah, F.M.; Latif, M.T.; Azid, A. Forecasting PM_{2.5} in Malaysia using a hybrid model. *Aerosol Air Qual. Res.* **2023**, *23*, 230006. [[CrossRef](#)]
11. Zhang, Y.; Sun, Q.; Liu, J.; Petrosian, O. Long-Term Forecasting of Air Pollution Particulate Matter (PM_{2.5}) and Analysis of Influencing Factors. *Sustainability* **2023**, *16*, 19. [[CrossRef](#)]
12. Zaini, N.a.; Ean, L.W.; Ahmed, A.N.; Abdul Malek, M.; Chow, M.F. PM_{2.5} forecasting for an urban area based on deep learning and decomposition method. *Sci. Rep.* **2022**, *12*, 17565. [[CrossRef](#)] [[PubMed](#)]
13. Institute of Public Health “Dr. Milan Jovanovic Batut”. *Health Statistical Yearbook of the Republic of Serbia 2022*; Institute of Public Health of Serbia “Dr Milan Jovanovic Batut”: Belgrade, Serbia, 2023.
14. Hadeed, S.J.; O’rourke, M.K.; Burgess, J.L.; Harris, R.B.; Canales, R.A. Imputation methods for addressing missing data in short-term monitoring of air pollutants. *Sci. Total Environ.* **2020**, *730*, 139140. [[CrossRef](#)] [[PubMed](#)]
15. Kim, T.; Kim, J.; Yang, W.; Lee, H.; Choo, J. Missing value imputation of time-series air-quality data via deep neural networks. *Int. J. Environ. Res. Public Health* **2021**, *18*, 12213. [[CrossRef](#)]
16. Marinov, E.; Petrova-Antonova, D.; Malinov, S. Time series forecasting of air quality: A case study of Sofia City. *Atmosphere* **2022**, *13*, 788. [[CrossRef](#)]
17. Ramadan, M.S.; Abuelgasim, A.; Al Hosani, N. Advancing air quality forecasting in Abu Dhabi, UAE using time series models. *Front. Environ. Sci.* **2024**, *12*, 1393878. [[CrossRef](#)]
18. Samal, K.K.R.; Babu, K.S.; Das, S.K.; Acharaya, A. Time series based air pollution forecasting using SARIMA and prophet model. In Proceedings of the 2019 International Conference on Information Technology and Computer Communications, Singapore, 4–6 April 2005; pp. 80–85.
19. Shen, J.; Valagolam, D.; McCalla, S. Prophet forecasting model: A machine learning approach to predict the concentration of air pollutants (PM_{2.5}, PM₁₀, O₃, NO₂, SO₂, CO) in Seoul, South Korea. *PeerJ* **2020**, *8*, e9961. [[CrossRef](#)] [[PubMed](#)]
20. Ye, Z. Air pollutants prediction in shenzhen based on arima and prophet method. *E3S Web Conf.* **2019**, *36*, 05001. [[CrossRef](#)]
21. Carlés, F.; Recalde, C.; Sauer, C.; Bernal, L.; Stalder, D. Air Quality Time Series Forecasting Using Machine Learning Algorithms. In Proceedings of the 2023 XLIX Latin American Computer Conference (CLEI), La Paz, Bolivia, 16–20 October 2023; pp. 1–9.
22. Espinosa, R.; Palma, J.; Jiménez, F.; Kamińska, J.; Sciavicco, G.; Lucena-Sánchez, E. A time series forecasting based multi-criteria methodology for air quality prediction. *Appl. Soft Comput.* **2021**, *113*, 107850. [[CrossRef](#)]
23. Samad, A.; Garuda, S.; Vogt, U.; Yang, B. Air pollution prediction using machine learning techniques—an approach to replace existing monitoring stations with virtual monitoring stations. *Atmos. Environ.* **2023**, *310*, 119987. [[CrossRef](#)]
24. Freeman, B.S.; Taylor, G.; Gharabaghi, B.; Thé, J. Forecasting air quality time series using deep learning. *J. Air Waste Manag. Assoc.* **2018**, *68*, 866–886. [[CrossRef](#)]
25. Belachsen, I.; Broday, D.M. Imputation of Missing PM_{2.5} Observations in a Network of Air Quality Monitoring Stations by a New k NN Method. *Atmosphere* **2022**, *13*, 1934. [[CrossRef](#)]
26. Chen, M.; Zhu, H.; Chen, Y.; Wang, Y. A novel missing data imputation approach for time series air quality data based on logistic regression. *Atmosphere* **2022**, *13*, 1044. [[CrossRef](#)]
27. Junninen, H.; Niska, H.; Tuppurainen, K.; Ruuskanen, J.; Kolehmainen, M. Methods for imputation of missing values in air quality data sets. *Atmos. Environ.* **2004**, *38*, 2895–2907. [[CrossRef](#)]
28. Norazian, M.N.; Shukri, Y.A.; Azam, R.N.; Al Bakri, A.M.M. Estimation of missing values in air pollution data using single imputation techniques. *ScienceAsia* **2008**, *34*, 341–345. [[CrossRef](#)]
29. Jiang, N.; Li, Y.; Zuo, H.; Zheng, H.; Zheng, Q. BiLSTM-A: A missing value imputation method for PM_{2.5} prediction. In Proceedings of the 2020 2nd International Conference on Applied Machine Learning (ICAML), Tianjin, China, 24–26 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 23–28.
30. Flores, A.; Tito-Chura, H.; Centty-Villafuerte, D.; Ecos-Espino, A. PM_{2.5} time series imputation with deep learning and interpolation. *Computers* **2023**, *12*, 165. [[CrossRef](#)]

31. Wijesekara, W.M.L.K.N.; Liyanage, L. Comparison of imputation methods for missing values in air pollution data: Case study on Sydney air quality index. In *Advances in Information and Communication, Proceedings of the 2020 Future of Information and Communication Conference (FICC), San Francisco, CA, USA, 5–6 March 2020*; Arai, K., Bhatia, R., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 2, pp. 257–269.
32. Kebalepile, M.M.; Dzikiti, L.N.; Voyi, K. Using Diverse Data Sources to Impute Missing Air Quality Data Collected in a Resource-Limited Setting. *Atmosphere* **2024**, *15*, 303. [\[CrossRef\]](#)
33. Alsaber, A.R.; Pan, J.; Al-Hurban, A. Handling complex missing data using random forest approach for an air quality monitoring dataset: A case study of Kuwait environmental data (2012 to 2018). *Int. J. Environ. Res. Public Health* **2021**, *18*, 1333. [\[CrossRef\]](#)
34. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
35. Shang, Z.; Yao, Z.; Liu, J.; Xu, L.; Xu, Y.; Zhang, B.; Guo, R.; Wei, Y. Automated Classification of Auroral Images with Deep Neural Networks. *Universe* **2023**, *9*, 96. [\[CrossRef\]](#)
36. Lian, J.; Liu, T.; Zhou, Y. Aurora Classification in All-Sky Images via CNN–Transformer. *Universe* **2023**, *9*, 230. [\[CrossRef\]](#)
37. Arnaut, F.; Kolarski, A.; Srećković, V.A.; Mijić, Z. Ionospheric Response on Solar Flares through Machine Learning Modeling. *Universe* **2023**, *9*, 474. [\[CrossRef\]](#)
38. Arnaut, F.; Kolarski, A.; Srećković, V.A. Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. *Universe* **2023**, *9*, 436. [\[CrossRef\]](#)
39. Cracknell, M.J.; Reading, A.M. Geological Mapping Using Remote Sensing Data: A Comparison of Five Machine Learning Algorithms, Their Response to Variations in the Spatial Distribution of Training Data and the Use of Explicit Spatial Information. *Comput. Geosci.* **2014**, *63*, 22–33. [\[CrossRef\]](#)
40. Arnaut, F.; Đurić, D.; Đurić, U.; Samardžić-Petrović, M.; Peshevski, I. Application of Geophysical and Multispectral Imagery Data for Predictive Mapping of a Complex Geo-Tectonic Unit: A Case Study of the East Vardar Ophiolite Zone, North-Macedonia. *Earth Sci. Inform.* **2024**, *17*, 1625–1644. [\[CrossRef\]](#)
41. Carranza, E.J.M.; Laborte, A.G. Random Forest Predictive Modeling of Mineral Prospectivity with Small Number of Prospects and Data with Missing Values in Abra (Philippines). *Comput. Geosci.* **2015**, *74*, 60–70. [\[CrossRef\]](#)
42. Carranza, E.J.M.; Laborte, A.G. Data-Driven Predictive Mapping of Gold Prospectivity, Baguio District, Philippines: Application of Random Forests Algorithm. *Ore Geol. Rev.* **2015**, *71*, 777–787. [\[CrossRef\]](#)
43. Zuo, R.; Carranza, E.J.M. Machine Learning-Based Mapping for Mineral Exploration. *Math. Geosci.* **2023**, *55*, 891–895. [\[CrossRef\]](#)
44. Waske, B.; Braun, M. Classifier Ensembles for Land Cover Mapping Using Multitemporal SAR Imagery. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 450–457. [\[CrossRef\]](#)
45. Stojanović, D.B.; Kleut, D.; Davidović, M.; Živković, M.; Ramadani, U.; Jovanović, M.; Lazović, I.; Jovašević-Stojanović, M. Data Evaluation of a Low-Cost Sensor Network for Atmospheric Particulate Matter Monitoring in 15 Municipalities in Serbia. *Sensors* **2024**, *24*, 4052. [\[CrossRef\]](#)
46. Đurić, M.; Vujović, D. Short-term Forecasting of Air Pollution Index in Belgrade, Serbia. *Meteorol. Appl.* **2020**, *27*, e1946. [\[CrossRef\]](#)
47. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing Data*; Wiley: Hoboken, NJ, USA, 2002.
48. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, Australia, 2018. Available online: <https://otexts.org/fpp2/> (accessed on 26 August 2024).
49. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
50. Junger, W.L.; De Leon, A.P. Imputation of Missing Data in Time Series for Air Pollutants. *Atmos. Environ.* **2015**, *102*, 96–104. [\[CrossRef\]](#)
51. Ali, M. PyCaret: An Open Source, Low-Code Machine Learning Library in Python. PyCaret Version 1.0.0. 2020. Available online: <https://www.pycaret.org> (accessed on 26 August 2024).
52. Taylor, S.J.; Letham, B. Forecasting at scale. *Am. Stat.* **2018**, *72*, 37–45. [\[CrossRef\]](#)
53. Wang, Y.; Xu, T.; Shi, G.; Yang, F.; Tang, X.; Zhao, X.; Wan, C.; Liu, S. Climatology of the Planetary Boundary Layer Height over China and Its Characteristics during Periods of Extremely Temperature. *Atmos. Res.* **2023**, *294*, 106960. [\[CrossRef\]](#)
54. Zareba, M.; Cogiel, S.; Danek, T.; Weglinska, E. Machine Learning Techniques for Spatio-Temporal Air Pollution Prediction to Drive Sustainable Urban Development in the Era of Energy and Data Transformation. *Energies* **2024**, *17*, 2738. [\[CrossRef\]](#)
55. Gokul, P.R.; Mathew, A.; Bhosale, A.; Nair, A.T. Spatio-Temporal Air Quality Analysis and PM2.5 Prediction over Hyderabad City, India Using Artificial Intelligence Techniques. *Ecol. Inform.* **2023**, *76*, 102067. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

DOI: <https://doi.org/10.15233/gfz.2022.39.15>
Original scientific paper



Improvement of 1D geoelectric sounding by narrowing the equivalence range and identification, quantification and reduction of lateral effects using the tri-potential technique

Filip Arnaut^{ORCID}, Branislav Sretenović and Vesna Cvetkov^{ORCID}

Department of Geophysics, Faculty of Mining and Geology, University of Belgrade, Belgrade, Serbia

Received 31 January 2022, in final form 26 September 2022

Apparent resistivity data from the Offset-Wenner array (Zemun, Serbia), the square array (Bogatić and Golubac, Serbia) and the Wenner tri-potential technique (Vrdnik, Serbia), were used to detect, measure, and reduce lateral effects in 1D inversion. Forward and inverse modelling with the Wenner α , β and γ arrays determined that the Wenner β array provided the most accurate estimate of the first- and second-layer resistivity, while the Wenner γ array provided the most accurate estimate of the high resistivity substratum. The survey on the Zemun loess plateau revealed that if the lateral index of inhomogeneity (LII) is low, the 1D interpretation of both Wenner arrays is justifiable. In addition, the averaging of resistances will result in an apparent resistivity curve that is devoid of lateral effects resulting from near-surface inhomogeneities. As demonstrated by the Vrdnik example, 1D inversion is inadequate when the values of LII and processing covariance (PC) are high. The survey in Golubac was conducted using the square array, which produced lower PC values than collinear arrays. Therefore, the quality of the averaged sounding curve was higher. Also, the interpolated values of the Offset Wenner array displayed reasonable accuracy, while the extrapolated values were inadequate when a low resistivity substratum was present.

Keywords: near-surface, 2D electrical resistivity tomography, inversion, Offset Wenner array, square array, effect of near-surface inhomogeneities on 1D measurements.

Introduction

The tri-potential technique was developed in the mid-twentieth century (Carpenter, 1955; Carpenter and Habberjam, 1956; Habberjam, 1969). However, it was never widely used because it required additional measurements that made this technique less effective

in comparison with traditional resistivity survey techniques. Improvements in data acquisition, which began in the late 1980s, made this technique more effective, and the additional data that this technique provides made it useful for 1D, 2D, and 3D interpretations in terms of lateral effects detection, quantification, and reduction, as well as reducing the number of subsurface equivalent models.

Simultaneous measurements of lateral and vertical resistivity changes using the tri-potential technique are now a part of modern 2D survey methods (electrical resistivity tomography – ERT). However, this type of measurement has not found widespread use in 1D vertical electrical sounding (VES). The additional data obtained by the tri-potential technique enable the detection, quantification, and reduction of lateral effects (Habberjam and Watkins, 1967a), as well as the improvement of the overall measurement accuracy using the relationship between the three measured resistivities ($\rho_a^\alpha, \rho_a^\beta, \rho_a^\gamma$).

In the 1D model case, where resistivity changes with depth only, the sounding curves ρ_a^α , ρ_a^β , and ρ_a^γ will exhibit the same general pattern. However, at different depths of investigation, changes in resistivity will be observed first on the ρ_a^γ sounding curve and lastly on the ρ_a^β sounding curve *i.e.*, the (Wenner) gamma array displays the greatest depth of investigation values (Szalai et al., 2014; Loke, 2022). In the case of resistivity increase with depth, the ρ_a^γ sounding curve will display the highest resistivity values, while the ρ_a^β sounding curve will show the lowest resistivity values. If the model is inverted (resistivity decrease with depth) the order will be reversed ($\rho_a^\gamma < \rho_a^\alpha < \rho_a^\beta$) (Acworth and Griffiths, 1985).

In this study, we use the apparent resistivity data measured by the Offset-Wenner array on the Zemun loess plateau (Serbia), as well as data from the square array (Bogatić and Golubac, Serbia) and the Wenner tri-potential technique (Vrdnik, Serbia). The results of the data processing used to detect, measure, and reduce lateral effects in 1D inversion were studied on the measured data while the lateral effects affecting the 1D, 2D, and 3D inversion were studied on the synthetic models.

2. Methodology

2.1. The Wenner tri-potential resistivity technique

Electrical resistivity measurements are based on the injection of an electric current into the ground by two current electrodes (C_1C_2) and measuring the potential difference by two potential electrodes (P_1P_2) (Carpenter, 1955; Zhou, 2016). The tri-potential technique involves measuring three apparent resistivities (or resistances) using a four-electrode array (Carpenter and Habberjam, 1956). Therefore, measurements with the Wenner tri-potential technique consist of measuring three apparent resistivities ($\rho_a^\alpha, \rho_a^\beta, \rho_a^\gamma$) with a four-electrode, colinear, equidistant array (Fig. 1a). Measured apparent resistivities are denoted by ρ_a^α ($C_1P_1P_2C_2$), ρ_a^β ($C_1C_2P_1P_2$) and ρ_a^γ ($C_1P_1C_2P_2$) (Carpenter, 1955; Carpenter and Habberjam, 1956; Acworth and Griffiths, 1985).

The percent of lateral effects in a certain segment of the apparent resistivity sounding curve is described as the ratio of lateral inhomogeneity (*LIR*). The interference of lateral effects can be measured for a part of the resistivity sounding curve using a series of *LIR* values. The root mean square of the lateral inhomogeneity ratio is used to measure the index of lateral inhomogeneity (*LII*):

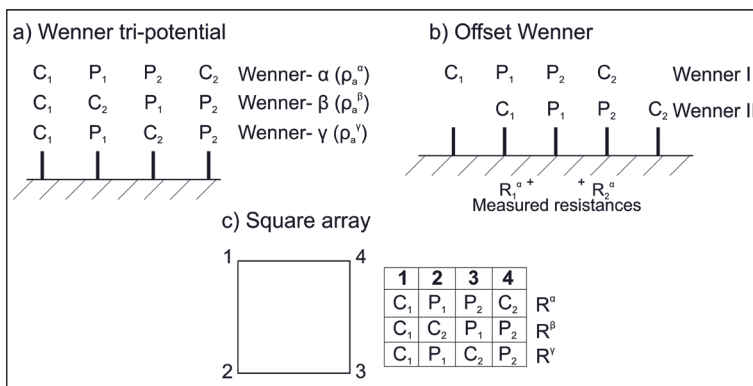


Figure 1. (a) The Wenner tri-potential technique; (b) The Offset Wenner array; (c) The square array.

$$LII = \sqrt{\frac{1}{m} \sum_{i=1}^m (LIR)^2}, \tag{1}$$

where m is the upper limit of summation. The load of the entire resistivity sounding curve by lateral effects is determined using the index of lateral inhomogeneity. The authors of the study on lateral effects identification, quantification, and reduction suggested criteria for classifying the lateral effects load Habberjam and Watkins (1967a):

- $0\% \leq LII \leq 20\%$ – good quality sounding curves,
- $20\% \leq LII \leq 50\%$ – medium quality sounding curves,
- $LII \geq 50\%$ – bad quality sounding curves.

The processed resistivity sounding curve (ρ_a°) is the result of lateral effect reduction, *i.e.*, it is a precise method of adjusting measured apparent resistivities to the nearest 1D model. The processing covariance (PC) is used to calculate the deviation of the processed sounding curve ρ_a° from the measured sounding curve ρ_a^m . Values of processing covariance are higher if the measured sounding curve is influenced by lateral effects. The processing covariance is calculated as:

$$PC = \frac{1}{m} \sum_{i=1}^m [(\log \rho_a^m)_i - (\log \rho_a^\circ)_i]^2, \tag{2}$$

where m is the upper limit of summation. If the value of processing covariance decreases, the results of quantitative 1D inversion become more accurate. Processing covariance values are dependent on measurement precision and local geology, *i.e.*, how much does the local geology deviate from a 1D subsurface model.

2.2. The Offset Wenner array

The Offset-Wenner array (Fig. 1b) was introduced in 1981 by Barker as a five electrode, colinear, non-focused, superposed array (Szalai and Szarka, 2008) consisting of two mutually shifted Wenner arrays (Barker, 1981; White and Scott, 1988). The Offset-

Wenner array consists of two distinct measurements, one with electrodes 1-2-3-4 (Wenner D) and the other with electrodes 2-3-4-5 (Wenner II). Measurement with the Offset-Wenner array can also be done in a tri-potential manner, as it was depicted in Fig. 1a where a total of six resistances can be obtained ($R_1^\alpha, R_2^\alpha, R_1^\beta, R_2^\beta, R_1^\gamma, R_2^\gamma$). The main advantage the Offset-Wenner array possesses is the detection and elimination of lateral resistivity effects (Barker, 1981; Nunn et al., 1983; Busby and Peart, 1997). The Offset-Wenner array also enables automatization of data acquisition as well as the extrapolation of the calculated apparent resistivity sounding curve (Barker, 1981). Several authors such as White and Scott (1988) and Chermali et al. (2018) considered the accuracy of the calculated interpolated data of the Offset-Wenner sounding curve, which was obtained on distances “ a ” generated by the relationship for the electrode spreading (S):

$$S = 3 \times 2^n \times a, n = 0, 1, 2, 3\dots \tag{3}$$

and the accuracy of the extrapolated data point, which was obtained for the distance of $2a$ with the respect to the last measured data point “ a ” (White and Scott, 1988; Chermali et al., 2018).

Electrical resistances R_1^α and R_2^α measured with two shifted Wenner arrays (Wenner I and Wenner II) that allow for the detection of near-surface lateral inhomogeneities and subsequent reduction of those effects by averaging R_1^α and R_2^α . The offset error (E_{of}) is used to calculate lateral effects:

$$E_{of} = \frac{R_1^\alpha - R_2^\alpha}{R_1^\alpha} \tag{4}$$

The electrode spreading (S) is done according to the rule:

$$S = 2^n \times a, n = 0, 1, 2, 3\dots, \tag{5}$$

with the inner electrode taking the place of the former outer electrode (Fig. 2).

The electrical resistances $R_a, R_b,$ and $R_c,$ which are measured with electrodes 1, 2, 4, and 5, are primarily used to calculate the resistance R_a value at electrode spacing according to the rule:

$$S = 3 \times 2^n \times a, n = 0, 1, 2, 3. \tag{6}$$

Even though the Offset Wenner and the tri-potential Wenner arrays are the most efficient for the reduction of lateral effects (apart from the square array), a limited number of data points on the sounding curve lessen the quality of detection and reduction of lateral effects for the Offset Wenner array.

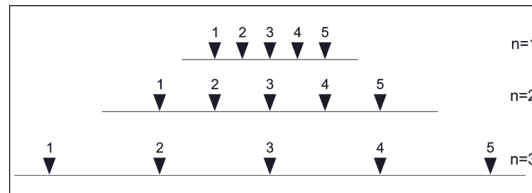


Figure 2. Offset Wenner array resistivity sounding.

2.3. The square array

The square array (Fig. 1c) was first introduced in 1967 as a non-collinear, non-superposed and non-focused array (Szalai and Szarka, 2008). The primary purpose of the square array was addressing directional dependence of resistivity measurements that arise with conventional collinear and symmetrical arrays such as the Wenner or Schlumberger array (Habberjam and Watkins, 1967a). The square array measurements are done in a tri-potential manner, with resistance measurements taken in two perpendicular directions (R^α , R^β) and a control measurement in the diagonal direction (R^γ). The square array can then be rotated by 45° , and measurements repeated, *i.e.*, the crossed square array. Alternatively, the square array can be expanded by $a\sqrt{2}$ and measurements for a larger electrode spacing can be performed. Apparent resistivities can be obtained from measured resistances by simple equations given in Habberjam and Watkins (1967b).

Averaging the ρ_a^α and ρ_a^β sounding curves, which were determined by the square array, Habberjam and Watkins (1967b) established a numerical method for reducing lateral effects.

In the case of a 1D subsurface model, the apparent resistivities measured in two perpendicular directions are identical (in fact, they vary with the value of the measurement error). Only the presence of inhomogeneities causes the two sounding curves to diverge. Such divergences are calculated for each data point in the sounding curve using a simple equation for azimuthal inhomogeneity ratio (*AIR*). The azimuthal inhomogeneity index is calculated by taking the root-mean-square value from the azimuthal inhomogeneity ratio:

$$AIR(a) = \frac{2R^\gamma(a)}{R^\alpha(a) + R^\beta(a)}, \quad (7)$$

$$AII(a) = \sqrt{\frac{1}{m} \sum_{i=1}^m (AIR)^2}, \quad (8)$$

where m is the upper limit of summation.

3. Results and discussion

3.1. Model equivalence analysis with the Wenner tri-potential technique

For two geoelectric sections, it can be said to be equivalent when the result of layering produces two practically identical sounding curves (Zohdy, 1949; Bhattacharya and Shalivahan, 2016). For example, two H-type curves can be equivalent (with respect to the longitudinal conductivity – S) if the middle layer has a small thickness and resistivity compared to the other two adjacent layers (Bhattacharya and Parta, 1968; Sanuade et al., 2019).

The magnitude of the minima of the sounding curves ρ_a^α , ρ_a^β and ρ_a^γ vary due to different depths of investigation (Acworth and Griffiths, 1985). The 1D inversion of the ρ_a^β sounding curve will determine the best true resistivity value of the middle layer, which is critical for a correct definition of the thickness of the second layer according to the principle of equivalence. For the 1D, 2D, and 3D geoelectrical surveys, the equivalence principle applies. The ambiguity of the inversion results increases when the thickness of the second layer is smaller than the thickness of the first (surface) layer.

Figure 3a demonstrates a case where the second layer resistivity is modified by 50% using the equivalence principle for contrast resistivities $\rho_2/\rho_3 \geq 10$ (Maillet, 1947; Koefoed, 1969), and the thickness h_2 changes linearly to 22.5 m (that is equivalent to the increase of 50% as well). In the presence of a thin layer ($h_2 \leq h_1$), and/or when the layer thicknesses in the model do not increase progressively with depth, the range of equivalence can be very broad.

To analyze the range of equivalence for the 1D model from Acworth and Griffiths (1985), sections of apparent resistivity for the Wenner α , β , and γ arrays were calculated, the dataset was then used in the 2D inversion. Figure 3b displays the 1D model from Acworth and Griffiths (1985) and the apparent resistivity section for the Wenner α (Fig. 3c) and β array (Fig. 3d).

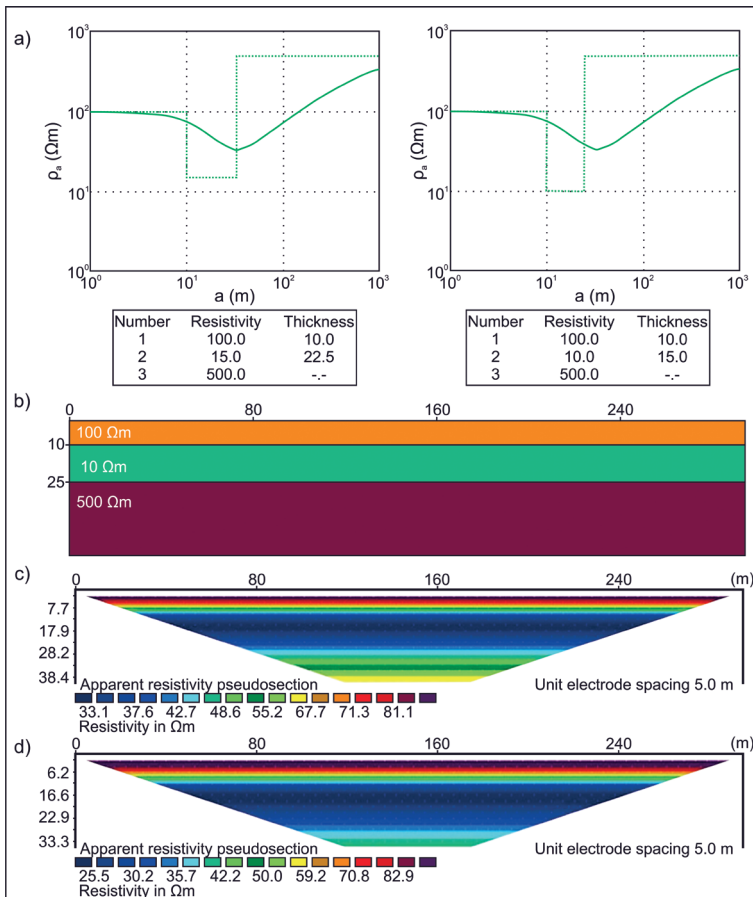


Figure 3. (a) Equivalent models obtained by the 1D inversion of the sounding curve from Acworth and Griffiths (1985); (b) 1D subsurface model from Acworth and Griffiths (1985); (c) apparent resistivity section for the Wenner α array; (d) apparent resistivity section for the Wenner β array; ρ_a – apparent resistivity.

The apparent resistivity section from Fig. 3c indicates that the minimal resistivity value is 33 Ωm . This is confirmed with the minimal resistivity from the Wenner α sounding curve from Acworth and Griffiths (1985). The minimum apparent resistivity value of 25.5 Ωm can be seen in the apparent resistivity section measured with the Wenner β array for the same subsurface model. The lower resistivity value is due to the Wenner- β array’s lower depth of investigation (Edwards, 1977; Loke, 2022), thus the measured apparent resistivities are more affected by the second low resistivity layer (10 Ωm).

The apparent resistivity pseudosections (Figs. 3c and 3d) were utilized for 2D inversion with a total of three iterations for *RMS* convergence. The results of 2D inversion obtained by Wenner α , β , and γ arrays (Fig. 4) confirm that inversion of the Wenner- β array yields the best true resistivity value and thickness of the second layer. The Wenner- γ array is the most fitting method for evaluating the true resistivity value of the substratum.

The previous example used the resistivity model from Acworth and Griffiths (1985), but the thickness of the first and second layers was altered to obtain additional information about the capabilities of the Wenner α , β and γ arrays and to conduct statistical analysis (Tab. 1). The initial model (M1) consisted of a 100 Ωm first layer with a 5 m relative thickness, while the second, low resistivity (10 Ωm) layer had a 20 m thickness. In each instance, the resistivity of the substratum remained the same (500 Ωm). Each model after the initial model (M1) had increased thickness of the first layer by 2.5 m and decreased thickness of the second layer by 2.5 m. Seven models were constructed in total.

Each model was constructed using the RES2DMOD forward modelling software (Loke, 2016) with 60 electrodes total, a unit electrode spacing of 5 m, and four nodes per

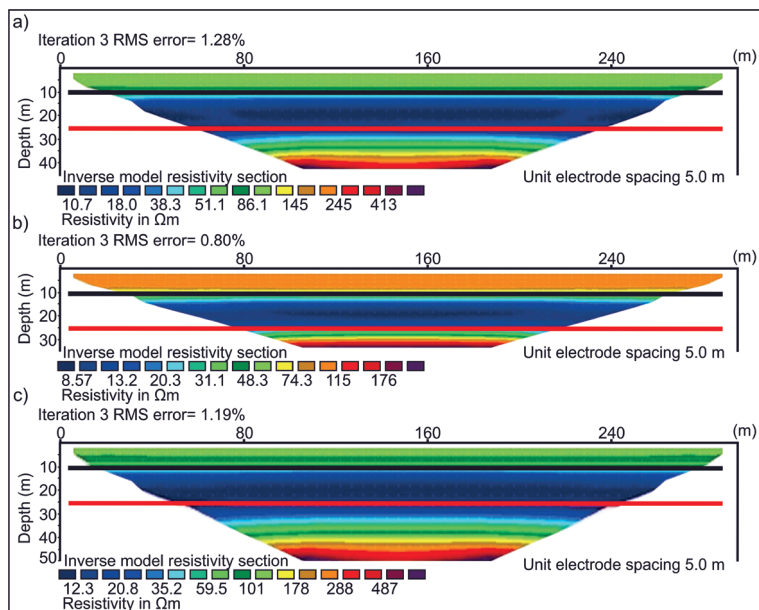


Figure 4. Results of 2D inversion for (a) Wenner α ; (b) Wenner β ; (c) Wenner γ array. Red line – the true depth to the high resistivity substratum; black line – the depth to the second layer.

Table 1. Forward model parameters for the statistical analysis of Wenner α , β and γ arrays based on the model given in Ackworth and Griffiths (1985); h_1 and h_2 – thickness of the first and layer; ρ_1 , ρ_2 and ρ_3 – resistivities of the first, second and third layer.

Model name	Model parameters				
	h_1 [m]	ρ_1 [Ωm]	h_2 [m]	ρ_2 [Ωm]	ρ_3 [Ωm]
M1	5	100	20	10	500
M2	7.5	100	17.5	10	500
M3	10	100	15	10	500
M4	12.5	100	12.5	10	500
M5	15	100	10	10	500
M6	17.5	100	7.5	10	500
M7	20	100	5	10	500

electrode spacing. In the first 30 meters of the forward model, the horizontal and vertical dimensions of the mesh cells were 2.5 m. After the 30th meter, the dimensions of the cells were generated automatically. To simulate real measurements, random noise of $\pm 2\%$ was added to each individual model after forward modeling. The least squares inversion parameters in RES2DINV (Loke, 1994) software for each model were constant. The maximum number of iterations for the RMS convergence were increased to five with the standard Gauss-Newton was the optimization method. The models obtained after inversion were exported to Surfer (Golden Software), thus allowing for additional statistical analysis. The data corresponding to the first, second, and third layers of the forward model were extracted from the model obtained after inversion and analyzed.

Table 2 provides a brief statistical summary of the minimum, maximum, and mean apparent resistivity of the Wenner α , β and γ arrays for all three layers of the seven models. To estimate the accuracy of the resistivity measurements for the three Wenner arrays, the absolute error (*ABSE*) was calculated using the difference between the calculated mean after inversion and the true resistivity of a given layer.

For the first layer, the Wenner β array has, on average, the smallest absolute error with a value of 5.51 Ωm , followed by the Wenner α array with 6.85 Ωm and then the Wenner γ array with 8.11 Ωm . In addition, the Wenner γ array displays (on average) the lowest mean resistivity value for the first layer with 93.32 Ωm , followed by the Wenner α array with 94.46 Ωm and the Wenner α array with 96.02 Ωm . Since we are aware that the second layer has a low resistivity, the values of the mean resistivity for the three arrays correspond well with the median depth of investigation for the three arrays that is given in Loke (2022). The array with the greatest resistivity corresponds to the Wenner β with the smallest depth of investigation, and vice versa.

In the case of the second layer, the Wenner β array has the smallest absolute error at 12.80 Ωm , while the Wenner α and γ arrays have values of 16.89 Ωm and 17.95 Ωm , respectively. Since the high resistivity substratum is first observed on the array with the greatest depth of investigation, the Wenner γ array now has the highest mean resistivity.

In the third layer case, the Wenner γ array with the lowest *ABSE* values was closest for determining the true substratum resistivity. Because the *ABSE* values are calculated based on the mean apparent resistivity, the *ABSE* values are quite high; however, the

Table 2. Statistical summary for the three Wenner arrays and the three layers of the model; ρ_{\min} – minimal resistivity; ρ_{\max} – maximal resistivity; ρ_{mean} – mean resistivity; ABSE – absolute error.

Array	Name	Layer 1				Layer 2				Layer 3			
		ρ_{\min} [Ωm]	ρ_{\max} [Ωm]	ρ_{mean} [Ωm]	ABSE [Ωm]	ρ_{\min} [Ωm]	ρ_{\max} [Ωm]	ρ_{mean} [Ωm]	ABSE [Ωm]	ρ_{\min} [Ωm]	ρ_{\max} [Ωm]	ρ_{mean} [Ωm]	ABSE [Ωm]
Wenner Alpha	M1	90.04	113.44	102.35	2.35	6.18	51.96	16.60	6.60	20.17	578.79	120.70	379.30
	M2	67.23	131.16	97.34	2.66	7.54	43.97	15.89	5.89	8.85	694.79	95.94	404.06
	M3	53.47	133.11	93.18	6.82	10.71	36.33	17.68	7.68	10.71	567.16	64.05	435.95
	M4	70.83	126.62	102.24	2.24	14.26	60.06	28.18	18.18	14.26	495.09	51.65	448.35
	M5	39.66	133.33	89.33	10.67	18.24	41.32	27.93	17.93	18.24	502.45	48.58	451.42
	M6	41.82	121.56	88.28	11.72	25.46	42.72	35.12	25.12	24.68	391.51	53.66	446.34
	M7	50.16	122.51	88.47	11.53	42.57	51.15	46.82	36.82	37.62	382.71	70.55	429.45
Average	59.03	125.96	94.46	6.85	17.85	46.79	26.89	16.89	19.22	516.07	72.16	427.84	
Wenner Beta	M1	96.76	115.85	105.36	5.36	5.95	48.27	16.17	6.17	30.83	325.67	107.52	392.48
	M2	66.54	137.87	97.69	2.31	7.25	41.98	15.08	5.08	21.09	487.50	110.70	389.30
	M3	52.21	127.06	92.56	7.44	8.74	38.50	16.91	6.91	12.93	414.57	85.87	414.13
	M4	49.82	122.49	91.06	8.94	10.83	35.69	19.21	9.21	11.83	317.50	64.48	435.52
	M5	41.18	126.71	90.88	9.12	12.95	36.84	22.01	12.01	13.56	302.99	54.19	445.81
	M6	60.51	128.53	98.16	1.84	18.45	52.26	32.54	22.54	18.45	210.04	52.21	447.79
	M7	50.17	131.37	96.46	3.54	26.20	52.59	37.69	27.69	26.20	230.85	64.88	435.12
Average	59.60	127.13	96.02	5.51	12.91	43.73	22.80	12.80	19.27	327.02	77.12	422.88	
Wenner Gamma	M1	98.99	113.42	105.00	5.00	6.41	50.86	17.19	7.19	31.65	704.80	156.50	343.50
	M2	70.48	133.22	98.96	1.04	8.36	45.12	16.77	6.77	17.88	1000.76	127.56	372.44
	M3	56.14	134.94	94.56	5.44	11.22	41.08	18.88	8.88	14.01	809.38	82.06	417.94
	M4	45.63	129.35	90.68	9.32	16.28	40.76	23.66	13.66	16.36	541.55	58.66	441.34
	M5	42.26	134.41	89.16	10.84	20.96	41.00	29.35	19.35	22.23	438.36	54.48	445.52
	M6	43.25	130.68	87.50	12.50	30.58	49.48	38.80	28.80	30.58	369.88	60.86	439.14
	M7	49.43	134.11	87.40	12.60	44.93	58.60	51.00	41.00	47.02	345.43	81.56	418.44
Average	58.02	130.02	93.32	8.11	19.82	46.70	27.95	17.95	25.68	601.45	88.81	411.19	

Wenner γ array exhibited the lowest *ABSE* values and the highest mean resistivity values with 88.81 Ωm , as well as the highest maximal resistivity in the third layer with the average maximal resistivity value at around 600 Ωm .

Using the data from the inverted models, a further statistical parameter was calculated. The coefficient of variation (*CV*) is a measure of data dispersion derived from the ratio of the standard deviation to the mean. The *CV* values for the three Wenner arrays on all three layers are displayed in Tab. 3.






The *CV* value for the first layer is the lowest on average among the three arrays, ranging from 18.78% to 21.8%. The *CV* value increases as the layers progress, reaching a maximum between 90% and 110% for the third layer. It should also be noted that the *CV* values for the second layer exhibit a decreasing trend, whereas the trends for the first and third layers are different. Consequently, Pearson's and Spearman's correlation coefficients could be utilized to conduct a correlation analysis between the parameters of the

Table 3. Coefficient of variation (CV) values for the three Wenner arrays and all the three layers of the model.

Array	Name	Layer 1 CV [%]	Layer 2 CV [%]	Layer 3 CV [%]
Wenner Alpha	M1	3.83	78.50	87.29
	M2	20.44	64.02	116.08
	M3	26.17	44.19	132.56
	M4	14.97	49.19	132.12
	M5	26.43	21.71	130.84
	M6	25.61	11.03	97.96
	M7	22.64	5.08	75.35
	Average	20.01	39.10	110.32
Wenner Beta	M1	4.31	74.26	70.29
	M2	21.87	60.18	90.81
	M3	24.47	53.06	109.08
	M4	23.80	42.41	106.56
	M5	23.40	31.77	103.61
	M6	15.04	31.80	83.19
	M7	18.60	20.74	72.95
	Average	18.78	44.89	90.93
Wenner Gamma	M1	3.11	75.03	87.33
	M2	19.47	60.23	119.33
	M3	24.53	44.26	134.32
	M4	27.76	28.18	120.81
	M5	28.06	18.03	103.76
	M6	25.97	11.81	80.05
	M7	23.74	5.59	59.50
	Average	21.80	34.73	100.73

Table 4. Correlation of the thickness of the second layer (h_2) and the coefficient of variation (CV); r – Pearson’s correlation coefficient; r_s – Spearman’s correlation coefficient.

Correlation of h_2 and CV						
Array/parameter	Layer 1		Layer 2		Layer 3	
	r	r_s	r	r_s	r	r_s
Wenner Alpha	-0.63	-0.50	0.98	0.96	0.24	0.25
Wenner Beta	-0.30	0.00	0.98	0.96	0.06	0.04
Wenner Gamma	-0.69	-0.54	0.98	1.00	0.56	0.54

Legend (after Schober et al., 2018)		
Description	Color	Range
Negligible		0.0–0.1
Weak		0.1–0.39
Moderate		0.4–0.69
Strong		0.7–0.89
Very strong		0.9–1.0

second layer thickness (h_2) and the CV. The correlation analysis can be seen in Tab. 4, where the correlation ranges from Schober et al. (2018) are used. For the first layer, the Wenner α and γ arrays exhibit a moderate, negative correlation, which indicates that as the thickness of the second layer decreases, the data dispersion (CV) increases. The Wenner β array yields contradictory results, as the Pearson correlation coefficient suggests a negative, weak correlation whereas the Spearman correlation coefficient suggests a nonexistent correlation. The second layer revealed a very strong correlation from all three Wenner arrays, *i.e.*, as the second layer’s thickness decreases, so does the resistivity data dispersion. The results of the third layer are comparable to the average depth of investigation. The Wenner β array with the shallowest depth of investigation has negligible correlation between the second layer’s thickness and data dispersion. In contrast, the array with the greatest depth of investigation has the strongest positive correlation with data dispersion in relation to the thickness of the second layer. When the thickness of the second layer is altered, the array with the greatest depth of investigation will have the largest data discrepancy of deeper layers.

The results presented here only illustrate one situation, in which resistivity decreases with depth, or the H-type curve if 1D resistivity sounding were performed. To provide a general recommendation for field surveys, additional modeling and statistical analysis are required in which various models with varying thicknesses are analyzed. In addition, modern survey techniques employing multicore cables enable relatively rapid data acquisition with various arrays. Due to this, all three Wenner arrays can be acquired with relative ease, where electrode switching is performed by a field computer and no additional effort is required by field technicians when switching from the Wenner α to the β or γ arrays. Additional models are not presented in this research paper, but the equivalence case illustrated by the modified Ackworth and Griffiths (1985) model should be the focus of future research.

3.2. Lateral effects and 1D inversion

It is obvious from the theoretically determined sounding curves (ρ_a^α , ρ_a^β , and ρ_a^γ) that there is no common general pattern in the presence of a vertical contact between two mediums of different resistivities. However, the sounding curves are better described by mutual convergence and divergence. The sounding curve for ρ_a^β diverges from the ρ_a^α and ρ_a^γ sounding curves, which are mutually comparable (Fig. 5a).

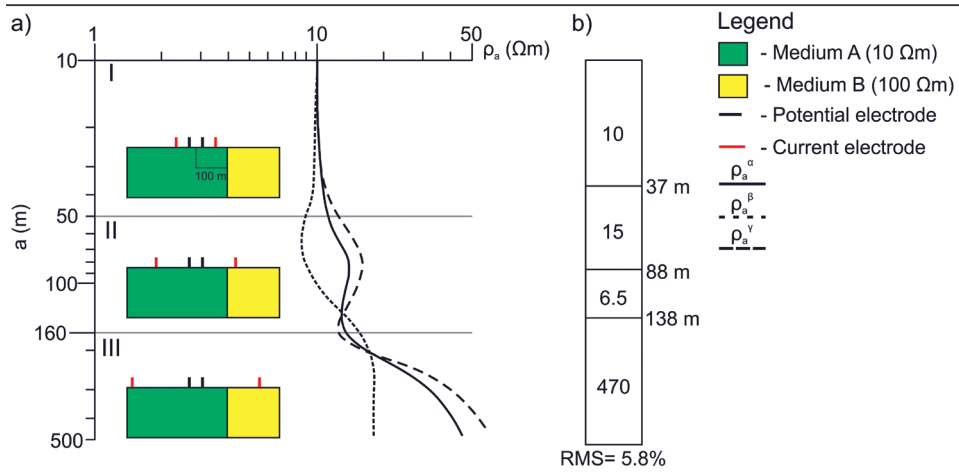


Figure 5. (a) Sounding curves of Wenner α , β and γ array over a vertical contact; (b) 1D inversion of the ρ_a^α sounding curve; ρ_a – apparent resistivity.

The additional sounding curves ρ_a^β and ρ_a^γ allow for the identification of lateral effects that are primarily bound to the current electrode, which is not possible with normal-sounding methodology. Figure 5a shows the divergence of the sounding curves ρ_a^α , ρ_a^β and ρ_a^γ as the current electrode approaches the vertical contact. The same divergence can also be seen as the current electrode transitions into the second medium with a higher resistivity value (segment I and II in Fig. 5a). The lateral effect could only be detected using the ρ_a^α sounding curve with the electrode spacing (a) greater than 160 m and when the potential electrode moves into the second medium. Only the slope of the last segment of the ρ_a^α sounding curve could suggest a vertical contact. The rest of the diagram could be interpreted as the sounding curve over a horizontally layered subsurface. Quantitative 1D inversion, which would be performed after smoothing of the ρ_a^α sounding curve, would result in an erroneous four-layer subsurface model (Fig. 5b). The last section of the ρ_a^α sounding curve is responsible for the higher root-mean-square (RMS) value of around 5.8%.

Aside from qualitative analysis of the data, quantitative determination of lateral effects is feasible. If the lateral effects are not overly articulated, a numerical adaptation of the data to a 1D subsurface model is possible. The lateral effect is quantified using a numerical approach proposed by Habberjam and Watkins (1967a) calculating the lateral inhomogeneity ratio (LIR) and the index of lateral inhomogeneity (LII).

Figure 6a shows the effects of geoelectrical sounding processing in Vrdnik (Serbia) using the Wenner tri-potential technique. Vrdnik is located in northern Serbia on the southern slope of the Fruška gora mountain. The geology of the Vrdnik locality consists of Miocene limestones, sandstones, conglomerates as well as freshwater lacustrine coal beds, while the uppermost layers are composed of Quaternary gravel, sands, and Pleistocene loess (Čičulić-Trifunović and Rakić, 1971). The example shows an increase in processing covariance as the index of lateral inhomogeneity rises.

The root-mean-square measurement accuracy value in the studied area is $\pm 4\%$. This implies that the minimal expected processing covariance value should be 0.0003 in the case of a 1D subsurface model. Any deviation of the local geology from a 1D model will lead to the increase of processing covariance. With values of the lateral inhomogeneity index of 110% (Fig. 6a), the processing covariance is 0.0920 confirming that the 1D inversion of sounding data is inadequate.

The value of processing covariance is significantly lower in Golubac (Serbia) (Figure 6b). Golubac is situated on the border between Serbia and Romania on the right bank of the Danube River (eastern Serbia). The sediments of the Golubac locality are comprised of middle Miocene sands, sandstones, conglomerates, and gravelly alluvium (Kalenić et al., 1973).

The measurements were carried out with the square array since the square array is less sensitive to lateral effects than colinear arrays (Habberjam, 1972a; Sretenović et al.,

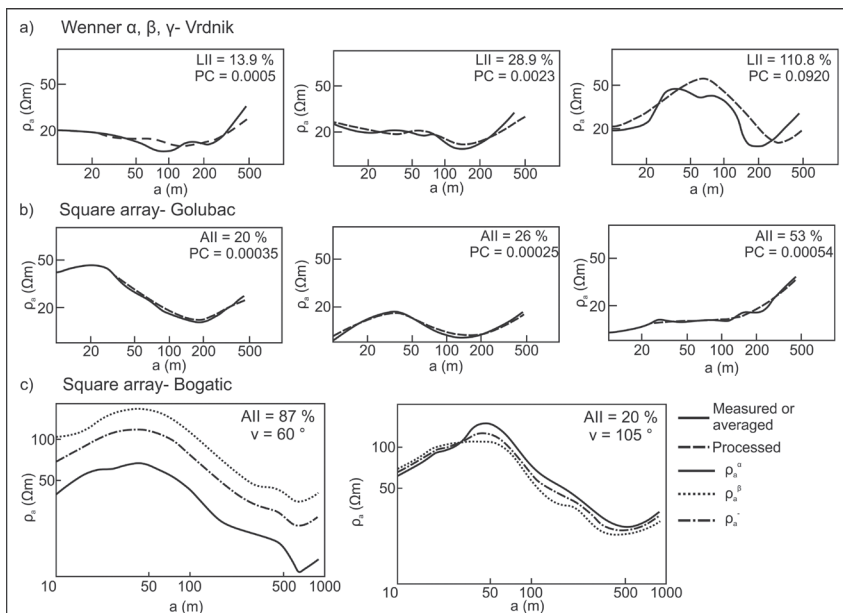


Figure 6. (a) Geoelectrical processing results with the Wenner tri-potential technique in Vrdnik (Serbia); (b) Geoelectrical processing results with the square array in Golubac (Serbia); (c) Geoelectrical sounding curve in Bogatić (Serbia) with two different square array orientations; ρ_a – apparent resistivity.

2014; Udosen and George, 2018; Sretenović and Arnaut, 2019). By averaging the ρ_a^α and ρ_a^β sounding curves, local near-surface inhomogeneities are reduced. The averaged sounding curve ρ_a^{ave} is of higher quality than sounding curves obtained with colinear electrode arrays. The processed sounding curve ρ_θ° is adapted to the conditions of a horizontally layered subsurface. Lower processing covariance values are a function of the numerical method for lateral effects reduction being based on the use of the square array, while the Wenner array requires two sub-arrays (α and γ).

Figure 6b shows the results of the square array geoelectrical sounding processing in Golubac. The processed sounding curves show that they are more adapted to the 1D horizontally layered subsurface half-space than data obtained using the Wenner tri-potential technique. The processing covariance is used to compare the results of the two arrays. The distinct differences between the ρ_a^α and ρ_a^β sounding curves are mirrored by high values of azimuthal inhomogeneity index (*AII*). The differences could be the result of anisotropy or deeper 2D structures. Regardless of the relatively low processing covariance values, interpretation with the 1D model assumption can be inadequate in both cases. In that case, further investigations with the crossed square array are required.

Because both the ratio and index of azimuthal inhomogeneity are dependent on the orientation of the square array relative to the subsurface strike or electrical anisotropy, low values of the *AII* do not guarantee that the subsurface is horizontally layered, which was supported by the Bogatić field example (Fig. 6c). Bogatić is situated in western Serbia, and its geology consists primarily of fluvial structures formed by the accumulation of sediments from the rivers Drina and Sava. The accumulated river terrace sediments have a relative thickness of 3 to 10 meters and are primarily composed of sandy gravel, gravel, siltstones, and sandstones which was deposited during the late Pleistocene and early Holocene. Deeper levels mainly consist of sands, silty sands, and clays. (Rajčević, 1978; Vrhovčić et al., 1986).

Both *AIR* and *AII* are theoretically equal to zero when the square array orientation relative to electrical strike is 45 degrees. This implies that the presence of a structure or electrical anisotropy will not be detected (Habberjam, 1975). The averaged sounding curve from four directions (ρ_x) is independent of the array orientation (Habberjam and Watkins, 1967b; Habberjam, 1972b; Habberjam and Jackson, 1974), and it reduces lateral effects by averaging apparent resistivity data from four directions. After analyzing electrical anisotropy using the crossed square array the mean apparent resistivity sounding curve (ρ_m) can be calculated. The ρ_m sounding curve is orientationally independent of the square array orientation but takes into consideration the effects of electrical anisotropy (Habberjam and Watkins, 1967b).

The Wenner tri-potential array cannot differentiate local 2D and 3D near-surface inhomogeneities from the effect of deeper 2D structures during lateral effects detection, quantification, and reduction. This unavoidable noise can be reduced during the measurement phase, but not during the phase of numerical processing (Barker, 1981; Barker, 1989). Because of the need for more efficient data acquisition (just half of the points on the sounding curve are measured while the rest are calculated), this technique was not chosen.

The offset Wenner array was used to perform an experimental survey. The survey included measuring resistances R_1^α and R_2^α in every data point of the apparent resistivity sounding curve, and then comparing measured and theoretically calculated apparent resistivity values using Barker's method (1981). Horizontal stratification was expected on the Zemun loess plateau, where the survey was conducted. The Zemun loess plateau

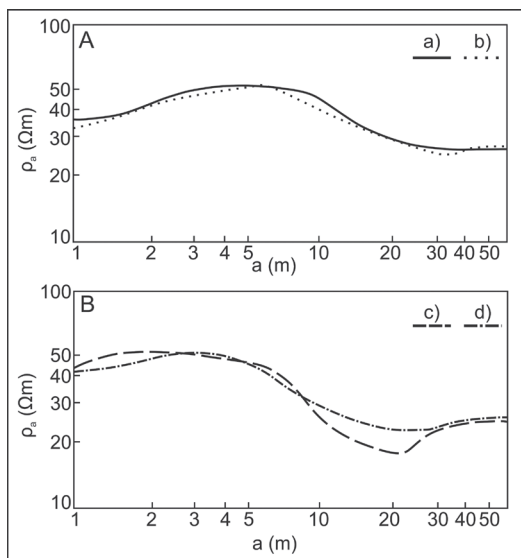


Figure 7. Sounding sites *A* and *B* during the verification of the offset Wenner array; *a*, *b*, *c*, and *d* – sounding curves; ρ_a – apparent resistivity.

is located in the Zemun municipality of Belgrade (Serbia) and consists of yellow, upper Pleistocene, sandy-clayey siltstones that formed during glacial periods (Marković et al., 1984). With that kind of local geology, the combined influence of deeper structures and local near surface inhomogeneities were avoided. To determine the subsurface architecture, sounding was carried out at two sounding sites, each with two perpendicular measurements. If the conditions of a 1D model are true, and there are no significant deviations of apparent resistivity sounding curves in both directions, the values of E_{of} can be related to local near-surface lateral inhomogeneities. Figure 7 displays the sounding curves R_1^α and R_2^α in both sounding sites with perpendicular measurements.

For sounding curves *a*, *b*, and *d*, the measured *RMS* values for E_{of} are around 6–10%, while the value for sounding curve *c* is around 35%. Since *c* is the sounding curve with the greatest lateral impact, the largest discrepancy between theoretically calculated and measured values is probable. Table 5 shows the previously mentioned divergences, as well as E_{inter} (interpolation error) and E_{extr} (extrapolation error).

Table 5. Estimated *RMS* values E_{of} and discrepancies between measured and theoretically calculated apparent resistivity values; E_{of} – offset error, E_{inter} – interpolation error, E_{extr} – extrapolation error; *a*, *b*, *c*, and *d* – sounding curves.

RMS/Sounding curve	a	b	c	d
E_{of} [%]	9.7	10.3	34.8	6.4
E_{inter} [%]	6.5	6.8	17.9	7.3
E_{extr} [%]	35.3	27	94.4	32.7

The highest values of discrepancy are determined for the sounding curve c which is the curve with the most lateral impact. For sounding curves, a, b, and d, the root-mean-square value E_{inter} is around 6.5%–7.5%, whereas, for diagram c, the value is around 18%. Extrapolation error (E_{extr}) *RMS* values are much larger, varying from 27% to 94%, which is unacceptable when measuring the extrapolated apparent resistivity value based on a twofold smaller value of a . Significantly higher E_{extr} values can not be explained solely by lateral effects that are insignificant on sounding curves a, b, and d. Furthermore, theoretically calculated extrapolated values of apparent resistivity show a significant difference from measured results. The reason for this is that the estimation of electrical resistance $R_a(2a)$ is based on the measurement of electrical resistances $R_a(a)$ and $R_c(a)$ (Eq. (9)), which can be nearly equivalent in the case of very small $R_a(2a)$. Calculating $R_a(2a)$ from $R_a(a)$ and $R_c(a)$ values that are nearly identical and very small will result in error magnification. When the apparent resistivity sounding curve is descending, the problem is more pronounced. However, when the apparent resistivity sounding curve is ascending and the difference between $R_a(a)$ and $R_c(a)$ is distinct, the problem is less pronounced.

$$R_a(2a) = 2[R_c(a) - R_a(a)]. \quad (9)$$

Discrepancies E_{inter} are primarily caused by lateral inhomogeneities near the surface. With larger electrode spacing (a) can lead to higher error values when measuring $R_b(a)$ and $R_b(2a)$ due to dipole-dipole array measurements ($C_1C_2P_1P_2$), where the electrical potential drops with the square of the distance, increasing the error when calculating the resistance $R_a(3a)$ (Eq. (10)).

$$R_a(2a) = 0.5R_a(2a) + R_b(2a) - R_b(a) + 0.5R_a(4a). \quad (10)$$

The sounding curve in Fig. 8d corresponds to the averaged sounding curve “a” in Fig. 8c. The measurements were done with the first (Wenner I) and second (Wenner II) array (α , β , and γ). The method of Habberjam and Watkins (1967a) is used to calculate lateral effects. The index of lateral inhomogeneities in measurements with the Wenner I array is 17.5%. The *LII* in measurements with the Wenner II array is 24.1%, confirming that the subsurface architecture closely resembles a 1D model. The *RMS* of the E_{of} , which is 9.7%, supports this model as well.

Figure 8d marks some of the lateral effects that are caused by the discrepancy from horizontal stratification. The *RMS* of a 1D inversion of the averaged ρ_a^{ave} sounding curve is 2%, implying that simple averaging is sufficient for lateral effects reduction (Fig. 8d).

3.3. 1D Lateral effects caused by 2D and 3D inhomogeneities

The development of fully automated 2D acquisition systems, which collect more resistivity data than conventional 1D survey systems, is attributable to the advancement of geoelectrical acquisition methodology. It was not possible to process large amounts of data in a short amount of time with 1D sounding methodology prior to the development of software packages such as the widely used RES2DINV. Due to this, 2D lateral effects can be easily displayed on models derived after inversion, yielding a more accurate depiction of the subsurface architecture than 1D surveys.

Figure 9a shows a horizontally layered subsurface model with 2D near-surface inhomogeneities. Sounding curves of apparent resistivity were calculated for that subsurface

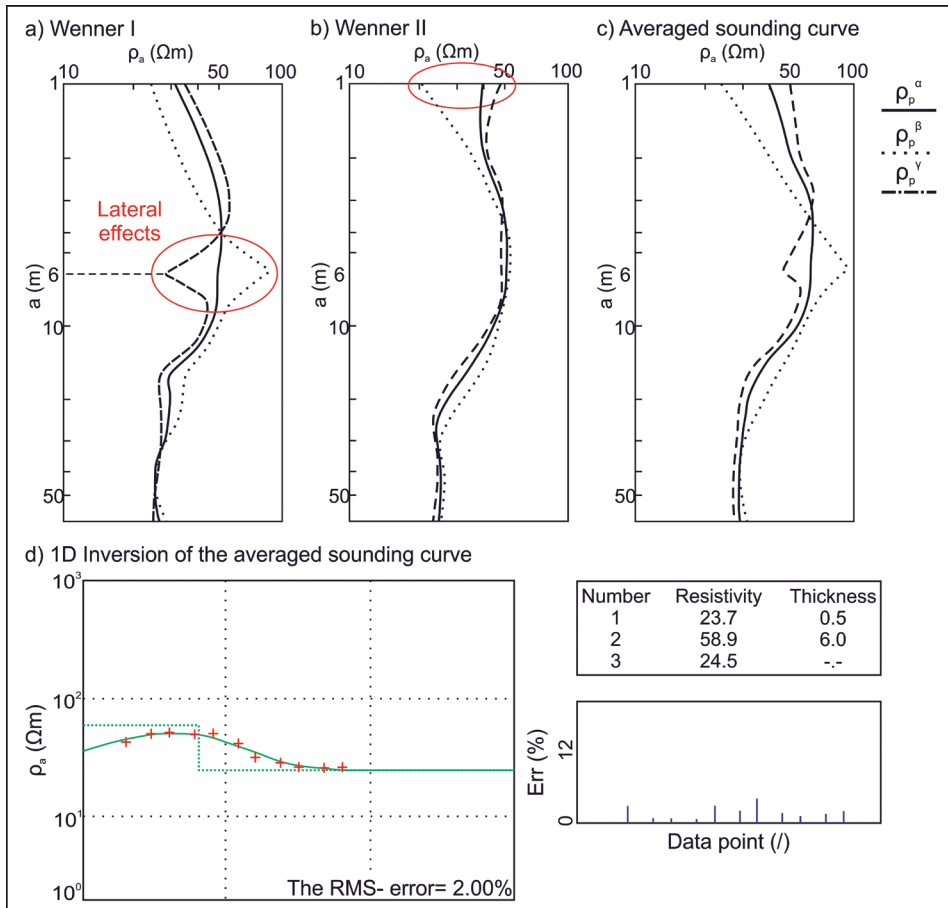


Figure 8. Sounding results with the offset Wenner array (Zemun loess plateau). Sounding curves α , β , and γ measured with (a) the Wenner I and (b) Wenner II array; (c) averaged sounding curves α , β , and γ from the measurements with the Wenner I and II array; (d) 1D inversion of the averaged sounding curve (Wenner I and II array averaged); ρ_a - apparent resistivity.

model using the offset Wenner array (α , β , and γ), as shown in Fig. 10. The lateral effects of 2D inhomogeneities that rise to the surface and are present at any depth level can be seen in sections of apparent resistivity. In the case of a 1D model, isolines of apparent resistivity should be horizontal. A distinct lateral effect was observed before the start of 2D inversion in RES2DINV. It was proposed that inversion be done with a cell size equal to half the size of the unit electrode spacing (a) in order to reduce the amount of lateral effects. The achieved *RMS* difference between observed and theoretically calculated data was 1.94%. Two additional parameters were applied:

1. vertical/horizontal flatness ratio = 0.3 and,
2. increase of damping factor with depth = 1.2.

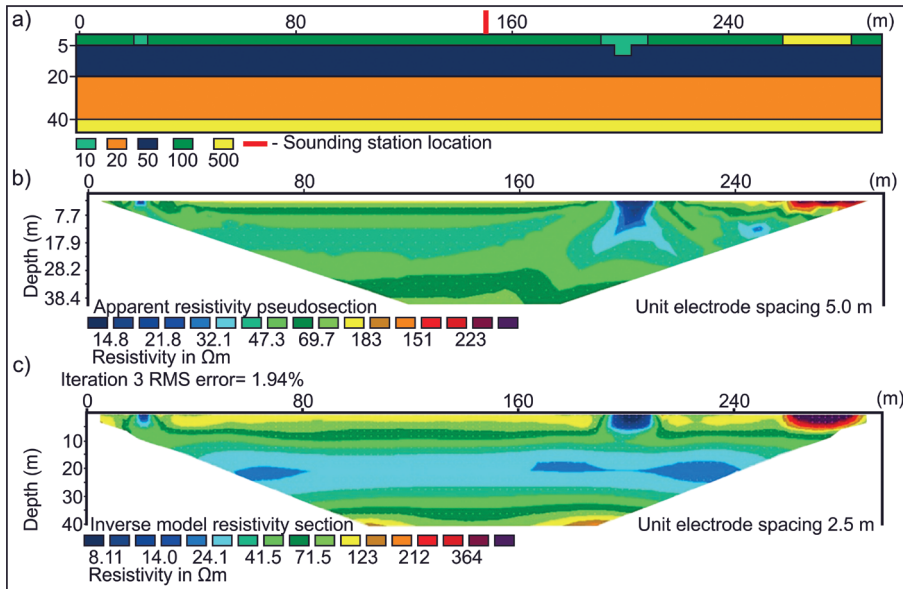


Figure 9. (a) Subsurface model; (b) Section of apparent resistivity; (c) 2D inversion results.

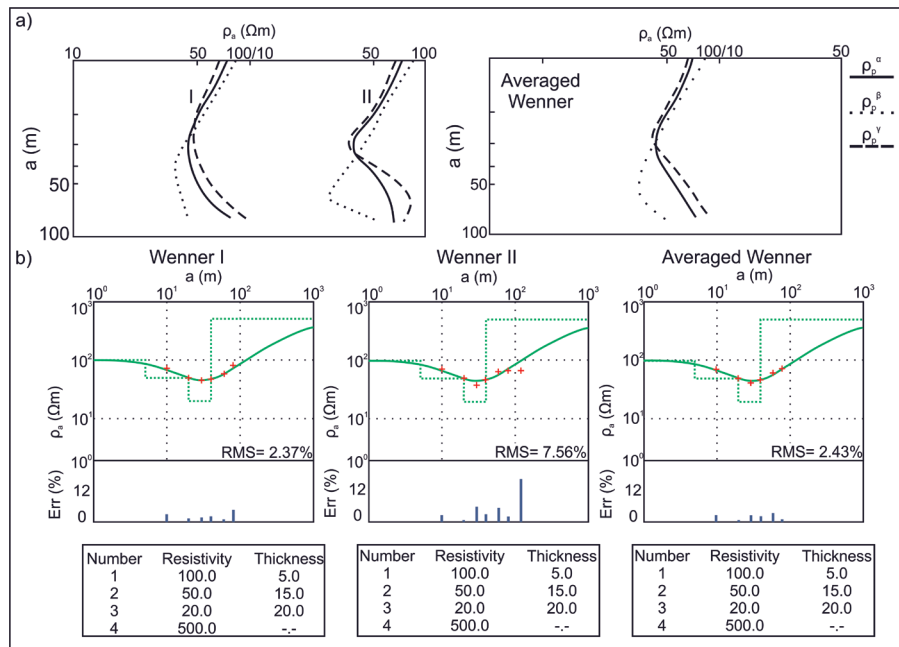


Figure 10. (a) Offset Wenner sounding curve for the α , β , and γ array; (b) 1D inversion results of Wenner I, Wenner II and Wenner averaged sounding curve; ρ_a – apparent resistivity.

Figure 10a shows the sounding curves of the Wenner α , β , and γ array in the case of averaging the apparent resistivity sounding curves of Wenner I and Wenner II to obtain two sounding curves. The sounding curves ρ_a^α , ρ_a^β , and ρ_a^γ , which correspond to the I and II Wenner arrays, indicate different lateral resistivities, which are later verified by processing results ($LII= 2.6\%$ and 6.2% , respectively). From a 1D inversion perspective, the *RMS* offset value of 10.85% closely resembles the 10% value that Barker (1981) suggests as good quality data. Figure 10b shows the discrepancies between apparent resistivity sounding curves calculated for the 1D subsurface without near-surface inhomogeneities and sounding curves with near-surface lateral inhomogeneities (Wenner I, Wenner II, and Wenner averaged). The sounding curve Wenner II has the greatest difference (*RMS* error = 7.6%), while the Wenner average and Wenner I data have a discrepancy of 2.4% . 1D inversion of the Wenner II array sounding curve did not create an *RMS* error of less than 5.5% .

Two boundary cases for detecting local 3D near-surface inhomogeneity (spherical shape) are shown in Fig. 11a. The apparent resistivity anomalies (α , β , and γ) cannot be observed when the depth to the sphere center (H) and the radius of the sphere (R) are approximately $H/R \geq 2$ and the geoelectrical profile passes through the projection of the center of the sphere to the surface ($B = 0$ m). This is because they do not reach the 10% threshold of the true resistivity of the surrounding material, in this case, $100 \Omega\text{m}$. This anomaly would be very small (around 5%) due to ambient noise and lateral effects from other near-surface inhomogeneities, so the sphere would not be detectable (Fig. 11a).

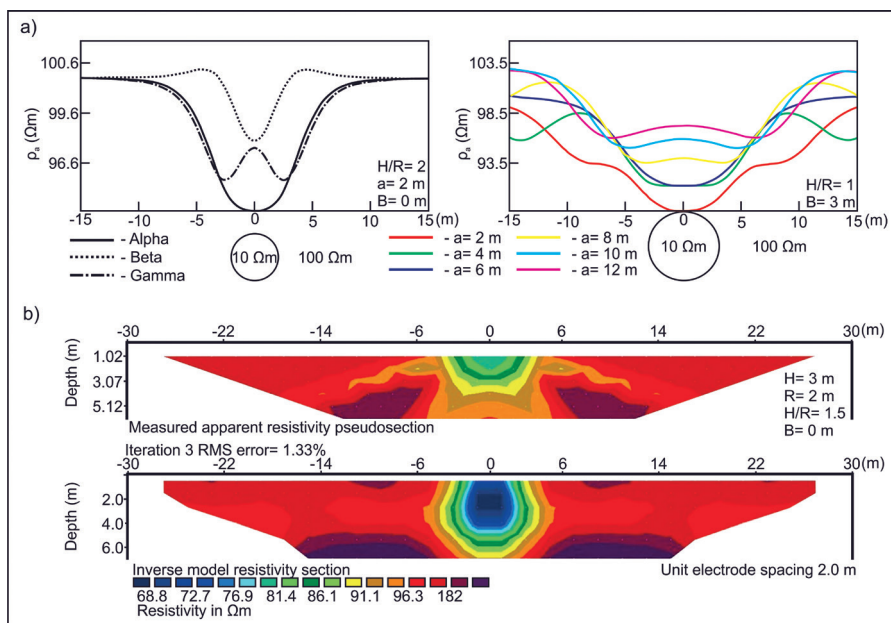


Figure 11. (a) Boundary cases during the detection of a spherical inhomogeneity ($H/R \geq 2$, (b) $H/R = 1$; (b) 2D survey over a spherical inhomogeneity section of apparent resistivity and a 2D inversion result; ρ_a – apparent resistivity.

The second boundary case is a spherical inhomogeneity that emerges to the surface ($H/R = 1$) when the detectability is certain. Anomaly intensity of a sphere is dependent on the Wenner electrode array separation (a) and the radius of the sphere (R). As a result, sounding curves for six depth levels were calculated, beginning with the smallest electrode spacing of $a_{min} = 2$ m and rising to $a_{max} = 12$ m. Figure 11a indicates that the anomaly intensity decreases as the ratio a/R increases. Since using larger measurements of the Wenner electrode array results in a higher impact of true resistivity of the environment ($100 \Omega\text{m}$) compared to the apparent resistivity.

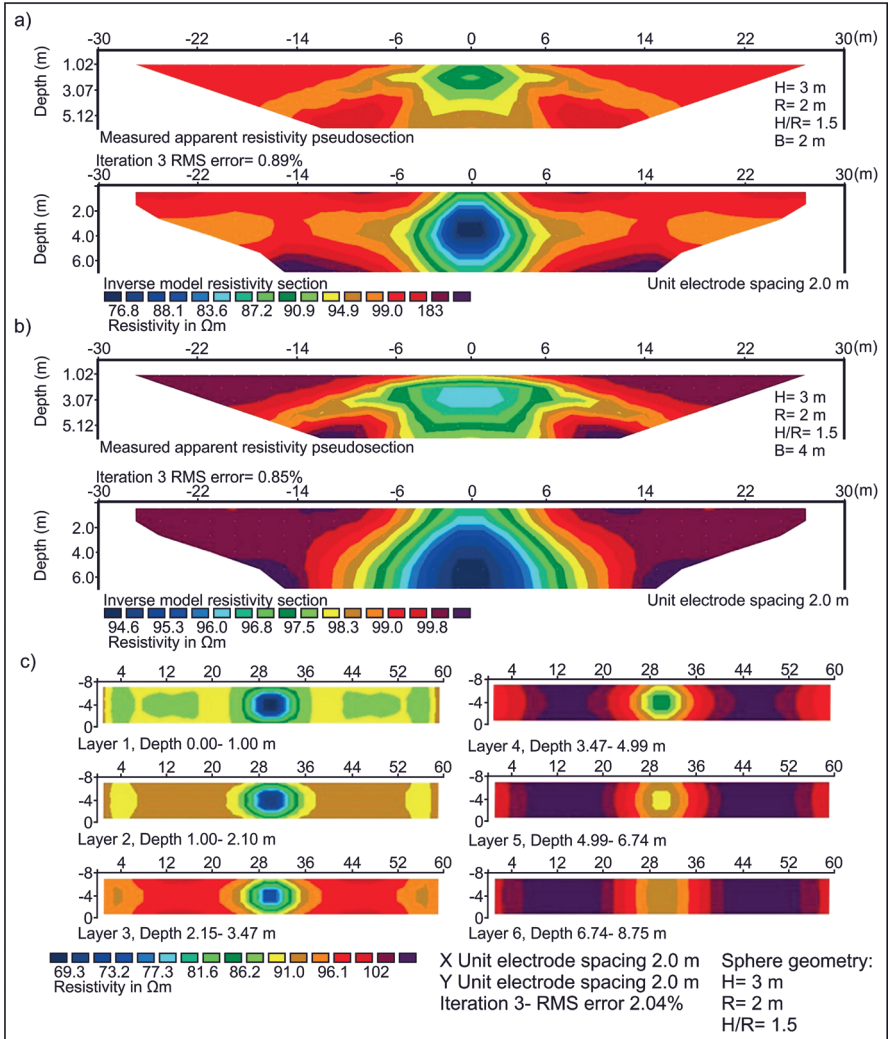


Figure 12. 2D survey over a spherical inhomogeneity. 2D profile with a shift of 2 m (a) and 4 m (b) from the projection of the sphere's center to the surface; (c) Conversion of 2D profiles to 3D format.

Figure 11b shows a 2D survey (ERT) over a spherical 3D near-surface inhomogeneity with a depth of $H = 3$ m and a radius of $R = 2$ m. The profile line passes through the center of the sphere's surface projection ($B = 0$ m). The existence of lateral effects that are symmetrically spread from the anomaly and distributed in all directions can be seen in the apparent resistivity section (Fig. 11b). The 2D inverse model shows an accurate depth to the center of the sphere as well as its radius, but not an accurate approximation of the true resistivity (60 Ωm instead of 10 Ωm). That effect is understandable provided that the anomaly from the sphere decreases rapidly as the ratio H/R (in this case $H/R = 1.5$).

Figures 12a and 12b show a two-dimensional ERT survey with a lateral shift from the sphere's projection to the surface ($B = 2$ m and $B = 4$ m). The anomalies caused by the existence of a 3D inhomogeneity (in this case a sphere) easily become undetectable.

The results of converting a 2D profile set to pseudo or semi 3D are shown in Fig. 12c. The position of the sphere is detectable (x , y , and z coordinates) in the first three layers (depth from 0 to 3.5 m), but the true resistivity is overestimated (70 Ωm instead of 10 Ωm). The last three layers (depths of 3.5 m to 8.75 m) steadily decrease true resistivity intensity and the sphere shape, which corresponds to the decrease in anomaly intensity induced by the sphere.

4. Conclusion

Resistivity surveys using the Wenner tri-potential technique, Offset Wenner, and the square array were presented to identify, quantify, and reduce lateral effects, as well as to narrow the equivalence range for various surveys in Serbia.

Model equivalence analysis revealed that the Wenner β array provides the best results for the first and second layer resistivities, whereas the Wenner γ array provides the highest apparent resistivity values and the lowest *ABSE* value for the third, highly resistive layer.

By averaging the electrical resistance values (R_1^a and R_2^a), it is possible to obtain an apparent resistivity sounding curve that is relatively free of lateral effects caused by 2D and 3D near-surface inhomogeneities. By averaging the values of all measured electrical resistances with the Offset Wenner array lateral effects reduction of local near surface inhomogeneities is achieved. In addition, the method developed by Habberjam and Watkins (1967b) permits a quantitative approximation and reduction based on these averaged values. A field investigation in the Zemun loess plateau demonstrates this possibility.

The Wenner tri-potential technique was applied in the Vrdnik survey (Serbia) where the *LII* values of 110% and the *PC* values of 0.092 confirmed that the 1D inversion was inadequate. In contrast, the *PC* values were lower in Golubac (Serbia), where the square array survey was applied. Because square array is less sensitive to lateral effects than collinear arrays, the averaged sounding curve was of higher quality than the sounding curves obtained from collinear arrays. When the 1D interpretation is inadequate due to electrical anisotropy or the effects of deeper structures, as reflected by high *AII* values, the crossed square array must be used to conduct additional research.

The Offset Wenner array was utilized to validate the interpolated and extrapolated values. It was determined that the interpolated apparent resistivity values were reasonably accurate. Extrapolated values are inadequate when resistivity decreases with depth, *i.e.*, when a low-resistivity substratum exists.

Due to the length of the paper, the equivalence scenario was not fully explored; ongoing research focuses exclusively on the equivalence principle with a two-layer starting model. This will enable better characterization of the depth of detectability and the possible narrowing of the equivalence range with different arrays.

Acknowledgements – This work was supported by the “Contract on realization and financing of scientific research of SRI in 2022”, Nr. 451-03-68/2022-14/200126.

References

- Acworth, R. I. and Griffiths, D. H. (1985): Simple data processing of tripotential apparent resistivity measurements as an aid to the interpretation of subsurface structure, *Geophys. Prospect.*, **33**, 861–887, <https://doi.org/10.1111/j.1365-2478.1985.tb00782.x>.
- Barker, R. D. (1989): Comment on “Examination of sounding curve extrapolation used by the offset wenner system” by White, P. A. and Scott, D. M., *Geophys. Prospect.*, **37**, 107–110, <https://doi.org/10.1111/j.1365-2478.1989.tb01824.x>.
- Barker, R. (1981): The offset system of electrical resistivity sounding and its use with a multicore cable, *Geophys. Prospect.*, **29**, 128–143, <https://doi.org/10.1111/j.1365-2478.1981.tb01015.x>.
- Bhattacharya, B. B. and Shalivahan, S. (2016): *Geoelectric methods: Theory and application*. McGraw Hill Education (India).
- Bhattacharya, P. K. and Patra, H. P. (1968): *Direct current geoelectric sounding: Principles and interpretation*. Elsevier Publishing Company.
- Busby, J. P. and Peart, R. J. (1997): Azimuthal resistivity and seismic measurements for the determination of fracture orientations, *Eng. Geol. Spec. Pub.*, Geological Society, London, **12**, 273–281, <https://doi.org/10.1144/GSL.ENG.1997.012.01.24>.
- Carpenter, E. W. and Habberjam, G. M. (1956): A tri-potential method of resistivity prospecting, *Geophys.*, **21**, 455–469, <https://doi.org/10.1190/1.1438247>.
- Carpenter, E. W. (1955): Some notes concerning the Wenner configuration, *Geophys. Prospect.*, **3**, 388–402, <https://doi.org/10.1111/j.1365-2478.1955.tb01384.x>.
- Chermali, M., Bellalem, F., Belgroun, M. W., Boudella, A. and Bounif, M. O. (2018): Examination of the validity of the computed intermediate apparent resistivity points in the Wenner and offset Wenner configuration of sounding, *Acta Geod. Geophys.*, **53**, 661–677, <https://doi.org/10.1007/s40328-018-0236-x>.
- Čičulić-Trifunović, M. and Rakić, M. O. (1971): *Basic geologic map of Former Yugoslavia 1:100 000, Explanatory booklet for the Sheet of Novi Sad*. Federal Geological Survey, Belgrade (in Serbian).
- Edwards, L. S. (1977): A modified pseudosection for resistivity and IP, *Geophys.*, **42**, 1020–1036, <https://doi.org/10.1190/1.1440762>.
- Habberjam, G. M. (1975): Apparent resistivity, anisotropy and strike measurements, *Geophys. Prospect.*, **23**, 211–247, <https://doi.org/10.1111/j.1365-2478.1975.tb01525.x>.
- Habberjam, G. M. and Jackson, A. A. (1974): A resistivity section of Holy Island (Anglesey), *Geol. J.*, **9**, 167–174, <https://doi.org/10.1002/gj.3350090205>.
- Habberjam, G. M. (1972a): The effects of anisotropy on square array resistivity measurements, *Geophys. Prospect.*, **20**, 249–266, <https://doi.org/10.1111/j.1365-2478.1972.tb00631.x>.
- Habberjam, G. M. (1972b): The point of assignment of an earth-resistance measurement, *Geoexpl.*, **10**, 141–148, [https://doi.org/10.1016/0016-7142\(72\)90029-4](https://doi.org/10.1016/0016-7142(72)90029-4).
- Habberjam, G. M. (1969): The location of spherical cavities using a tripotential resistivity technique, *Geophys.* **34**, 780–784. <https://doi.org/10.1190/1.1440049>.
- Habberjam, G. M. and Watkins, G. E. (1967a): The use of a square configuration in resistivity prospecting, *Geophys. Prospect.*, **15**, 445–467, <https://doi.org/10.1111/j.1365-2478.1967.tb01798.x>.

- Haberjam, G. M. and Watkins, G. E. (1967b): The reduction of lateral effects in resistivity probing, *Geophys. Prospect.*, **15**, 221–235, <https://doi.org/10.1111/j.1365-2478.1967.tb01785.x>.
- Kalenić, M., Hadži-Vuković, M., Dolić, D., Lončarević, Č. and Rakić, M. O. (1973): *Basic geologic map of Former Yugoslavia 1:100 000, Explanatory booklet for the Sheet of Kučevo*. Federal Geological Survey, Belgrade (in Serbian).
- Koefoed, O. (1969): An analysis of equivalence in resistivity sounding, *Geophys. Prospect.*, **17**, 327–335, <https://doi.org/10.1111/j.1365-2478.1969.tb02087.x>.
- Loke, M. H. (2022): Tutorial: 2-D and 3-D electrical imaging surveys. Unpublished course notes (available at: <https://www.geotomosoft.com/downloads.php>).
- Loke, M. H. (2016): RES2DMOD ver. 3.03. Rapid 2D resistivity forward modelling using the finite difference and finite-element methods. Wenner (alpha, beta, gamma), inline & equatorial dipole-dipole, pole-pole, pole-dipole and Wenner-Schlumberger. Freeware courtesy of M. H. Loke (available from <https://www.geotomosoft.com/downloads.php>).
- Loke, M. H. (1994): The inversion of two-dimensional apparent resistivity data. Unpubl. Ph.D. thesis, University of Birmingham, U.K.
- Maillet, R. (1947): The fundamental equations of electrical prospecting, *Geophys.*, **12**, 529–556, <https://doi.org/10.1190/1.1437342>.
- Marković, B., Veselinović, M., Anđelković, J., Stevanović, P., Roglić, Č. and Obradinović, Z. (1984): *Basic geologic map of Former Yugoslavia 1:100 000, Explanatory booklet for the Sheet of Belgrade*. Federal Geological Survey, Belgrade (In Serbian).
- Nunn, K. R., Barker, R. D. and Bamford D. (1983): In situ seismic and electrical measurements of fracture anisotropy in the Lincolnshire Chalk, *Q. J. Eng. Geol. Hydrogeol.*, **16**, 187–198, <https://doi.org/10.1144/GSL.QJEG.1983.016.03.03>.
- Rajčević, D. (1978): *Basic Geologic Map of Former Yugoslavia 1:100 000, Explanatory booklet for the Sheet of Šabac* Federal Geological Survey, Belgrade (in Serbian).
- Sanuade, O. A., Amosun, J. O., Oyeyemi, K. D., Oloajo, A. A., Fagbemigun, T. S. and Faloyo, J. I. (2019): Analysis of principles of equivalence and suppression in resistivity sounding technique, *J. Phys: Conf. Ser.*, 1299, <https://doi.org/10.1088/1742-6596/1299/1/012065>.
- Schober, P., Boer, C. and Schwarte, L. A. (2018): Correlation coefficients: appropriate use and interpretation, *Anesth. Analg.*, **126**, 1763–1768, <https://doi.org/10.1213/ANE.0000000000002864>.
- Sretenović, B. and Arnaut, F. (2019): Inadequacy of 1D, 2D and 3D resistivity inverse modeling in the presence of electrical anisotropy, *J. Earth. Sci.*, **8**, 102–116, <https://doi.org/10.11648/j.earth.20190802.14>.
- Sretenović, B. B., Vasiljević, I. M. and Cvetkov, V. V. (2014): The application of geoelectrical attributes based on crossed square array resistivity sounding data at 'Pudarske Kolibe' site in eastern Serbia, *Teh.*, **69**, 604–609, <https://doi.org/10.5937/tehnika1404604S>.
- Szalai, S., Lemperger, I., Metwaly, M., Kis, A., Wesztergom, V., Szokoli, K. and Novák, A. (2014): Multiplication of the depth of detectability using γ_{11n} arrays, *Appl. Geophys.*, **107**, 195–206, <https://doi.org/10.1016/j.jappgeo.2014.06.003>.
- Szalai, S. and Szarka, L. (2008): On the classification of surface geoelectric arrays, *Geophys. Prospect.*, **56**, 159–175, <https://doi.org/10.1111/j.1365-2478.2007.00673.x>.
- Udosen, N. I. and George, N. J. (2018): Characterization of electrical anisotropy in North Yorkshire, England using square arrays and electrical resistivity tomography, *Geomech. Geophys. Geo-energ. Geo-resour.*, **4**, 215–233, <https://doi.org/10.1007/s40948-018-0087-5>.
- Vrhovčić, J., Buzaljko, R., Mojićević, M., Prtoljan, B., Galovilć, I., Marković, B., Anđelković, J., Pavlović, Z. and Rajčević, D. (1986): *Basic Geologic Map of Former Yugoslavia 1:100 000, Explanatory booklet for the Sheet of Bijeljina*, Federal Geological Survey, Belgrade (in Serbian).
- White, P. A. and Scott, D. M. (1988): Examination of sounding curve extrapolation used by the offset Wenner system, *Geophys. Prospect.*, **36**, 194–200, <https://doi.org/10.1111/j.1365-2478.1988.tb02159.x>.

- Zhou, B. (2016): Electrical resistivity tomography: A subsurface-imaging technique, in: *Applied geophysics with case studies on environmental, exploration and engineering geophysics*, edited by Kanli, A. I. IntechOpen, London, <https://doi.org/10.5772/intechopen.81511>.
- Zohdy, A. A. (1949): Use of Dar Zarrouk curves in the interpretation of vertical electrical sounding data, *Geol. Surv. Bullet. 1313-D*, US Department of the Interior, Geological Survey, Washington, DC, <https://doi.org/10.3133/b1313D>.

SAŽETAK

Poboljšanje 1D geoelektričnog sondiranja sužavanjem raspona ekvivalencije i indentifikacija, kvantifikacija i redukcija lateralnih efekata korišćenjem tri-potencijalne metodologije

Filip Arnaut, Branislav Sretenović i Vesna Cvetkov

Podaci o prividnoj specifičnoj električnoj otpornosti prikupljeni Offset-Wenner dispozitivom (Zemun, Srbija), kvadratnim dispozitivom (Bogatić i Golubac, Srbija) i Wenner tri-potencijalnom metodologijom (Vrdnik, Srbija) korišteni su za detekciju, mjerenje i smanjenje bočnih nuspojava tijekom 1D inverzije. Direktno i inverzno modeliranjem s Wenner α , β i γ dispozitivima pokazalo je da Wenner β dispozitiv daje najbolju procjenu specifične električne otpornosti prvog i drugog sloja, dok je Wenner γ dispozitiv dao najbolju procjenu trećeg sloja, visokootpornog substratuma. Istraživanja na Zemunskom lesnom platou pokazala su da kada je indeks bočne nehomogenosti (*LII*) nizak, opravdana je 1D interpretacija oba Wenner dispozitiva. Također, usrednjavanje otpora će proizvesti krivulju otpora koja je oslobođena bočnih učinaka od blisko površinskih nehomogenosti. Da je 1D inverzija neadekvatna kada su *LII* i vrijednosti kovarijance procesinga (*PC*) velike prikazano je na primjeru Vrdnika. Primenom kvadratnog dispozitiva na lokalitetu Golubac dobijene su niže vrednosti *PC* od kolinearnih dispozitiva. Stoga je kvalitet osrednjene krivulje sondiranja bila veća. Interpolirane vrijednosti Offset-Wenner dispozitiva imale su dovoljnu preciznost, dok su ekstrapolirane vrijednosti bile neadekvatne u prisutnosti niskootpornog substratuma.

Ključne riječi: blisko-površinska ispitivanja, 2D geoelektrična tomografija, inverzija, Offset Wenner dispozitiv, kvadratni dispozitiv, efekti blisko-površinskih nehomogenosti na 1D ispitivanja

Corresponding author's address: Filip Arnaut, Department of Geophysics, Faculty of Mining and Geology, University of Belgrade, Đušina 7, 11120 Belgrade, Serbia; tel: +381 63 83 89 237; e-mail: filip.arnaut@rgf.rs



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Machine Learning Classification Workflow and Datasets for Ionospheric VLF Data Exclusion

Filip Arnaut ^{*}, Aleksandra Kolarski  and Vladimir A. Srećković 

Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; aleksandra.kolarski@ipb.ac.rs (A.K.); vlada@ipb.ac.rs (V.A.S.)

* Correspondence: filip.arnaut@ipb.ac.rs

Abstract: Machine learning (ML) methods are commonly applied in the fields of extraterrestrial physics, space science, and plasma physics. In a prior publication, an ML classification technique, the Random Forest (RF) algorithm, was utilized to automatically identify and categorize erroneous signals, including instrument errors, noisy signals, outlier data points, and the impact of solar flares (SFs) on the ionosphere. This data communication includes the pre-processed dataset used in the aforementioned research, along with a workflow that utilizes the PyCaret library and a post-processing workflow. The code and data serve educational purposes in the interdisciplinary field of ML and ionospheric physics science, as well as being useful to other researchers for diverse objectives.

Dataset: <https://zenodo.org/records/10433194> (accessed on 26 December 2023)

Dataset License: CC-BY 4.0

Keywords: machine learning; open data; ionospheric anomaly classification; open-source software



Citation: Arnaut, F.; Kolarski, A.; Srećković, V.A. Machine Learning Classification Workflow and Datasets for Ionospheric VLF Data Exclusion. *Data* **2024**, *9*, 17. <https://doi.org/10.3390/data9010017>

Academic Editor: Giuseppe Ciaburro

Received: 26 December 2023

Revised: 15 January 2024

Accepted: 17 January 2024

Published: 18 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

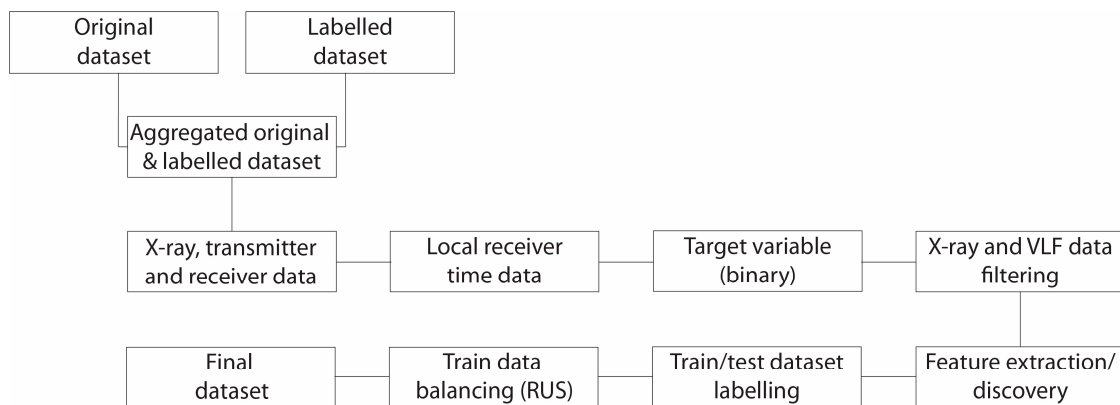
1. Summary

Numerous machine learning (ML) algorithms and pre-processing techniques have been made possible by rapid advancements in computer science, data science, and data analysis. It can be noted that it takes a lot of time and effort to manually verify, review, and exclude data from an ionospheric very-low-frequency (VLF) investigation during intense occurrences [1,2]. Nevertheless, ML classification methods can be used to automate this job. We evaluated the Random Forest (RF) algorithm [3] in our prior publication [4] with the purpose of automatically classifying erroneous ionospheric VLF amplitude data points during solar flare (SF) investigation/detection. These erroneous data points were categorized as representing SF events, instrumentation errors, or noisy signals. Due to its ease of use and simplicity (few hyperparameters to tune and a reduced likelihood of overfitting the model due to averaging/voting [5] and the law of large numbers [3]), the RF algorithm is considered a first choice for various ML tasks [6,7]. Consequently, it was a suitable selection for the given research purpose. However, as stated in the research paper [4], it is advantageous to extend the original dataset and test additional classification algorithms in order to possibly increase the predictive power of the algorithms.

This data report fulfills two objectives: firstly, it will catalog and provide a link to the data employed in this study, thereby making them accessible to a broader range of researchers, professionals, and others, and secondly, it will include a workflow that integrates the PyCaret library [8], enabling the comparison and testing of fifteen models in total (data and code available at Supplementary material). The chapter methods, i.e., the workflow description, will provide a comprehensive overview of the workflow utilized in conjunction with the data. A synopsis of the pre-processing steps performed during the construction of the original dataset is also available in [4].

2. Data Description

The datasets that were originally obtained and labeled (erroneous values were filtered out) for other research purposes in September and October of 2011 provided a favorable opportunity to evaluate ML classification on this type of data. The original and labeled samples were combined into a single data source, to which additional data (X-ray irradiance, transmitter and receiver data, and local receiver time) were added (Figure 1). The target variable was obtained from the labeled, i.e., filtered, dataset where each instance which was filtered out of the original dataset was annotated as 1 (anomalous data class) and the data that remained were annotated as 0 (normal data class). The feature extraction process was performed by analyzing statistical features of the VLF amplitude and X-ray irradiance signals. The statistical features utilized included rolling window statistics such as mean, standard deviation, and median, with three different window sizes (5, 20, and 180 min). Additionally, lagged signals and other measures such as rate of change, first- and second-order difference, etc., were also employed. Due to the imbalanced nature of the given ML task, random undersampling [9–11] was performed to balance the distribution of the target labels. As a result, the final dataset was generated. To obtain a more comprehensive explanation of the data pre-processing, refer to [4].



RUS- Random Undersampling

Figure 1. Data pre-processing workflow (modified after [4]).

The dataset was divided into two separate sets: the training dataset and the test dataset. Both of these sets have been pre-processed and are provided as links. The workflow described below can be readily applied to these pre-processed versions of the dataset.

3. Methods (Workflow Description)

Prior to executing the code, the user must specify the input variables, which include the training and test datasets, as well as the visualization range. The initial stage of the workflow involves ML modeling, where the PyCaret library employs the training dataset to conduct a comparison among 15 ML algorithms (Figure 2 and Table 1). After conducting the comparison, the model with the best evaluation metrics and statistics is selected as the overall best model. This model is then used for the hyperparameter tuning process to further optimize the model. The last step involves employing the optimized model to make predictions on the given test dataset. This process generates an output file that includes the predictions made by the most effective and fine-tuned ML algorithm, along with the features and target variables. The post-processing workflow can be summarized in four steps: decoding the data from each transmitter and receiver, separating the individual transmitter–receiver pairs due to the presence of 19 pairs in the test dataset, computing per-class evaluation metrics for each pair, and finally visualizing the true and predicted data labels using the specified input range at the start of the workflow.

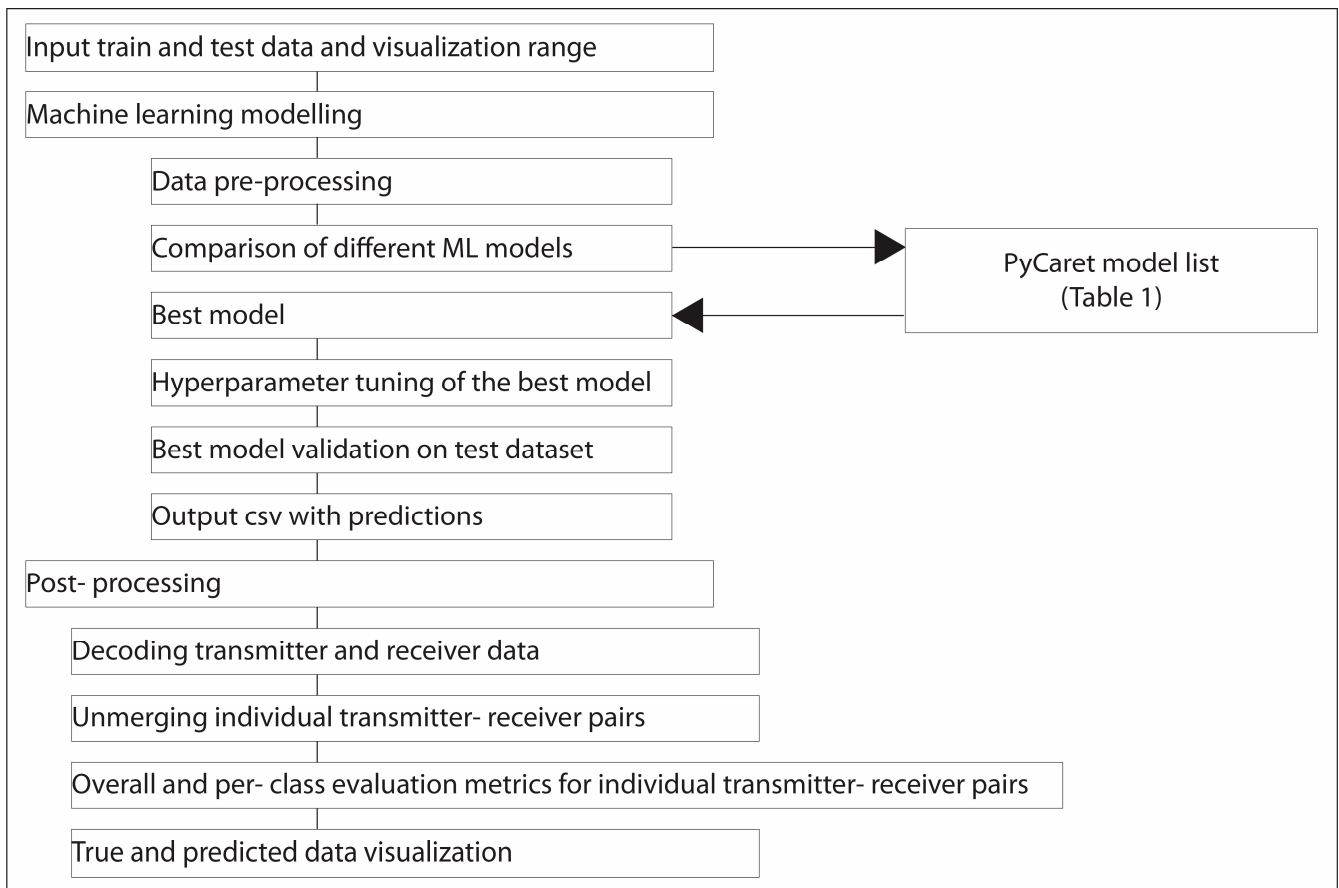


Figure 2. Workflow for ML modeling and post-processing.

The evaluation metrics employed for the workflow consist of the confusion matrix, as well as the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values for each class. Furthermore, the workflow presents accuracy, precision, and F1-score values for each class as well. The overall metrics are also displayed as an output of the workflow. However, due to the highly imbalanced ML task in question, greater emphasis should be given to per-class evaluation metrics. Furthermore, a comprehensive overview of all evaluation metrics can be found in [4]. However, for the purpose of this brief data descriptor, we will provide a short definition of the F1-score. The F1-score is calculated as the harmonic mean of the true positive (TP) rate, also known as recall, and the precision parameter. When evaluating imbalanced ML tasks, the F1-score is typically preferred over accuracy [12,13].

The ML workflow's results are displayed in Figure 3b. All three panels represent the signal obtained from the NAA-Walsenburg transmitter–receiver pair. The signal's duration in Figure 3 spans 600 min, beginning on 19 October 2011 at 14:37 UT and concluding on 20 October 2011 at 1:37 UT. In addition, the workflow also provides the evaluation metrics for each transmitter–receiver pair, specifically the F1-score. In the given example shown in Figure 3b, the anomalous data class has an F1-score of 0.65, while the normal data class has a score of 0.96. This results in a total F1-score of 0.93.

A comparison of the outcomes produced by the workflow integrating the PyCaret library and the transmitter–receiver pair utilized by [4] employing the RF algorithm reveals a distinction. The PyCaret algorithm ascertains that the Extra Trees Classifier (ET) is the optimal overall model for the given task. The comparison of the outputs clearly illustrates that, at least for the instance depicted in Figure 3, the ET classifier is more suitable. However, additional investigation is required to ascertain the specific conditions that require a certain model.

Table 1. Models in the PyCaret library. For more details please see [3,14–23].

#	Model Name	Python Function	More Information at:
1	Logistic Regression	lr	[14]
2	Ridge Classifier	ridge	[14]
3	Linear Discriminant Analysis	lda	[15]
4	Random Forest Classifier	rf	[3]
5	Naive Bayes	nb	[14]
6	Gradient boosting Classifier	gba	[16]
7	Adaboost Classifier	ada	[17]
8	Extra Trees Classifier	et	[18]
9	Quadratic Discriminant Analysis	qda	[19]
10	Light Gradient Boosting Machine	lightgbm	[20]
11	K Neighbors Classifier	knn	[21]
12	Decision Tree Classifier	dt	[22]
13	Extreme Gradient Boosting	xgboost	[23]
14	Dummy Classifier	dummy	[15]
15	SVM Linear Kernel	svm	[14]

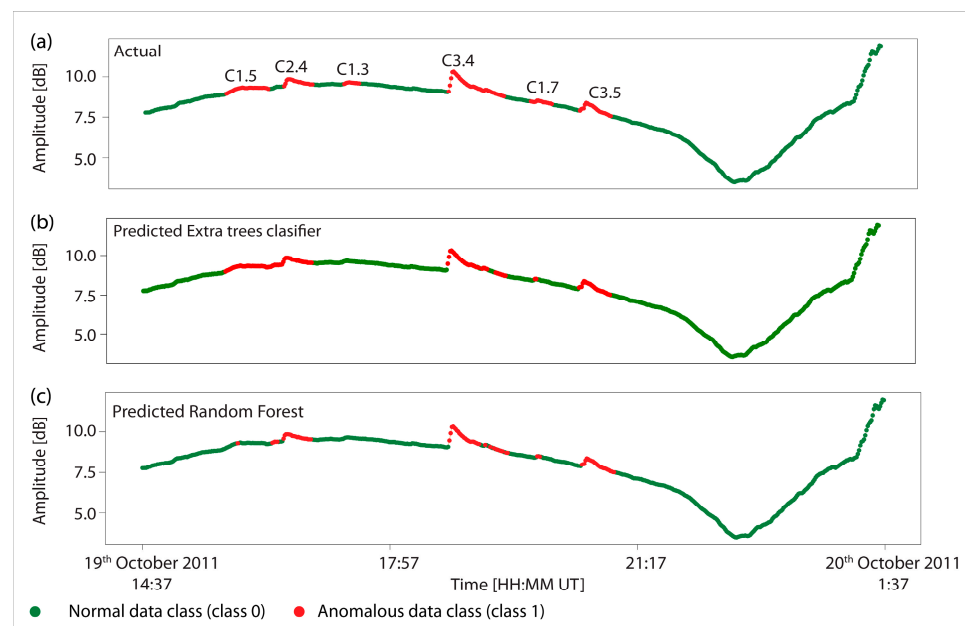


Figure 3. (a) Visualization of actual class labels for the NAA_Walsenburg transmitter–receiver pair from 19 October 2011 14:37 to 20 October 2011 01:37, obtained from [4]; (b) predictions made by the Extra Trees Classifier from the PyCaret library for the same time period; (c) predictions made by the Random Forest Classifier from the same time period, obtained from [4].

In addition to conducting additional research to identify the most suitable model for different scenarios, it is crucial to undertake a comprehensive data acquisition endeavor to further enhance the predictive capabilities of a model. Additional data collection would allow the model to acquire observations from a wider array of events and varying degrees of potential noise levels. For example, data from different time periods within one or a couple of solar cycles can be utilized, etc. Acquiring this level of detailed data requires the collaboration of a larger team of researchers to label and verify the data in a semi-

manual manner. Following this undertaking, the model has the potential to be significantly enhanced and additional solutions can be devised to cater to a larger research community, for instance, the creation of standalone software with a user-friendly interface that can be utilized by a diverse group of researchers for data VLF pre-processing.

4. User Notes

Researchers may benefit from this database and method/analysis. The code and data serve educational purposes in the interdisciplinary field of ML and ionospheric physics science, as well as being useful to other researchers for diverse objectives. The benefits of these data are the fact that manual data labeling is a laborious undertaking that requires the time of a few researchers. This dataset fulfills the function of being a publicly accessible, annotated dataset that can be employed by researchers to experiment with various research perspectives, thereby conserving both time and labor. The study of space weather (which includes VLF ionospheric research) is highly significant for the broader community (for more information see [24]), as it offers a valuable understanding of the sun–earth connection [25]. This research can enhance our comprehension of space weather phenomena, which in turn will have significant implications for navigation systems, telecommunications, and other related fields.

Supplementary Materials: The following supporting information can be downloaded at <https://zenodo.org/records/10433194> (accessed on 26 December 2023).

Author Contributions: Conceptualization, F.A.; formal analysis, F.A., A.K. and V.A.S.; visualization, F.A.; writing—original draft, F.A.; writing—review and editing, V.A.S., A.K. and F.A. All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Institute of Physics Belgrade through a grant by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia. This article/publication is based upon work from COST Action CA22162—A TRANSDISCIPLINARY NETWORK TO BRIDGE CLIMATE SCIENCE AND IMPACTS ON SOCIETY (FutureMed)—supported by COST (European Cooperation in Science and Technology).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data and the workflow are available at <https://zenodo.org/records/10433194>.

Acknowledgments: The WALDO database (<https://waldo.world>, accessed on 1 October 2023) provides VLF data. It is run collaboratively by the University of Colorado Denver and the Georgia Institute of Technology, utilizing data gathered from Stanford University and these two institutions. It has been made possible by numerous grants from the Department of Defense, NASA, and the NSF.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. McRae, W.M.; Thomson, N.R. VLF Phase and Amplitude: Daytime Ionospheric Parameters. *J. Atmos. Sol.-Terr. Phys.* **2000**, *62*, 609–618. [[CrossRef](#)]
2. Šulić, D.M.; Srećković, V.A.; Mihajlov, A.A. A Study of VLF Signals Variations Associated with the Changes of Ionization Level in the D-Region in Consequence of Solar Conditions. *Adv. Space Res.* **2016**, *57*, 1029–1043. [[CrossRef](#)]
3. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
4. Arnaut, F.; Kolarski, A.; Srećković, V.A. Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. *Universe* **2023**, *9*, 436. [[CrossRef](#)]
5. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random Forests for Classification in Ecology. *Ecology* **2007**, *88*, 2783–2792. [[CrossRef](#)] [[PubMed](#)]
6. Hatwell, J.; Gaber, M.M.; Azad, R.M.A. CHIRPS: Explaining Random Forest Classification. *Artif. Intell. Rev.* **2020**, *53*, 5747–5788. [[CrossRef](#)]

7. Bartz-Beielstein, T.; Chandrasekaran, S.; Rehbach, F.; Zaefferer, M. Case Study I: Tuning Random Forest (Ranger). In *Hyperparameter Tuning for Machine and Deep Learning with R*; Springer Nature: Berlin/Heidelberg, Germany, 2023; pp. 187–220. [[CrossRef](#)]
8. Ali, M. PyCaret: An Open Source, Low-Code Machine Learning Library in Python. PyCaret Version 1.0.0. 2020. Available online: <https://www.pycaret.org> (accessed on 1 October 2023).
9. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *ACM SIGKDD Explor. Newsl.* **2004**, *6*, 20–29. [[CrossRef](#)]
10. Hasanin, T.; Khoshgoftaar, T. The Effects of Random Undersampling with Simulated Class Imbalance for Big Data. In Proceedings of the 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, 6–9 July 2018. [[CrossRef](#)]
11. Saripuddin, M.; Suliman, A.; Syarmila Sameon, S.; Jorgensen, B.N. Random Undersampling on Imbalance Time Series Data for Anomaly Detection. In Proceedings of the 4th International Conference on Machine Learning and Machine Intelligence, Hangzhou, China, 17–19 September 2021. [[CrossRef](#)]
12. Hossin, M.; Sulaimani, M.N. A Review on Evaluation Metrics for Data Classification Evaluations. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1–11. [[CrossRef](#)]
13. Joshi, M.V. On Evaluating Performance of Classifiers for Rare Classes. In Proceedings of the 2002 IEEE International Conference on Data Mining, Maebashi City, Japan, 9–12 December 2002. [[CrossRef](#)]
14. Bonaccorso, G. *Machine Learning Algorithms*; Packt Publishing Ltd.: Birmingham, UK, 2017; pp. 1–368.
15. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
16. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
17. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [[CrossRef](#)]
18. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely Randomized Trees. *Mach. Learn.* **2006**, *63*, 3–42. [[CrossRef](#)]
19. Friedman, J.H. Regularized Discriminant Analysis. *J. Am. Stat. Assoc.* **1989**, *84*, 165–175. [[CrossRef](#)]
20. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems*; The MIT Press: Cambridge, MA, USA, 2017; Volume 30.
21. Cover, T.; Hart, P. Nearest Neighbor Pattern Classification. *IEEE Trans. Inf. Theory.* **1967**, *13*, 21–27. [[CrossRef](#)]
22. Quinlan, J.R. Induction of Decision Trees. *Mach. Learn.* **1986**, *1*, 81–106. [[CrossRef](#)]
23. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. Available online: <https://arxiv.org/abs/1603.02754v3> (accessed on 14 October 2023).
24. Hapgood, M. Societal and Economic Importance of Space Weather. In *Machine Learning Techniques for Space Weather*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 3–26. [[CrossRef](#)]
25. Kolarski, A.; Srećković, V.A.; Arnaut, F. Low Intensity Solar Flares’ Impact: Numerical Modeling. *Contrib. Astron. Obs. Skaln. Pleso.* **2023**, *53*, 176–187. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Short-term forecasting of PM₁₀ and PM_{2.5} concentrations with Facebook's Prophet Model at the Belgrade-Zeleno brdo

Filip Arnaut¹ , Vesna Cvetkov¹ , Dragana Đurić¹ 
and Mileva Samardžić-Petrović² 

¹ Department of Geophysics, Faculty of Mining and Geology, University of Belgrade, Belgrade, Serbia

² Department of Geodesy and Geoinformatics, Faculty of Civil Engineering, University of Belgrade, Belgrade, Serbia

Received 2 December 2022, in final form 25 October 2023

We demonstrate the use of Facebook's Prophet (usually just called Prophet) model for short-term air quality forecasting at Belgrade-Zeleno brdo monitoring station. To address missing data, we applied minimally-altering data distribution imputation techniques. Linear interpolation proved effective for short-term gaps (1–3 hours), hourly mean method for mid-term gaps (24–26 hours), and Hermite interpolation polynomial for long-term gaps (132–148 hours). The most significant data change was a 3.4% shift in skewness. Partitioning the time series enabled a detailed quality assessment of the Prophet model, with PM_{2.5} predictions being more precise than PM₁₀. Using the longest time series for forecasting yielded absolute errors of 6.5 µg/m³ for PM₁₀ and 2.7 µg/m³ for PM_{2.5}. Based on 173 forecasts, we anticipate Prophet model root-mean-square values under 6.26 µg/m³ and 9.99 µg/m³ for PM_{2.5} and PM₁₀ in 50% of cases. The Prophet model demonstrates several advantages and yields satisfactory results. In future research, the results obtained from the Prophet model will serve as benchmark values for other models. Additionally, the Prophet model is capable of providing satisfactory air quality forecasting results and will be utilized in future research.

Keywords: time series analysis, time series forecasting, air quality, suspended particles

1. Introduction

The Republic of Serbia (RS) has seen an increase in the demand for electric energy, in alignment with the global trend. This trend is particularly pronounced in the densely populated and urbanized region of the capital city of Belgrade. In the RS, lignite thermal power plants are the primary source of electricity (Energy

Agency of the Republic of Serbia, 2020), resulting in the production of large quantities of suspended particles, especially PM₁₀ and PM_{2.5}. Most of the electricity produced in the RS is derived from fossil fuels such as coal and natural gas, with renewable sources accounting for 20% of the total (Official Gazette of the Republic of Serbia, 156/2020). Notably, five of the ten most polluting thermal power plants are in the RS, while the average Western Balkan thermal power plant emits 16 times more suspended particles than the average European Union thermal power plant (Puljić et al., 2019).

The ecological problem posed by suspended particles in the RS, resulting from an insufficiently regulated energy sector, is further exacerbated in urban environments by traffic and heating materials. A substantial body of research has established a strong link between increased suspended particle concentrations and cardiovascular diseases, as well as an increased susceptibility to allergens (Jovanović, 2020; Dockery et al., 1992; Araujo, 2011; Bernstein et al., 2011). According to calculations by the World Health Organization for the year 2019, based on data from 2012, approximately 5400 fatalities and 126,000 lost years of life can be attributed to the rise in airborne particles (World Health Organization, 2019).

The knowledge of the spatial distribution of suspended particles in the future can be a crucial piece of information for decision-makers. Since 1927 (Bisgaard and Kulachi, 2011), modern methods of time series forecasting have been developed, and today there are, numerous models and methods for the forecasting of time series data, ranging from the simplest to the most complex, such as machine learning approaches.

The Prophet model was selected for this study due to its numerous successful applications in forecasting air pollution parameters in large urban areas worldwide. Additionally, other studies have shown that the Prophet model exhibits notable capabilities in predicting various time-series data, such as district heating load (Shakeel et al., 2023a, b) and the percentage of COVID-19 infected population (Papastefanopoulos et al., 2020). Notably, the Prophet model has exhibited superior performance compared to traditional machine learning models, particularly in scenarios with limited data availability. For instance, research conducted in Seoul (South Korea) demonstrated short-term and long-term forecasts of PM₁₀, PM_{2.5}, O₃, NO₂, SO₂, and CO concentrations using the Prophet model (Shen et al., 2020). The mean absolute error values obtained for PM_{2.5} and PM₁₀ concentrations were 12.6 µg/m³, and 19.6 µg/m³, respectively. The data were two to four times more accurate than similar models' forecasts (Shen et al., 2020). In Zhao et al. (2018), the Prophet model was instrumental in revealing data trends that might otherwise be challenging to discern. The study found that the highest annual PM_{2.5} levels occur in January, with the highest weekly levels on Fridays, and the lowest concentrations on Sundays and Mondays (Zhao et al., 2018). Hybrid forecasting methods were also shown to yield more accurate results compared to simple methods when forecasting PM_{2.5} concentrations (Ejohwomu

et al., 2022). The forecasts were quantified using the mean absolute scaled error (Hyndman and Koehler, 2006), a reliable method for assessing forecast model quality when comparing multiple models. Samal et al. (2019) demonstrated that the Prophet model without transformations provided more accurate air quality forecasts in Bhubaneswar, India, compared to the Seasonal ARIMA (SARIMA) model. The addition of log transformation further improved the Prophet model's accuracy (Samal et al., 2019). Ye (2019) introduced a hybrid, weight-based model for air quality forecasting in Shenzhen, China, combining the ARIMA and Prophet models for a three-day forecasting horizon. The limitations of the ARIMA model include the time-consuming search for optimal (p , d , q) parameters. The forecasting method, in general, faces challenges when predicting over long horizons, with the potential for significant model errors (Ye, 2019). The automation of the lengthy search for optimal ARIMA parameters can be achieved through the utilization of the `pmdarima` library in Python (Smith and Taylor, 2017). In Nanjing, China, Zhou et al. (2020) developed a hybrid Prophet-LSTM (Long Short-Term Memory) model that improved the Prophet model's accuracy, particularly for high-value predictions with a one-year forecasting horizon. Tejasvini et al. (2020) found that when comparing the Prophet model to ARIMA and Naïve Bayes, the Prophet model required the least time for making predictions and had the lowest calculated error parameter when forecasting atmospheric CO.

The primary objective of this study is to apply the Prophet model to concentration data of PM_{10} and $PM_{2.5}$ suspended particles at the Belgrade-Zeleno brdo monitoring station. In this study, a more conservative approach was taken to forecasting, *i.e.*, the forecasting horizon wasn't set too far in the future but was limited to 24 hours. This approach allows us to assess the Prophet model's performance in providing satisfactory one-day-ahead forecasts—a best-case scenario. As demonstrated by Shen et al. (2020), extending the forecasting horizon can compound error terms, making a shorter horizon a more reliable measure of the model's effectiveness.

2. Methodology and data

Time series forecasting encompasses a range of techniques for predicting future observations based on past (historical) data. Typically, time series forecasting techniques rely on equally spaced data (Kovačić, 1995; Lütkepohl and Kratzig, 2004; Agrawal and Adhikary, 2013). While the Prophet model does not necessitate equally spaced data, this research addresses missing observations to facilitate future correlation with other models that do require equally spaced data.

The data preprocessing scheme requires information on three key parameters:

- the number of data points; the quantity of data points contained in the data file,
- the total number of data points for the investigated period; the total number of data points for the investigated period minus the number of data

points that include missing observations, or the number of data points that can be found in the time interval if no observations are missed,

- the total number of missing observations; the number of observations that were either not measured (due to instrument failure) or were eliminated from the data set (erroneous values which can also be due to instrumental failure).

The choice between data imputation, data amputation, or a combination of these methods depends on the number of missing observations and their location in the data set.

Methods for imputation of missing data can be straightforward, such as last observation carried forward, previous measurement carried backward, average method (Kang, 2013), K nearest neighbors, linear (or any other) interpolation, etc. It is essential to note that there is no “one-size-fits-all” approach to data preprocessing that is appropriate in every circumstance. Usually, imputation methods are selected by trial and error. Suitable imputation methods for this study were those that do not alter the basic statistical properties of the data *i.e.*, the data distribution after preprocessing (Fig. 1). Additionally, data amputation methods involve eliminating observations from the data set.

Both before and after data preprocessing, descriptive statistics were used to calculate various statistics, including the mean, median, mode, minimum, maximum, skewness, kurtosis, and Pearson’s and Spearman’s correlation coefficients between variables (Schober et al., 2018). Furthermore, a visual inspection is always valuable for checking data and its distribution (Bisgaard, Kulachi, 2011). As part of the descriptive statistics step, a quantitative approach to testing data distribution modality, based on Hartigan and Hartigan’s unimodality test (Hartigan, Hartigan, 1985), was also implemented (refer to Fig. 1).

For the quantitative evaluation of the chosen model, the in-sample forecasting technique was utilized for statistical validation. In-sample forecasting was performed by removing a specific amount of data (in this case, the last 24 hours), and statistical validation was based on the discrepancy between the forecast and the removed in-sample data.

The quantitative assessment parameters for the model were the absolute error (*AE*), mean absolute error (*MAE*), absolute percentage error (*APE*), mean absolute percentage error (*MAPE*), and root mean square error (*RMSE*) are.

Absolute error (*AE*) and absolute percentage error (*APE*) are given as:

$$AE = |y_{forecast} - y_{true}| \text{ [}\mu\text{g/m}^3\text{]} \quad (1)$$

$$APE = \frac{|y_{forecast} - y_{true}|}{y_{true}} \cdot 100 \text{ [%]} \quad (2)$$

where $y_{forecast}$ is the value predicted by the Prophet model and y_{true} is the true value used for validation.

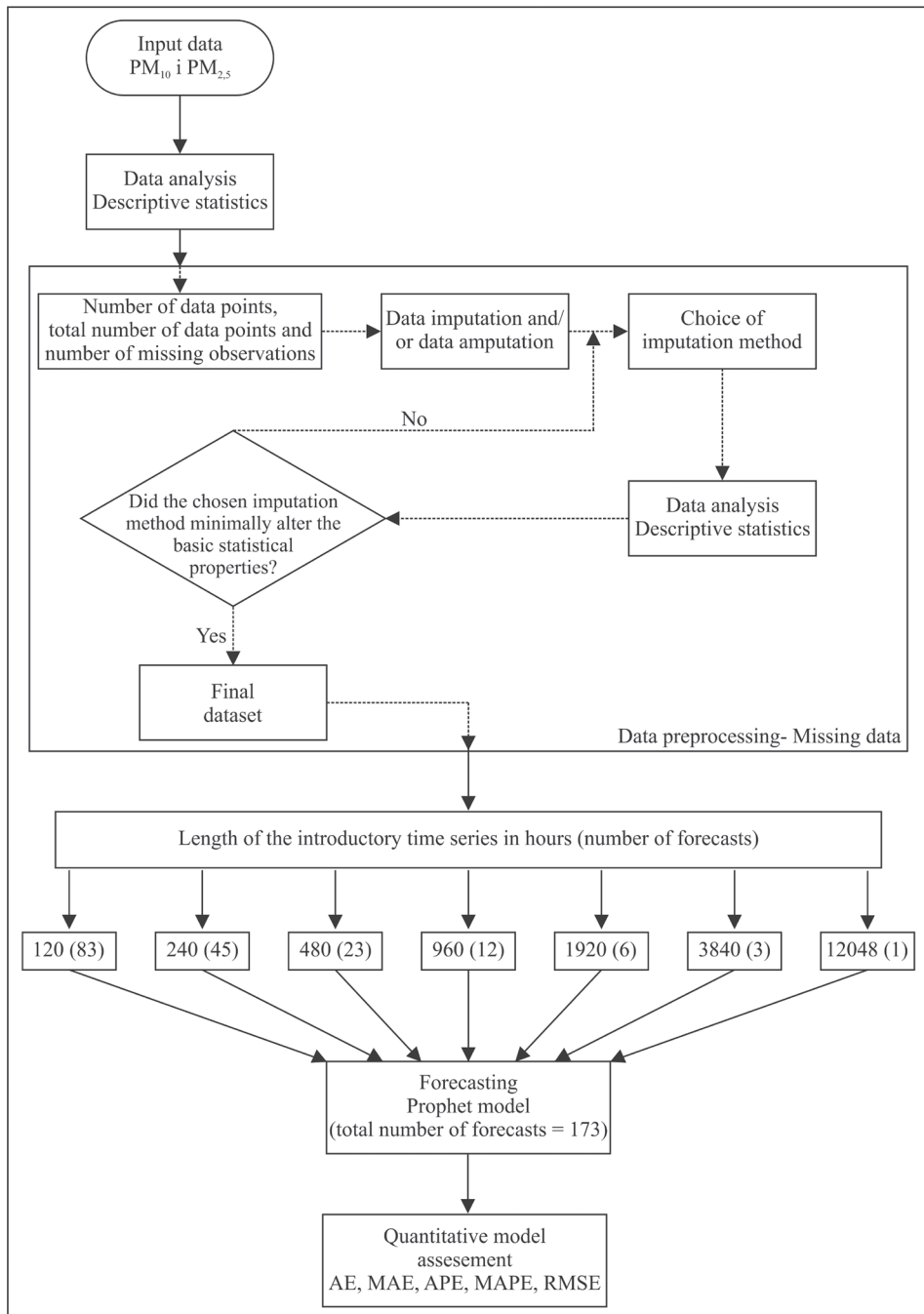


Figure 1. Research methodology.

If the length of the forecasting horizon is greater than one, two additional parameters can be obtained: the mean absolute error (*MAE*) and the mean absolute percentage error (*MAPE*), which are given as:

$$MAE = \frac{\sum_{i=1}^N |y_{forecast} - y_{true}|}{N} \left[\mu\text{g}/\text{m}^3 \right] \quad (3)$$

$$MAPE = \frac{\sum_{i=1}^N \frac{|y_{forecast} - y_{true}|}{y_{true}}}{N} \cdot 100 [\%] \quad (4)$$

where N is the length of the forecasting horizon.

The root mean square error (*RMSE*) is defined by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_{true} - y_{forecast})^2}{N}} \left[\mu\text{g} / \text{m}^3 \right] \quad (5)$$

The regression-based Prophet model (Taylor and Letham, 2018) used in this study is represented by the following equation:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (6)$$

where $g(t)$ is the trend, $s(t)$ is the periodic, *i.e.*, seasonal variations, $h(t)$ is the non-equally spaced holiday effect, and ε_t are changes to which the model does not adapt to (Taylor and Letham, 2018). The Prophet model's advantages can be seen in its simplicity and its speed in processing large amounts of data.

Furthermore, it is important to note that the Prophet model does not fall under the category of machine learning models. Instead, it is more closely related to classical time series forecasting methods. Moreover, the Prophet model does not necessitate the use of a machine learning format that involves features and targets, as is typically seen in supervised machine learning algorithms. The Prophet model requires a minimum of two values as input: the ds value, which represents the date value for each data point, and the y value, which represents the measured or observed parameter that is the main focus of the forecasting process.

The PM₁₀ and PM_{2.5} concentrations for the monitoring station Belgrade-Zeleno brdo are obtained from the Environmental Protection Agency of the Republic of Serbia for the period from January 1st, 2021, to August 17th, 2022. Measurements at the monitoring station Zeleno brdo are conducted at an hourly interval. The monitoring station is located at an elevation of 240 m in a suburban area and is part of the SEPA network (Environmental Protection Agency, 2020). Furthermore, the Zeleno brdo measuring station was selected due to its location in a relatively undisturbed and suburban area, which is advantageous for benchmark research. The measured values of PM₁₀ and PM_{2.5} were inspected by the Environmental Protection Agency for possible erroneous values.

The Iowa Environmental Mesonet (2022) provided hourly temperature measurements from Nikola Tesla International Airport in Belgrade, Serbia. While data from a location closer to the particle monitoring station was unavailable, data from Belgrade Airport was deemed sufficient. Notably, the distance by air between the international airport in Belgrade and the monitoring station is 18.7 km, which is acceptable under the circumstances

3. Results and discussion

3.1. Data preprocessing and exploratory data analysis

The graphical representation (Fig. 2) illustrates four distinct categories of missing observations during the data period. The number of short-term missing observations was six, with the first category encompassing missing data for 1 to 3 hours. Mid-term missing observations spanned 24 to 26 hours. The final two categories represented long-term (132 to 148 hours) and extremely long-term missing observations (2,171 hours). In total, there were 2,509 missing observations, equivalent to 17.6% of the total data.

The initial step in data preprocessing consisted of a conditional data amputation, *i.e.*, the reduction of the data set so that the first data would be collected on April 2, 2021. The number of missing observations has been reduced to 338

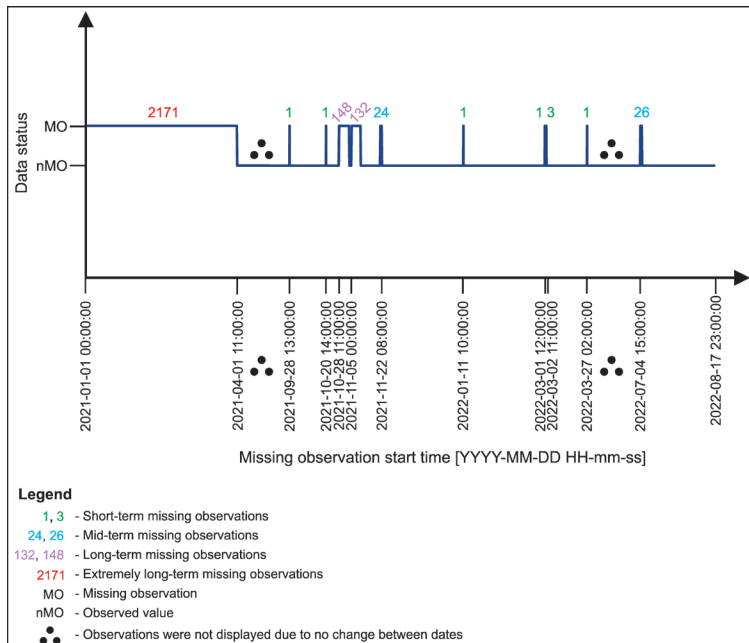


Figure 2. The distribution of missing observations.

hours, or 2.8% of the reduced time interval. In the first stage, short-term missing observations were imputed, and in the final stage, long-term missing observations were imputed. The amputated data set formed the basis for the statistical parameters used in data imputation (Tab. 1).

For short-term missed observations, the last observation carried forward method was primarily used. This method was less efficient than the linear interpolation method because linear interpolation did not alter the basic statistical parameters, it was the optimal method for short-term imputation.

The method employed for addressing mid-term missing observations involved mean substitution with the condition that hourly means were calculated. The hourly mean substitution method permits hourly variations in the PM₁₀ and PM_{2.5} parameters throughout the day, whereas the conventional mean substitution method does not. In the median value of the PM₁₀ parameter, as well as the skewness and kurtosis parameters for both the PM₁₀ and PM_{2.5} parameters, the discrepancy between the basic descriptive statistics computed prior to data imputation for the mid-term missing observations was evident. However, these discrepancies did not exceed 0.7%.

The application of K nearest neighbors and the mean substitution method for long-term data imputation first exhibited its inadequacies in the modification of the PM₁₀ and PM_{2.5} parameter distributions. Visual analysis and Hartigan and Hartigan's unimodality test (Hartigan and Hartigan, 1985) indicated that the fundamental descriptive statistical parameters exhibited a unimodal distribution with a tail extending toward positive values (positively skewed). However, after imputation using the mean substitution method and K nearest neighbors for long-term missing observations, the distribution became bimodal. Consequently, these

Table 1. Statistical parameters during data imputation; T- the distribution is unimodal; CC- correlation coefficient.

Method	Input data		Linear interpolation		Hourly average		Hermite polynomial	
	PM ₁₀	PM _{2.5}	PM ₁₀	PM _{2.5}	PM ₁₀	PM _{2.5}	PM ₁₀	PM _{2.5}
Length of missing observations	/		Short-term		Mid-term		Long-term	
Parameter	PM ₁₀	PM _{2.5}	PM ₁₀	PM _{2.5}	PM ₁₀	PM _{2.5}	PM ₁₀	PM _{2.5}
Number of data points [I]	11746	11746	11754	11754	11804	11804	12084	12084
Minimum [µg/m ³]	2.97	1.94	2.97	1.94	2.97	1.94	2.97	1.94
Maximum [µg/m ³]	218.00	156.00	218.00	156.00	218.00	156.00	218.00	156.00
Mean [µg/m ³]	26.44	18.69	26.44	18.69	26.44	18.69	26.53	18.85
Median [µg/m ³]	22.40	14.60	22.40	14.60	22.50	14.60	22.50	14.70
Mod [µg/m ³]	21.50	11.00	21.50	11.00	21.50	11.00	21.50	11.00
Skewness [I]	2.39	2.07	2.39	2.07	2.40	2.08	2.37	2.07
Kurtosis [I]	11.20	6.30	11.20	6.30	11.26	6.34	10.82	6.13
Unimodality test [I]	T		T		T		T	
Pearson's CC [I]	0.87		0.87		0.87		0.87	
Spearman's CC [I]	0.87		0.87		0.87		0.87	

two methods were automatically dismissed. Several interpolation methods were assessed, and Hermite’s interpolation polynomial was chosen based on its minimal deviation from the fundamental statistical properties. The kurtosis parameter, with a change of 3.39%, exhibited the most significant alteration. For a visual representation of the preprocessed data for the PM_{10} and $PM_{2.5}$ parameters, refer to Fig. 3.

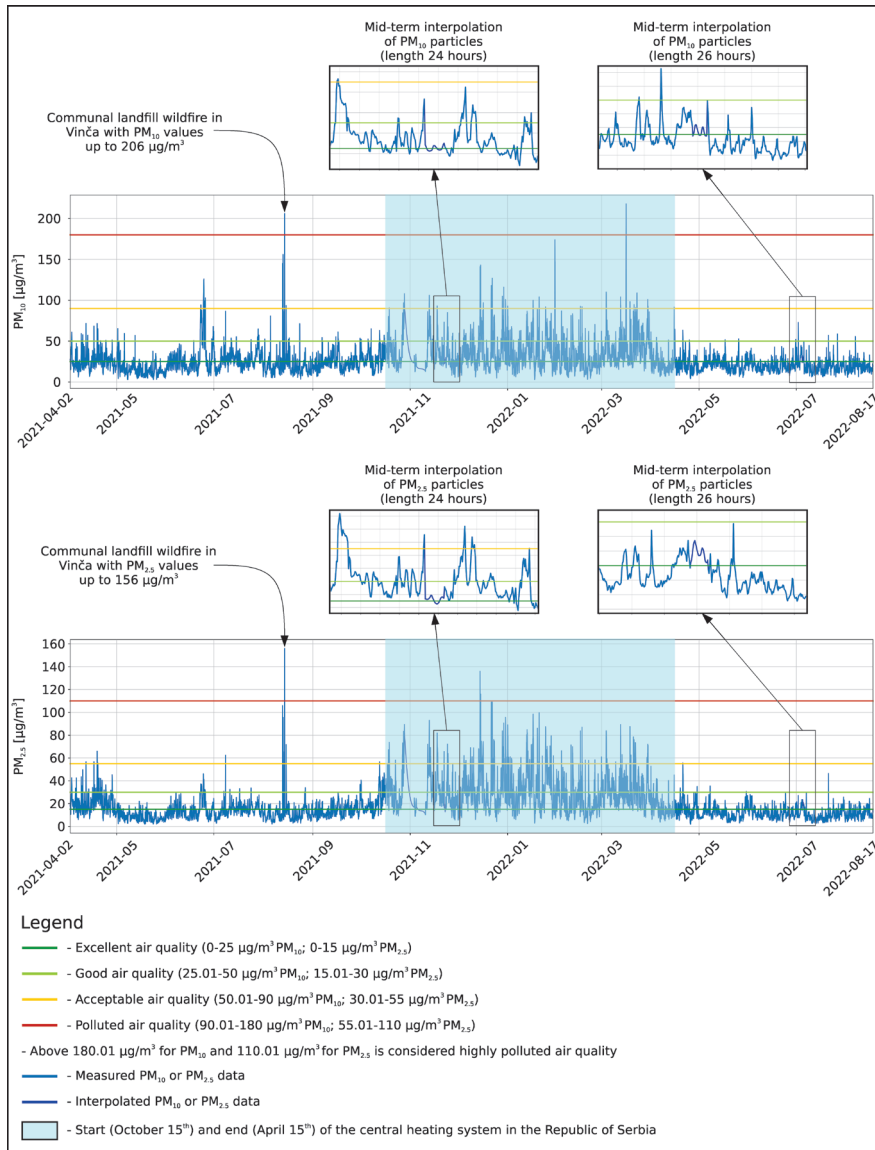


Figure 3. Processed PM_{10} (top) and $PM_{2.5}$ (bottom) data.

Additionally, at the Zeleno brdo monitoring station, the Pearson's and Spearman's correlation coefficients show 0.87 correlation between PM_{10} and $PM_{2.5}$ particles. This correlation falls within the range defined as a strong positive correlation according to Schober et al. (2018).

The processed PM_{10} and $PM_{2.5}$ data (Figs. 3b and 3d) exhibit an intriguing trend from October 2021 to May 2021. The increased values of suspended particles measured at the Belgrade-Zeleno brdo monitoring station may be attributable to the increased use of home heating materials during the winter months. Correlation analysis, using Pearson's and Spearman's correlation coefficients, revealed that PM_{10} particles have a -0.25 (Pearson) and -0.26 (Spearman) correlation with temperature, while the correlation between $PM_{2.5}$ particles and temperature is -0.52 (Pearson) and -0.56 (Spearman). According to the correlation ranges provided by Schober et al. (2018), PM_{10} particles have a weakly negative correlation with temperature, whereas $PM_{2.5}$ particles have a moderately negative correlation. The calculated correlation can be interpreted because the Pearson and Spearman correlation coefficients for both variables are comparable, *i.e.*, are within a close range. At the Belgrade-Zeleno brdo monitoring station, the onset of winter is accompanied by a moderate increase in $PM_{2.5}$ particles and a slight increase in the heavier PM_{10} particles, which can be attributed to the increased use of home heating materials. In this context, it's worth noting that home heating materials have a greater effect on $PM_{2.5}$ particles than PM_{10} particles.

Based on the processed PM_{10} and $PM_{2.5}$ data and the air pollution index from the Environmental Protection Agency (2022), the duration of each index was displayed in Tab. 2.

With the separation of data and the enabling of the central heating system, it was possible to express the same idea in a different way. In the RS, the central heating system is activated on October 15th and remains in operation until April 15th. The data was separated into two categories, one for the period when the central heating was active and the other for the period when it was not active (Tab. 2). Even though the number of hours measured when the central heating system is active (4,386 hours or 36% of the data set) is significantly less than the number of hours measured when the central heating system is not active (7,704 hours or 64% of the data set), the majority of polluted, and very polluted air quality indexes for the $PM_{2.5}$ particles were accounted for by the data measured during active central heating. In contrast, the situation was less severe for PM_{10} particles, where the ratio of heating season to non-heating season for the extremely polluted air quality index was 2:1. (for the $PM_{2.5}$ particles the ratio is 4:1). The ratio of polluted air quality indexes during the heating season to non-heating season was 23:1 (365:16) for $PM_{2.5}$ and 3:1 (67:21) for PM_{10} particles.

In congruence with the correlation analysis displayed previously, when the annual temperature drops and households and the central heating system were

Table 2. The air quality index and its duration at the monitoring station in Belgrade-Zeleno brdo during heating and non-heating season.

Air quality index	Non heating season PM ₁₀ [h]	Heating season PM ₁₀ [h]	Both heating and non-heating season PM ₁₀ [h]
Excellent	5456	1677	7133
Good	2050	1908	3958
Acceptable	176	714	890
Polluted	21	67	88
Very polluted	1	2	3
Air quality index	Non heating season PM _{2.5} [h]	Heating season PM _{2.5} [h]	Both heating and non-heating season PM _{2.5} [h]
Excellent	5332	882	6214
Good	2167	1790	3957
Acceptable	188	1327	1515
Polluted	16	365	381
Very polluted	1	4	5

active, a greater effect should be observed for PM_{2.5} particles compared to PM₁₀ particles, and the material used for household heating has an undeniable impact on the air quality index at the Belgrade- Zeleno brdo measuring station.

3.2. Forecasting of PM₁₀ and PM_{2.5} concentrations

As both PM₁₀ and PM_{2.5} values were provided as a single time series with a total length of 12,084 hours, forecasting was conducted using multiple shorter time series, each with a fixed 24-hour forecasting horizon (Fig. 4). This approach allowed for more forecasts with shorter time series, enabling a more comprehensive statistical validation of the selected model. Figure 4 depicts shorter time series, each with duration of 120 hours. The first step was the first 120 hours, or from April 2nd to April 6th, 2021, with a forecast horizon of 24 hours, or the entire 7th of April. Five quantitative assessment parameters (*AE*, *MAE*, *APE*, *MAPE*, *RMSE*) and the forecasted value of PM₁₀ and PM_{2.5} were the outputs of such a forecast. The next step was the next 120 hours, or from the 8th to the 12th of April, with the 13th of April being the forecasted and validation date. Such iterations are performed for the entire data set, with the length of *n* (in this case 120 hours). There were a total of 83 distinct time series and forecasts made for a 120 hour period. The length of the introductory time series was then extended from 120 hours to 240 hours and further until it covered the entire dataset of 12,048 hours.

A parameter can be introduced that represents the ration of the introductory time series (number of historical data) relative to the length of the forecasting

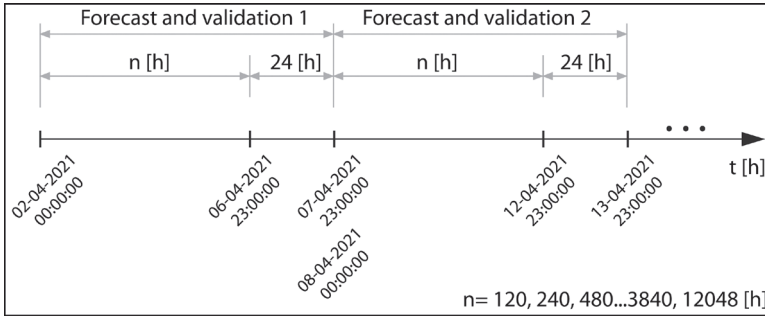


Figure 4. The principle of separating a time series into multiple shorter time series.

horizon. This ratio increases as the number of historical data increases while the forecasting horizon remains the same. For the introductory time series with a length of 120 hours and a forecasting horizon of 24 hours, the ratio was defined as 5:1, meaning that five historical data points were utilized to generate one future data point. The maximum value of such a ratio, however, was 501:1 when the introductory time series has a length of 12,048 hours, and the forecasting horizon was not altered. Table 3 presents such a ratio, the number of predictions made for the given ratio, and the minimum, mean, median, and maximum values of the *RMSE* calculated based on all predictions made for the ratio. The ratio 12048:24 was excluded from Tab. 3 as it was described in greater detail individually.

The *RMSE* values obtained from Tab. 3 for PM_{10} and $PM_{2.5}$ suspended particles suggest that the Prophet model may, in some cases, result in significant deviations between predicted and true values. In the case of a ratio of 480:24 for PM_{10} concentration, the *RMSE* in one forecast was $50.78 \mu\text{g}/\text{m}^3$. In that situation, the *MAE* was $49.54 \mu\text{g}/\text{m}^3$ and the *MAPE* was 57.34%. The validation window for this time series, otherwise, contained values that were unexpectedly high. The mean PM_{10} value for the validation window was about $85 \mu\text{g}/\text{m}^3$, whereas the mean for the entire data set was $26.53 \mu\text{g}/\text{m}^3$. In addition, for both the PM_{10} and $PM_{2.5}$ parameters, an introductory time series of 1920 hours yielded the lowest calculated error parameters. Based on the number of predictions made for the given duration (6), it cannot be stated with a high degree of certainty that the event was not random. Additional modelling and forecasting should be done in the future to confirm the optimal length of the introductory time series for the Prophet model.

The mean *RMSE* values were lower for $PM_{2.5}$ particle concentrations across all ratios compared to PM_{10} , and the same holds true for the median *RMSE* value. The calculated *RMSE* values indicate that the model displayed more accurate $PM_{2.5}$ concentration values. Based on all 173 forecasts, 50% of forecasts will have a lower *RMSE* value than $6.26 \mu\text{g}/\text{m}^3$ for $PM_{2.5}$ concentrations and $9.99 \mu\text{g}/\text{m}^3$ for PM_{10} concentrations.

Table 3. Root mean square error for shorter time series.

Parameter	Ratio	Number of forecasts	$RMSE_{min}$ [$\mu\text{g}/\text{m}^3$]	$RMSE_{mean}$ [$\mu\text{g}/\text{m}^3$]	$RMSE_{median}$ [$\mu\text{g}/\text{m}^3$]	$RMSE_{max}$ [$\mu\text{g}/\text{m}^3$]
PM _{2.5}	120:24 (5:1)	83	1.04	9.20	6.35	33.81
	240:24 (10:1)	45	1.41	8.45	5.82	38.71
	480:24 (20:1)	23	2.85	11.42	9.04	46.02
	960: 24 (40:1)	12	1.96	10.49	8.29	31.03
	1920:24 (80:1)	6	1.98	4.83	4.87	7.65
	3840:24 (160:1)	3	2.00	6.24	3.78	12.93
PM ₁₀	120:24 (5:1)	83	1.64	12.85	10.79	41.21
	240:24 (10:1)	45	2.30	11.01	8.21	46.11
	480:24 (20:1)	23	3.77	16.19	15.26	50.78
	960: 24 (40:1)	12	3.77	15.23	11.52	49.25
	1920:24 (80:1)	6	3.12	6.47	6.17	9.39
	3840:24 (160:1)	3	4.49	8.97	7.04	15.39

Figure 5 depicts the forecast for the entire data set, using 12048 hours of data were used as the introductory time series to obtain 24-hour future values. Figure 5a depicts the PM₁₀ particle forecast, the *AE* and *APE* parameters. The *AE* fluctuates between 2 $\mu\text{g}/\text{m}^3$ and 10 $\mu\text{g}/\text{m}^3$, with a mean value of 6.5 $\mu\text{g}/\text{m}^3$. The *APE* ranges from 10.5% to 93.0%, with an average of 46.45%. Otherwise, as expected, the *AE* values for PM_{2.5} particles were lower, ranging from 0.73 $\mu\text{g}/\text{m}^3$ to 5 $\mu\text{g}/\text{m}^3$ with a mean *MAE* of 2.7 $\mu\text{g}/\text{m}^3$. Similarly, the *MAPE* was less than the *MAPE* for PM₁₀ particles. High *APE* and *MAPE* values for PM₁₀ and PM_{2.5} particles were not alarming. Even though the *AE* was small, the *APE* can be relatively large when predicting low values, demonstrating the inadequacy of using the two aforementioned errors. Figure 5b's 15th hour PM_{2.5} particle forecast can be used as an example. The true value was 7.04 $\mu\text{g}/\text{m}^3$, while the predicted value was 8.84 $\mu\text{g}/\text{m}^3$. In this instance, the *AE* was relatively small at 1.8 $\mu\text{g}/\text{m}^3$, but the *APE* was high at 25%. The contrast also applies to the prediction of high values. As a result, we advise to calculate and display additional quantitative assessment parameters, which in this case was the *RMSE*. When interpreting *AE* and *APE* values, it is crucial to consider the predicted order of magnitude.

The quality of the Prophet model was also evident when displaying daily variations in PM₁₀ and PM_{2.5} concentrations, an essential forecasting parameter. In both cases illustrated in Fig. 5, the model displayed daily PM₁₀ and PM_{2.5} variations. While the predicted values may differ from the actual measurements, they adequately represent the general trend of daily concentration increases and

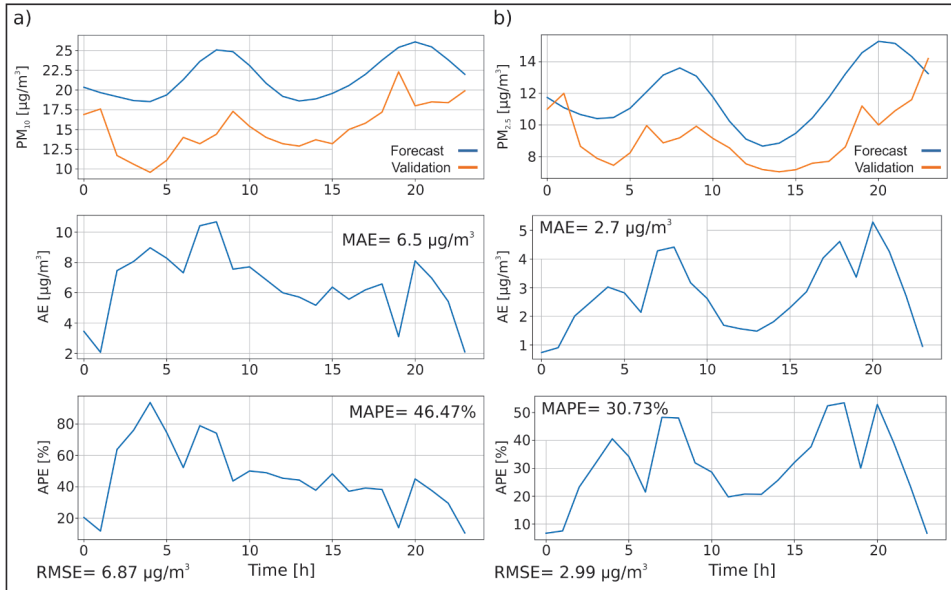


Figure 5. (a) Forecast of the PM_{10} concentration, absolute and absolute percentage error; (b) Forecast of the $PM_{2.5}$ concentration, absolute and absolute percentage error.

decreases. The conformity of the true and forecasted curves from Fig. 5 can be adequately seen in the trend component, even though the values are not absolutely accurate. This research can be compared to the work of Shen et al. (2020); both studies agree that the Prophet model provides more accurate forecasts for $PM_{2.5}$ particles compared to PM_{10} particles.

Future research in the field of $PM_{2.5}$ air quality forecasting could incorporate various novel forecasting techniques, including Uber's EN-RNN (Smyl, 2020), FFORMA (Manso et al., 2020), forecasHybrid (Shaub, 2020), and ShoTS (Thomakos et al., 2022), tailored for shorter time series and other time series forecasting methods. Aside from conventional time series forecasting methods and machine learning techniques.

4. Conclusion

Accurately predicting future concentrations of suspended particles and other air quality parameters is crucial information that authorities can utilize to mitigate the risks of cardiovascular and respiratory diseases. In this research, the Prophet model was utilized for short-term (24-hour in advance) PM_{10} and $PM_{2.5}$ particle forecasting. Based on 173 forecasts with varying lengths of the introductory time series, it is expected that Prophet model will produce *RMSE* values that are less than $6.26 \mu\text{g}/\text{m}^3$ and $9.9 \mu\text{g}/\text{m}^3$ for $PM_{2.5}$ and PM_{10} particles, respectively.

The calculated error parameters for the longest introductory time series (12048 hours) confirmed that the *RMSE* is smaller for the $PM_{2.5}$ parameter ($2.99 \mu\text{g}/\text{m}^3$) than for the PM_{10} parameter ($6.87 \mu\text{g}/\text{m}^3$). Additionally, the Prophet model adequately reflected daily variations in PM_{10} and $PM_{2.5}$ parameters, even though the values were not forecasted with absolute precision.

This research should not be interpreted as a stand-alone study, but rather as an introduction to a larger study that will include additional monitoring stations in Belgrade and other air quality parameters as well as hourly traffic conditions and other meteorological data. Furthermore, various models should be evaluated using various data transformation techniques (such as normalization and stationarity). Developing a method for accurate forecasts would unquestionably increase the likelihood of successfully mitigating air pollution, which could lead to a reduction in cardiovascular and respiratory diseases. Future research should also examine whether the Prophet model is suitable for longer-term forecasts.

Acknowledgements – The authors express gratitude to the Environmental Protection Agency of the Republic of Serbia for providing data for this research. Additionally, the research was realized with financial support from the Society of Exploration Geophysicists (Norman and Shirley Domenico Scholarship).

References

- Agrawal, R. K. and Adhikari, R. (2013): An introductory study on time series modeling and forecasting. New York, CoRR, pp. 67, <https://doi.org/10.48550/arXiv.1302.6613>.
- Araujo, J. A. (2011): Particulate air pollution, systemic oxidative stress, inflammation, and atherosclerosis, *Air Qual. Atmos. Health.*, **4**, 79–93, <https://doi.org/10.1007%2Fs11869-010-0101-8>.
- Bernstein, J. A., Alexis, N., Barnes, C., Bernstein, I. L., Nel, A., Peden, D., Diaz-Sanchez, D., Tarlo, M. T. and Williams, P. B. (2004): Health effects of air pollution, *J. Allergy Clin. Immunol.*, **114**, 1116–1123, <https://doi.org/10.1016/j.jaci.2004.08.030>.
- Bisgaard, S. and Kulahci, M. (2011): *Time series analysis and forecasting by example*. John Wiley & Sons, pp 400.
- Dockery, D. W., Schwartz, J. and Spengler, J. D. (1992): Air pollution and daily mortality: associations with particulates and acid aerosols, *Environ. Res.*, **59**, 362–373, [https://doi.org/10.1016/s0013-9351\(05\)80042-8](https://doi.org/10.1016/s0013-9351(05)80042-8).
- Ejohwomu, O. A., Shamsideen Oshodi, O., Oladokun, M., Bukoye, O. T., Emekwuru, N., Sotunbo, A. and Adenuga, O. (2022): Modelling and forecasting temporal $PM_{2.5}$ concentration using ensemble machine learning methods, *Buildings*, **12**, 46, <https://doi.org/10.3390/buildings12010046>.
- Energy Agency of the Republic of Serbia (2020): Report on the work of the Energy Agency for the year 2020. *Energy Agency of the Republic of Serbia* (in Serbian).
- Environmental Protection Agency of the Republic of Serbia (2022): Air pollution index. *Environmental Protection Agency of the Republic of Serbia* (last accessed on 15 September 2022, available at: <http://www.amskv.sepa.gov.rs/kriterijumi.php>) (Online, in Serbian).
- Hartigan, J. A. and Hartigan, P. M. (1985): The dip test of unimodality, *Ann. Stat.*, **13**, 70–84, <https://doi.org/10.1214/aos/1176346577>.
- Hyndman, R. J. and Koehler, A. B. (2006): Another look at measures of forecast accuracy, *Int. J. Forecast.*, **22**, 679–688, <https://doi.org/10.1016/j.ijforecast.2006.03.001>.

- Iowa Environmental Mesonet (2022): *Iowa Environmental Mesonet ASOS Network*. Iowa State University (last accessed on 15 September 2022, available at: <https://mesonet.agron.iastate.edu/request/download.phtml>).
- Jovanović, M. V. (2020): *Chemical content and oxidative potential of respirable particulate matter in urban and in industrial environments*. PhD dissertation, Faculty of Chemistry, University of Belgrade, pp 106.
- Kang, H. (2013): The prevention and handling of the missing data, *Korean. J. Anesthesiol.*, **64**, 402–406, <https://doi.org/10.4097%2Fkjae.2013.64.5.402>.
- Kovačić, J. Z. (1995): *Time series analysis*. Faculty of Economics, University of Belgrade, pp 354 (in Serbian).
- L. Zhou, M. Chen and Q. Ni. (2020): A hybrid Prophet-LSTM model for prediction of air quality index, *2020 IEEE Symposium Series on Computational Intelligence*. 595–601, <https://doi.org/10.1109/SSCI47803.2020.9308543>.
- Lütkepohl, H. and Krätzig, M. (2004): *Applied time series econometrics*. Cambridge University Press, pp 323.
- Montero-Manso, P., Athanasopoulos, G., Hyndman, R. J. and Talagala, T. S. (2020): FFORMA: Feature-based forecast model averaging, *Int. J. Forecast.*, **36**, 86–92, <https://doi.org/10.1016/j.ijforecast.2019.02.011>.
- Official Gazette of the Republic of Serbia (156/2020): Decision on determining the energy balance of the Republic of Serbia for 2021 (156/2020).
- Papastefanopoulos, V., Linardatos, P. and Kotsiantis, S. (2020): COVID-19: A comparison of time series methods to forecast percentage of active cases per population, *Appl. Sci.*, **10**(11), 3880, <https://doi.org/10.3390/app10113880>.
- Puljić, V. M., Jones, D., Moore, C., Myllyvirta, L., Gierens, R., Kalaba, I., Ciuta, I., Gallop, P. and Risteska, S. (2019): *Chronic coal pollution – EU action in the Western Balkans will improve health and economies across Europe*. HEAL, CAN Europe, Sandbag, CEE Bankwatch Network i Europe Beyond Coal (in Serbian).
- Samal, K. K. R., Babu, K. S., Das, S. K. and Acharaya, A. (2019): Time series based air pollution forecasting using SARIMA and prophet model, in: *Proceedings of the 2019 International Conference on Information Technology and Computer Communications*, 80–85, <https://doi.org/10.1145/3355402.3355417>.
- Schober, P., Boer, C. and Schwarte, L. A. (2018): Correlation coefficients: appropriate use and interpretation, *Anesth. Analg.*, **126**, 1763–1768, [10.1213/ANE.0000000000002864](https://doi.org/10.1213/ANE.0000000000002864).
- Shakeel, A., Chong, D. and Wang, J. (2023a): Load forecasting of district heating system based on improved FB-Prophet model, *Energy*, **278**, 127637, <https://doi.org/10.1016/j.energy.2023.127637>.
- Shakeel, A., Chong, D. and Wang, J. (2023b): District heating load forecasting with a hybrid model based on LightGBM and FB-prophet, *J. Clean. Prod.*, **409**, p. 137130, <https://doi.org/10.1016/j.energy.2023.137130>.
- Shaub, D. (2020): Fast and accurate yearly time series forecasting with forecast combinations, *Int. J. Forecast.*, **36**, 116–120, <https://doi.org/10.1016/j.ijforecast.2019.03.032>.
- Shen, J., Valagolam, D. and McCalla, S. (2020): Prophet forecasting model: A machine learning approach to predict the concentration of air pollutants (PM_{2.5}, PM₁₀, O₃, NO₂, SO₂, CO) in Seoul, South Korea. *Peer J.*, **8**, <https://doi.org/10.7717/2Fpeerj.9961>.
- Smith, Taylor G., et al. (2017): pmdarima: ARIMA estimators for Python, <http://www.alkaline-ml.com/pmdarima> (Online; last accessed on 18 October 2023).
- Smyl, S. (2020): A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting, *Int. J. Forecast.*, **36**, 75–85, <https://doi.org/10.1016/j.ijforecast.2019.03.017>.
- Taylor, S. J. and Letham, B. (2018): Forecasting at scale, *Am. Stat.*, **72**, 37–45, <https://doi.org/10.1080/00031305.2017.1380080>.
- Tejasvini, K. N., Amith, G. R. and Shilpa, H. (2020): Air pollution forecasting using multiple time series approach, in: *Proceedings of the Global AI Congress 2019*, 91–100, https://doi.org/10.1007/978-981-15-2188-1_8.

- Thomakos, D., Wood, G., Ioakimidis, M. and Papagiannakis, G. (2023): ShoTS forecasting: Short time series forecasting for management research, *Brit. J. Manage.*, <https://doi.org/10.1111/1467-8551.12624>
- World Health Organization (2019): *Health impact of ambient air pollution in Serbia: A call to action*. World Health Organization.
- Ye, Z. (2019): Air pollutants prediction in Shenzhen based on ARIMA and Prophet method, *E3S Web Conf.*, **136**, 1–5, <https://doi.org/10.1051/e3sconf/201913605001>.
- Zhao, N., Liu, Y., Vanos, J. K. and Cao, G. (2018): Day-of-week and seasonal patterns of PM_{2.5} concentrations over the United States: Time-series analyses using the Prophet procedure, *Atmos. Environ.*, **192**, 116–127, <https://doi.org/10.1016/j.atmosenv.2018.08.050>.

SAŽETAK

Kratkoročno predviđanje koncentracija PM₁₀ i PM_{2.5} pomoću Facebookovog Prophet modela na Beograd-Zeleno brdo

Filip Arnaut, Vesna Cvetkov, Dragana Đurić i Mileva Samardžić-Petrović

Demonstriramo korištenje Facebookovog Prophet (obično samo Prophet) modela za kratkoročnu prognozu kvalitete zraka na mjernoj stanici Beograd-Zeleno brdo. Kako bismo riješili podatke koji nedostaju, primijenili smo tehnike imputacije podataka koje minimalno mijenjaju distribuciju. Linearna interpolacija pokazala se učinkovitom za kratkoročne praznine (1–3 sata), metoda satnih srednjaka za srednjoročne praznine (24–26 sati), a Hermite interpolacijski polinom za dugoročne praznine (132–148 sati). Najznačajnija promjena podataka bio je pomak od 3,4% u asimetričnosti. Podjelom vremenskog niza omogućena je detaljna procjena kvalitete modela Prophet, pri čemu su predviđanja PM_{2.5} bila preciznija od PM₁₀. Korištenje najdužeg vremenskog niza za prognoziranje dalo je apsolutne pogreške od 6,5 µg/m³ za PM₁₀ i 2,7 µg/m³ za PM_{2.5}. Na temelju 173 predviđanja, predviđamo srednje kvadratne vrijednosti Prophetovog modela ispod 6,26 µg/m³ i 9,99 µg/m³ za PM_{2.5} i PM₁₀ u 50% slučajeva. Model Prophet pokazuje nekoliko prednosti i daje zadovoljavajuće rezultate. U budućim istraživanjima, rezultati dobiveni iz Prophet modela poslužiti će kao referentne vrijednosti za druge modele. Osim toga, model Prophet može pružiti zadovoljavajuće rezultate predviđanja kvalitete zraka i koristiti će se u budućim istraživanjima.

Ključne riječi: analiza vremenskih nizova, predviđanje vremenskih nizova, kvaliteta zraka, lebdeće čestice.

Corresponding author's address: Filip Arnaut, Department of Geophysics, Faculty of Mining and Geology, University of Belgrade, Đušina 7, 11120 Belgrade, Serbia; tel: +381 63 83 89 237; e-mail: filip.arnaut@rgf.rs



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Energetic solar flare events in relation with subionospheric impact on 6-10 September 2017: data and modeling

A. Kolarski, V.A. Srećković , M. Langović and F. Arnaut

*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
11080 Belgrade, Serbia, (E-mail: aleksandra.kolarski@ipb.ac.rs)*

Received: September 15, 2023; Accepted: October 11, 2023

Abstract. Solar flares are among main extraterrestrial events that are well known to severely affect both space weather and near-Earth surrounding conditions. Under incident X-ray solar flare radiation, ionospheric plasma undergoes perturbations clearly and distinctly observable on ground-based monitoring systems' recordings. Mid-latitude lower ionosphere under influence of energetic solar flare events was examined by employing different numerical modeling procedures, including Machine Learning, relying on ground-based Very Low Frequency (VLF) radio signal recordings from Belgrade VLF database and solar soft X-ray irradiance satellite measurements taken from Geostationary Operational Environmental Satellite (GOES) database.

Key words: Solar activity – Solar X-ray flares – radio signal perturbations – GOES – data – modeling

1. Introduction

Through complex solar–terrestrial interactions, Earth is continually under Sun's emitted radiation influences, both of the electromagnetic and corpuscular nature (Kelly, 2009). As the major source of numerous and diverse driving agents related with Earth's magnetospheric and ionospheric perturbations, Sun's varying activity is under ongoing increasingly extensive research (see Hayes et al., 2021; Rycroft et al., 2000, and references therein). As modern society progressively becomes more and more dependent on complex technological systems including satellite, telecommunication and power grid networks, solar activity as potentially hazardous to these systems and consequently potentially threatening to an urban way of life, gained also in recent years more attention outside pure scientific community (Yasyukevich et al., 2018).

As still reliably unpredictable to current knowledge, solar activity especially in terms of high and extreme energetic solar events' occurrences, such as e.g. energetic eruptions of high class solar flare (SF) events and coronal mass ejections (CMEs) and their interactions, is often presented through analyses of separate events which make case studies very important in the sense of a comprehensive

overview of evolution of such complex natural phenomena and their impacts on near Earth's surroundings (Srećković et al., 2021a; Kolarski et al., 2023). Here, solar activity during the most active period of the descending branch of the solar cycle (SC) 24, with the respect of high class X-ray solar flare events with geo-effective implications, is presented as a case study of solar conditions throughout a period enclosing the strongest SF events in September 2017, as observed by ground-based systems located in Belgrade (Serbia) at the Institute of physics Belgrade (44.85°N; 20.38°E), through subionospheric Very Low Frequency (VLF) propagation disturbances (e.g. Šulić et al., 2016) recorded on signal emitted from UK, employing four numerical techniques developed for retrieving lower ionospheric plasma parameters (Silber & Price, 2017; Arnaut et al., 2023).

Table 1. X-class SFs from September 2017, from GOES15 database.

Fare date	Class	Region	Time UT			I_x^{max} (10^{-4} Wm^{-2})
			Start	Max.	End	
6 September 2017	X2.2	2673	08:57	09:10	09:17	2.2658*
6 September 2017	X9.3	2673	11:53	12:02	12:10	9.3293
7 September 2017	X1.3	2673	14:20	14:36	14:55	1.3880
10 September 2017	X8.2	2673	15:35	16:06	16:31	8.2808

* GOES13 data

2. Observations

Year of 2017 belongs to the descending branch of the SC 24, the cycle characterized by morphology notably decreased compared to several previous cycles, and with the intensity that is fourth smallest since the solar cycle 1 (Kolarski et al., 2022; Grodji et al., 2022). Even though it is placed close to the solar minimum between SCs 24 and 25, regarding solar flare activity year of 2017 stands out from other surrounding years from this descending branch, especially taking into consideration very active periods during months of September (when in total 99 SFs occurred, with 27 M-class and 4 X-class events, from which 2 of them were the strongest of SC 24) and April (when in total 52 SFs occurred, with 7 M-class and without X-class events). Solar conditions during September 2017 are mainly related to active region 2673 (AR12673) from which the strongest, 4 X-class SFs of intensity in range X1.3-X9.3 occurred during days of 06, 07 and 10 September (with X2.2 & X9.3 on 06 September, X1.3 on 07 September and X8.2 on 10 September). This active region with complex evolution appeared on 29 August on the south-eastern limb of the Sun and on 11 September disappeared from the Earth-view of the Sun, producing in total 77 SFs, of which 1 B (B9.8), 45 C in range C1-C9.8, 27 M in range M1-M8.1 and 4 X-class flares in range X1.3-X9.3. It is important to note that all M- and X-class SFs from September 2017 are related to this active region. Regarding the rest of the September, only

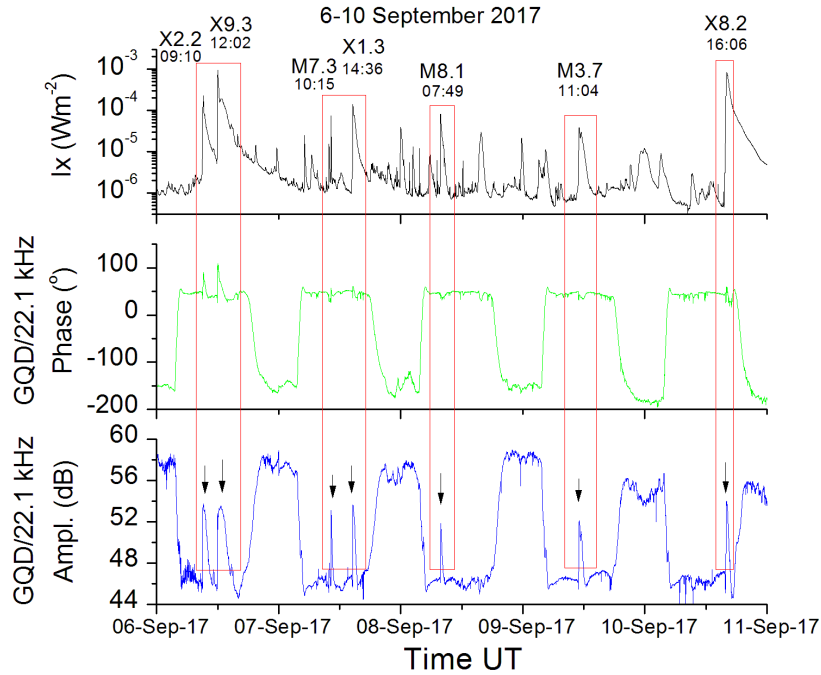


Figure 1. GQD signal perturbations registered in Belgrade during period of high solar activity from 06-10 September 2017 and GOES solar X-ray flux data as shown in the panels from lower to upper, respectively. Detected strong solar flares are marked with red rectangle.

weak flare activity of B- and C-class related to active regions 2677, 2680, 2681, 2682 (re-designated former active region 2673 which reappeared on 24 September and became again visible at the eastern limb of the Sun) and 2683 was reported, with total of 38 SFs in range B1-C3, with just 5 B-class SFs in range B2.2-B4.2 from active region 2682. There were also some days without reported flare activity, of which 6 days in a row 14-19 September and 2 days of 21-22 September. Main characteristics of X-class SFs from September 2017 are given in Table 1, with solar soft X-ray flux (0.1-0.8 nm) taken from Geostationary Operational Environmental Satellite (GOES) archive database (<https://satdat.ngdc.noaa.gov/sem/goes/data/avg/>).

Solar flare event of class X9.3, that occurred during solar cycle's declining branch relatively close to the solar minimum, particularly is interesting since event of X-class was not occurred as far back as on 05 May 2015 (X2.7), and of similar strength as far back as on 05 December 2006 (X9). If we exclude some weaker X-class SFs (e.g. as on 06 August 2011 (X6.9), 06 December 2006 (X6.5), 09 September 2005 (X6.2) and 08 September 2005 (X5.4)) and some significantly stronger SFs (e.g. as on 07 September 2005 (X17+), 04 November

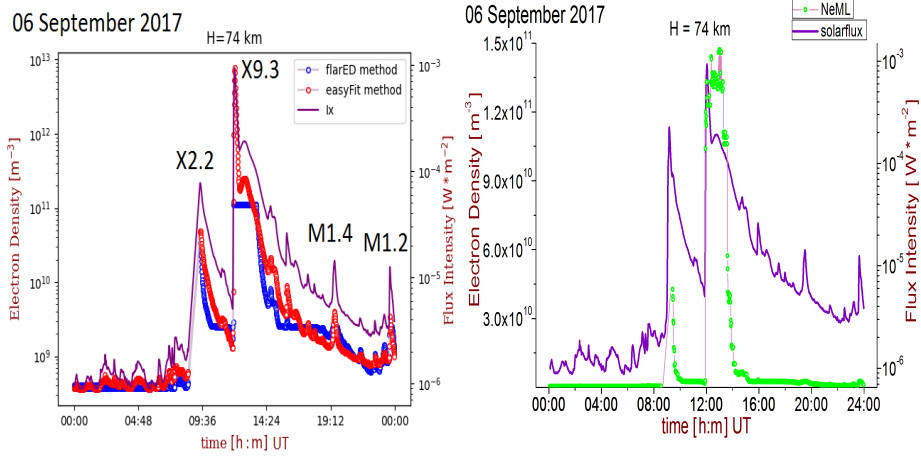


Figure 2. Variations of GOES X-ray flux (in violet) and Ne ($h = 74$ km) in function of UT on 06 September 2017 obtained by FlarED and easyFit methods (in blue and red, respectively) on the left and from ML method (in green) on the right.

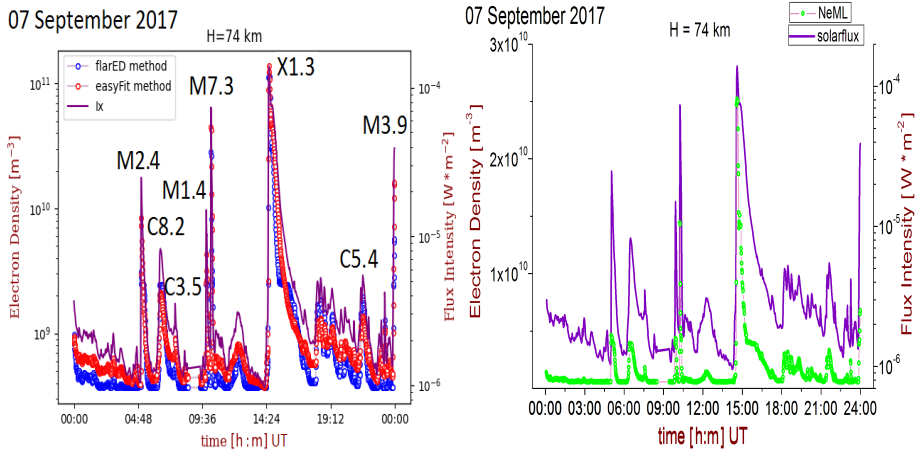


Figure 3. Variations of GOES X-ray flux (in violet) and Ne ($h = 74$ km) in function of UT on 07 September 2017 obtained by FlarED and easyFit methods (in blue and red, respectively) on the left and from ML method (in green) on the right.

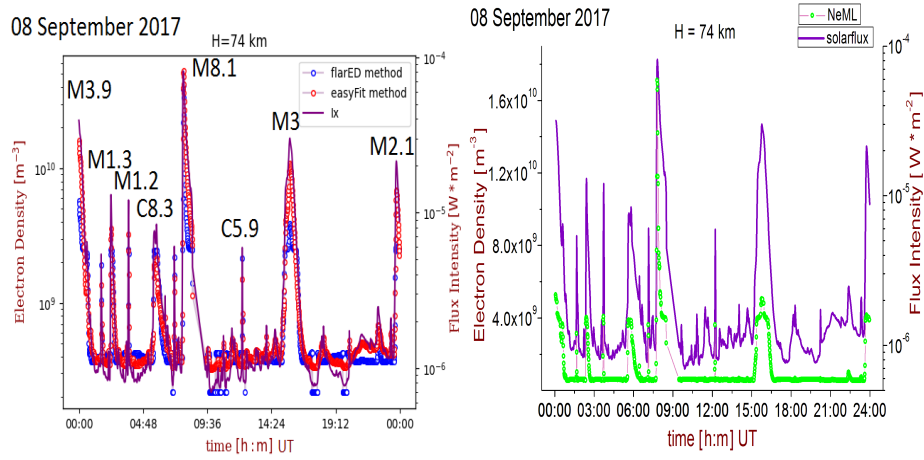


Figure 4. Variations of GOES X-ray flux (in violet) and Ne ($h = 74$ km) in function of UT on 08 September 2017 obtained by FlarED and easyFit methods (in blue and red, respectively) on the left and from ML method (in green) on the right.

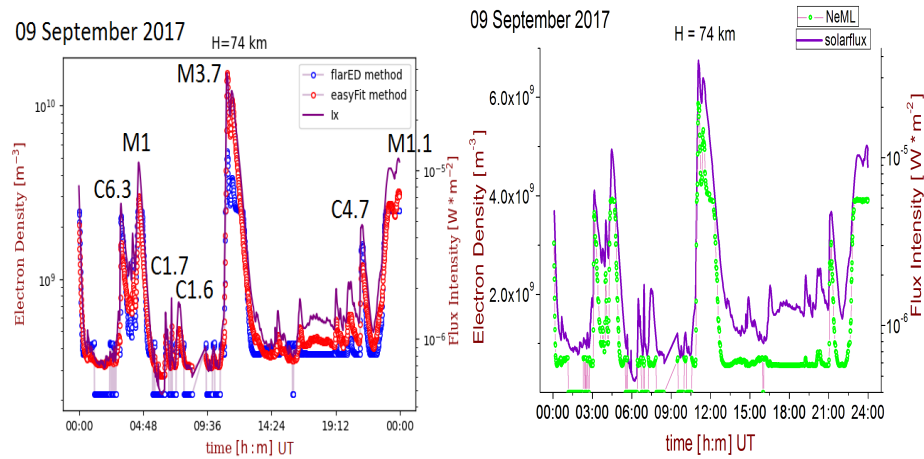


Figure 5. Variations of GOES X-ray flux (in violet) and Ne ($h = 74$ km) in function of UT on 09 September 2017 obtained by FlarED and easyFit methods (in blue and red, respectively) on the left and from ML method (in green) on the right.

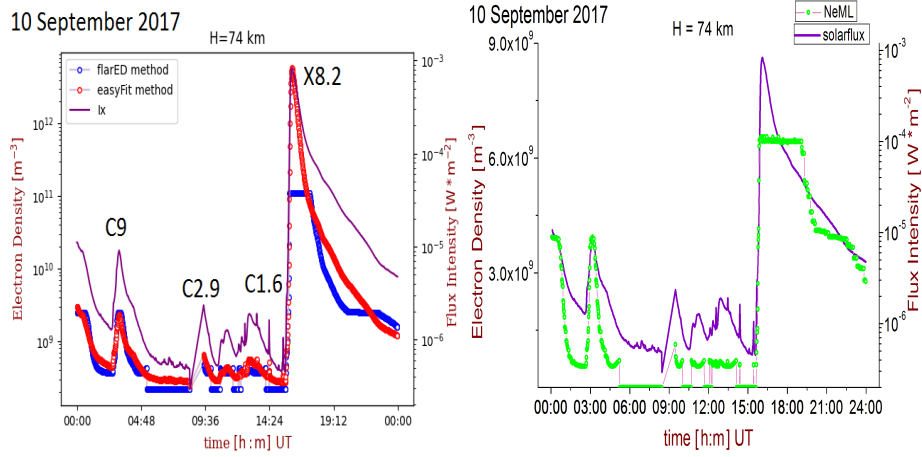


Figure 6. Variations of GOES X-ray flux (in violet) and N_e ($h = 74$ km) in function of UT on 10 September 2017 obtained by FlarED and easyFit methods (in blue and red, respectively) on the left and from ML method (in green) on the right.

2003 (X28+), 28 October 2003 (X17.2+), 02 April 2001 (X20+) and 15 April 2001 (X14.4)), SF of similar strength as X9.3 from September 2017 occurred about two decades in the past i.e. on 06 November 1997 (X9.4). During 06, 07 and 10 September R3 radio blackouts related to X-class SFs were reported.

3. Results and discussion

VLF signal's response to strong solar X-ray flare events from September 2017 was monitored in Belgrade (44.85°N; 20.38°E) on signal transmitted from UK (GQD/22.1 kHz), using recordings from BEL AWESOME (Atmospheric Weather Electromagnetic System for Observation Modeling and Education) receiving system operating in narrow-band mode (Šulić et al., 2016). Response of GQD signal to these X-class SF together with corresponding solar X-ray irradiances during the most active days in September 2017 are given in Figure 1. Under strong solar X-ray radiation in soft range, GQD signal's amplitude and phase responded by following incident X-ray flux with time delay corresponding to the sluggishness (e.g. Mitra, 1974; Hayes et al., 2021; Žigman et al., 2023), with perturbations of complex morphology, characteristic for this signal as recorded in Belgrade (Kolarski & Grubor, 2014; Grubor et al., 2008). Amplitude and phase perturbations induced by these 4 X-class SFs reached up to 8 dB in amplitude and up to a few tens of degrees in phase, compared to unperturbed ionospheric conditions. Based on recorded GQD signal's amplitude and phase perturbations, modeling of propagation parameters (Wait & Spies, 1964) was conducted using Long Wave Propagation Capability (LWPC) software (Ferguson, 1998), with

goal that modeled and real conditions within waveguides match as close as possible. Aside LWPC, for obtaining lower ionospheric plasma parameters, three other numerical techniques were applied: FlareED and easyFit (Srećković *et al.*, 2021a,b) and also novel method based on Machine Learning (ML) approach.

Table 2. Analyzed intense SF events from September 2017.

(dd/mm/yy) Class I_x^{max} time (UT)	N_{eLWPC} (m^{-3})	$N_{easyFit}$ (m^{-3})	$N_{FlareED}$ (m^{-3})	N_{ML} (m^{-3})
06/09/17 X2.2 09:10	$2.06 \cdot 10^{11}$	$3.49 \cdot 10^{11}$	$1.12 \cdot 10^{11}$	$1.23 \cdot 10^{11}$
06/09/17 X9.3 12:02	$1.15 \cdot 10^{12}$	$7.67 \cdot 10^{12}$	$1.12 \cdot 10^{11**}$	$1.33 \cdot 10^{11*}$
07/09/17 M7.3 10:15	$1.11 \cdot 10^{11}$	$4.45 \cdot 10^{10}$	$2.85 \cdot 10^{10}$	$1.44 \cdot 10^{10}$
07/09/17 X1.3 14:36	$1.76 \cdot 10^{10}$	$1.37 \cdot 10^{11}$	$1.12 \cdot 10^{11}$	$2.52 \cdot 10^{10}$
08/09/17 M8.1 07:49	$6.59 \cdot 10^{10}$	$5.25 \cdot 10^{10}$	$4.17 \cdot 10^{10}$	$1.66 \cdot 10^{10}$
10/09/17 X8.2 16:06	$2.12 \cdot 10^{11*}$	$5.79 \cdot 10^{12}$	$1.12 \cdot 10^{11**}$	$6.44 \cdot 10^9*$

*Unreliable result

**Saturation

In contrast to other energetic SFs reported within this active period, which were directed towards the Earth, main difference of X8.2 SF is in terms of geoeffective potential, since this event was the only one SF that erupted almost on the far side of the Sun's surface and did not directly affected the Earth. However, since this event was highly energetic, Earth was still affected in great deal, which can be clearly seen in perturbations of subionospherically propagating GQD signal. Another specific feature related to this event is that during the occurrence of X8.2 SF on GQD signal recorded in Belgrade there was a transition period from stable daytime to stable nighttime ionospheric conditions related to terminator period of sunset. Although process of modeling lower ionospheric response related to high class SFs based on VLF parameters is challenging by itself, some special circumstances, such as e.g. SF's occurrence during stable

daytime ionospheric conditions (such in case of e.g. X2.2 SF) and relatively not so high absolute amount of amplitude and/or phase perturbations (such in cases of some mid class SFs) etc., can go in favor especially when classical approach using LWPC software is employed. However, in conditions related to terminator periods modeling of SFs is especially challenging, which is emphasized in much greater deal when energetic ones are processed. Since classical approach in cases like X8.2 SF event is far from an ideal choice giving unreliable results, application of other numerical techniques, such as e.g. FlareED and easyFit methods, can be beneficial providing needed data.

FlareED and easyFit methods are already proven as efficient both in cases of moderate and high-class SF events (Barta et al., 2022; Kolarski et al., 2023). Together with this two methods, a novel technique employing ML procedures is developed and applied to the same cases of chosen energetic SFs from inspected period 06-10 September 2017, with aim to be tested and verified in conditions of strong solar X-ray radiation. Output results from FlareED and easyFit methods of electron density variations (at the arbitrary reference height of 74 km) covering entire days including cases of energetic SFs from inspected period 06-10 September 2017 are given on the left in Figures 2-6, with soft X-ray flux in violet and electron densities in blue and red for FlareED and easyFit methods, respectively. Results from novel technique relying on ML procedures are presented on the right in Figures 2-6, with soft X-ray flux in violet and electron density variations in green. Both FlareED and easyFit methods are sensitive to X-ray flux variations related to moderate and high-class SF events, with main difference that in case of very energetic SFs, like X9.3 and X8.2, FlareED method gave electron density that went into saturation, compared to output from easyFit method that successfully mimics input X-ray irradiance and gives much higher electron density values as well. Method relying on ML has also proven itself to be an efficient in cases of strong SFs, as other two methods, but also gave some results that went into saturation (especially related to X8.2 and in some sense to X9.3). Comparison between results, in cases of strongest SFs from inspected active period, obtained with applied different numerical methods is given in Table 2.

EasyFit method (Srećković et al., 2021a) gave electron densities higher than FlareED (Srećković et al., 2021b), while ML results are closer to these from FlareED than to easyFit. When compared to LWPC results, it can be said that easyFit method works better than other applied methods, except in the cases of X1.3 where results differ for about an order of magnitude and M7.3 about half an order of magnitude. Unfortunately in the case of M8.2 results from LWPC approach are not reliable, while results from FlareED and ML methods both went into saturation, just as in the case of X9.3, so only available result is from easyFit method. EasyFit gave result for X8.2 of the same order of magnitude but slightly lower than in the case of X9.3. In general, technique relying on ML procedures gives promising results, but also requires some further adjustments in order to provide correct output especially in cases of very strong SF events, like X9.3 and

X8.2. Developing novel numerical techniques that successfully represent lower ionospheric responses especially in cases of very strong SF events is of great importance since classical techniques such as LWPC has model limitations on one hand and VLF parameters are not always favourable on the other.

4. Conclusions and future studies

In this contribution numerical modeling of lower ionospheric response to energetic SF events during September 2017, in relation to subionospheric VLF propagation perturbations with analyzed GQD signal recordings from BEL VLF station, was conducted by employing four different numerical procedures: by classical approach using LWPC software, by approximate FlareED and easy-Fit methods and by novel procedure relying on ML techniques. Focus was on period 06-10 September 2017 covering especially active solar conditions with 4 X-class SFs in range X1.3-X9.3 and also including 2 strongest M-class SFs, of intensity M7.3 and M8.1, from September 2017. Results related to electron density variations obtained by FlareED, easyFit and ML methods are compared to each other and to these from classical approach using LWPC software. In general, when compared with LWPC output electron densities, best results were obtained by easyFit method. ML procedures gave promising results, but further work is necessary in order to provide better results just in cases of very energetic SFs, like these X9.3 and X8.2.

Acknowledgements. This work was funded by the Institute of Physics Belgrade through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia. Authors appreciate comments expressed by referees, which improved this paper.

References

- Arnaut, F., Kolarski, A., & Srećković, V. A., Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. 2023, *Universe*, **9**, DOI: 10.3390/universe9100436
- Barta, V., Natras, R., Srećković, V., et al., Multi-instrumental investigation of the solar flares impact on the ionosphere on 05–06 December 2006. 2022, *Frontiers in Environmental Science*, **10**, DOI: 10.3389/fenvs.2022.904335
- Ferguson, J. 1998, Computer programs for assessment of long-wavelength radio communications, version 2.0: User's guide and source files, Tech. rep., Space and naval warfare systems center San Diego CA
- Grodji, O. D. F., Doumbia, V., Amaechi, P. O., et al., A Study of Solar Flare Effects on the Geomagnetic Field Components during Solar Cycles 23 and 24. 2022, *Atmosphere*, **13**, DOI: 10.3390/atmos13010069
- Grubor, D., Šulić, D., & Žigman, V., Classification of X-ray solar flares regarding their effects on the lower ionosphere electron density profile. 2008, *Annales Geophysicae*, **26**, 1731, DOI: 10.5194/angeo-26-1731-2008

- Hayes, L. A., O'Hara, O. S. D., Murray, S. A., & Gallagher, P. T., Solar Flare Effects on the Earth's Lower Ionosphere. 2021, *Solar Physics*, **296**, 157, DOI: 10.1007/s11207-021-01898-y
- Kelly, M. C. 2009, *The Earth's Ionosphere: Plasma Physics and Electrodynamics, Second Edition*
- Kolarski, A. & Grubor, D., Sensing the Earth's low ionosphere during solar flares using VLF signals and goes solar X-ray data. 2014, *Advances in space research*, **53**, 1595, DOI: 10.1016/j.asr.2014.02.022
- Kolarski, A., Srećković, V. A., & Mijić, Z. R., Monitoring solar activity during 23/24 solar cycle minimum through VLF radio signals. 2022, *Contributions of the Astronomical Observatory Skalnaté Pleso*, **52**, 105, DOI: 10.31577/caosp.2022.52.3.105
- Kolarski, A., Veselinović, N., Srećković, V. A., et al., Impacts of Extreme Space Weather Events on September 6th, 2017 on Ionosphere and Primary Cosmic Rays. 2023, *Remote Sensing*, **15**, 1403, DOI: 10.3390/rs15051403
- Mitra, A. P. 1974, *Ionospheric effects of solar flares* (Springer, Berlin/Heidelberg)
- Rycroft, M., Israelsson, S., & Price, C., The global atmospheric electric circuit, solar activity and climate change. 2000, *Journal of Atmospheric and Solar-Terrestrial Physics*, **62**, 1563, DOI: [https://doi.org/10.1016/S1364-6826\(00\)00112-7](https://doi.org/10.1016/S1364-6826(00)00112-7)
- Silber, I. & Price, C., On the use of VLF narrowband measurements to study the lower ionosphere and the mesosphere–lower thermosphere. 2017, *Surveys in Geophysics*, **38**, 407, DOI: 10.1007/s10712-016-9396-9
- Srećković, V. A., Šulić, D. M., Ignjatović, L., & Vujčić, V., Low Ionosphere under Influence of Strong Solar Radiation: Diagnostics and Modeling. 2021a, *Applied Sciences*, **11**, 7194, DOI: 10.3390/app11167194
- Srećković, V. A., Šulić, D. M., Vujčić, V., Mijić, Z. R., & Ignjatović, L. M., Novel Modelling Approach for Obtaining the Parameters of Low Ionosphere under Extreme Radiation in X-Spectral Range. 2021b, *Applied Sciences*, **11**, 11574, DOI: 10.3390/app112311574
- Šulić, D. M., Srećković, V. A., & Mihajlov, A. A., A study of VLF signals variations associated with the changes of ionization level in the D-region in consequence of solar conditions. 2016, *Advances in Space Research*, **57**, 1029, DOI: 10.1016/j.asr.2015.12.025
- Žigman, V., Dominique, M., Grubor, D., Rodger, C. J., & Clilverd, M. A., Lower-ionosphere electron density and effective recombination coefficients from multi-instrument space observations and ground VLF measurements during solar flares. 2023, *Journal of Atmospheric and Solar-Terrestrial Physics*, **247**, 106074, DOI: 10.1016/j.jastp.2023.106074
- Wait, J. R. & Spies, K. P. 1964, *Characteristics of the Earth-ionosphere waveguide for VLF radio waves* (US Department of Commerce, National Bureau of Standards, Boulder, CO, USA)
- Yasyukevich, Y., Astafyeva, E., Padokhin, A., et al., The 6 September 2017 X-Class Solar Flares and Their Impacts on the Ionosphere, GNSS, and HF Radio Wave Propagation. 2018, *Space Weather*, **16**, 1013, DOI: 10.1029/2018SW001932

Low intensity solar flares' impact: numerical modeling

A. Kolarski^{}, V.A. Srećković^{} and F. Arnaut^{}

*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
11080 Belgrade, Serbia (E-mail: aleksandra.kolarski@ipb.ac.rs)*

Received: September 29, 2023; Accepted: October 26, 2023

Abstract. Solar flares, as strong explosions on the Sun's surface, are well known driving agents that severely affect the near-Earth environment, producing additional ionization within the sunlit Earth's atmospheric layers. X-ray solar flares can be classified regarding their effects on the lower ionosphere and its electron density profile. The focus of this research is on the study of disturbances induced by X-ray solar flares in order to predict the impact of possible weak solar events. In this paper we examined solar activity of lower intensity by conducting numerical modeling using several models and based on data obtained by very low frequency radio signals and from the Geostationary Operational Environmental Satellite (GOES) database on solar X-ray radiation.

Key words: Solar activity – Solar X-ray flares – Radio signal perturbations – GOES

1. Introduction

Solar flares (SFs) are powerful bursts of electromagnetic radiation originating from the Sun's surface [Bothmer et al. \(2007\)](#). Among the categories of SF from A to X-class, B- and C-class SFs represent a low to low-to-moderate level of intensity ([Grubor et al., 2008](#); [Hayes et al., 2021](#)). Although B- and C-class flares are considered to have low level intensity, they still possess the potential to cause disruptions in communication and navigation systems, if they are directed towards the Earth and especially during periods of quiet solar activity such as periods of solar minimum(s). During periods of solar minimum, which refers to the phase of the 11-year solar cycle with minimal solar activity, B- and C-class SFs become significant, as they often constitute the majority of solar flares observed in these periods.

Low class SFs, especially during periods of solar minimum, can induce disturbances in the ionosphere, leading to noticeable effects on terrestrial communication. The primary effect of SFs, in general, on the ionosphere is the augmentation in electron density. The intense X-ray and ultraviolet radiation released during SFs can ionize the upper atmosphere, creating additional free

electrons (Šulić et al., 2016; Curto, 2020; Žigman et al., 2007). These free electrons can affect radio wave propagation by altering the refractive properties of the ionosphere. As a result, the density of the ionosphere increases temporarily, influencing the behavior of radio signals passing through it (Kelly, 2009).

Very Low Frequency (VLF) technology plays a crucial role in studying solar flares and their effects on the lower ionosphere. VLF technology utilizes a range of electromagnetic waves with frequencies from 3 to 30 kHz, with the ability to propagate over long distances without losses and penetrate through both soil and water, making them bounce within the Earth-ionosphere waveguide (Ratcliffe et al., 1972; Silber & Price, 2017). VLF technology finds particular relevance in the study of solar flares due to its ability to capture and analyze the effects of SFs and to indirectly provide data of relevance to ionospheric plasma property variations during these events (Kolarski et al., 2011; Šulić & Srećković, 2014; Srećković et al., 2017; Arnaut et al., 2023). By observing such alterations, valuable insights into the Sun-Earth “connection” can be gained and an enhanced overall understanding of space weather phenomena. The study of SFs of low intensity during solar minimum and their impact on the ionosphere provides valuable insights into the complex interactions between Sun and Earth’s upper atmosphere (Barta et al., 2022; Grodji et al., 2022).

2. Observations

Transit between solar cycle (SC) 23 and 24 implies the period of solar activity is characterized by quiet Sun, i.e., solar minimum. The solar minimum of the 23/24 solar cycle often refers to the period mainly related to years 2008 and 2009, i.e. a period with quiet Sun between the descending branch of the SC 23 and the ascending branch of the SC 24. During this period there were not many SFs reported. More precisely only 8 C class SF events occurred in January, April, November and December and 1 M class SF occurred in March during 2008. Likewise 28 C class SF events occurred in July, September, October and December of 2009. When mutually compared in terms of background solar radiation, year 2008 was a bit calmer with X-ray background flux up to A8.1 compared to 2009 with X-ray background flux up to B1.4, but both of low background solar radiation conditions enabled studying of low class SFs and their influence on Earth’s ionosphere. Results of monitoring solar activity during the 23/24 solar minimum using VLF technology and the European mid-latitude lower ionospheric response to some of the low class SFs from this period based on records obtained by BEL VLF station in Belgrade (Serbia) and estimated electron densities using the Long Wave Propagation Capability (LWPC) methodology can be found in Kolarski et al. (2022a).

In this paper focus is on the utilization of two numerical methods, so called FlareED and easyFit that were developed by Srećković et al. (2021a,b) on the cases of low intensity SFs (lower C- and upper B-class SFs). Moreover the aim

of present paper is to test methods' efficiency and sensitivity when applied to cases of SFs from the lower part of the soft X-ray irradiance spectrum. We note that initially FlareED and easyFit methods were developed for SF events of mid to high intensity (upper C-, M- and lower X-class SFs, e.g. [Srećković et al. \(2021b\)](#); [Kolarski et al. \(2022b\)](#)).

Table 1. Estimated electron densities at nominal height of 74 km by employed different approaches: classical using LWPC, FlarED, easyFit and ML methods.

dd/mm/yyyy	27/10/2009	16/12/2009	23/12/2009
SF class	B8.3	C3.7	C6.4
max X-ray flux time (UT)	13:11	13:02	10:17
Ne LWPC (m^{-3})	$3.79 \cdot 10^8$	$1.37 \cdot 10^9$	$1.17 \cdot 10^9$
Ne easyFit (m^{-3})	$4.20 \cdot 10^8$	$1.01 \cdot 10^9$	$2.33 \cdot 10^9$
Ne FlarED (m^{-3})	$3.47 \cdot 10^8$	$9.64 \cdot 10^8$	$1.64 \cdot 10^9$
Ne ML (m^{-3})	$4.20 \cdot 10^8$	$1.02 \cdot 10^9$	$2.28 \cdot 10^9$

Along with these two methods that were already used for purposes of exploring the influences of stronger SFs on mid-latitude lower ionospheric plasma properties and are used in this study as well, here we also present results from a novel numerical method relying on machine learning (ML) techniques that is applied to low intensity SFs. Machine learning (ML) techniques have demonstrated extensive applicability in various domains, including research, engineering, and industry. Therefore, these techniques offer significant informational value that is not readily identifiable through other statistical methods, thereby offering a novel perspective on the data. One prominent category of widely used algorithms comprises tree, or forest-based techniques, such as Random Forest (RF) ([Breiman, 2001](#)) or XGBoost (XGB) ([Chen & Guestrin, 2016](#)). A typical ML workflow involves the initial preparation of the training and testing datasets, which include both the features and the target variables. Subsequently, the model is trained using the train dataset, following which the evaluation phase can be initiated using the test dataset, i.e., testing the trained model on a test dataset and constructing evaluation metric statistics. Once satisfactory model (hyper)parameters have been obtained, the model can be utilized for predicting the target variable on other datasets. The XGB algorithm was utilized in this study to derive the β and H' parameters (the so called Wait's parameters, [Wait & Spies 1964](#)) (target variables; multi-output ML task) from X-ray irradiance data and other features that were derived from the statistical properties of the X-ray irradiance data (feature data). The XGB model was fine-tuned in relation to the number of trees and the learning rate.

From period 2008 - 2009 several days with reported X-ray solar activity were singled out and inspected in detail: 03 November 2008, 05 July 2009, 27 October 2009, 10 December 2009, 16 December 2009, 21 December 2009 and

23 December 2009. Days were primarily chosen for analysis based on reported X-ray SFs that were “visible” on BEL VLF recordings, in order to provide it possible to compare results related to estimated electron densities obtained by the classical approach (such as [Kolarski et al. 2022a](#)) and by using FlareED and easyFit methods for interpretable cases, on the one hand and to avoid cases of stronger C-class events on the other (like M1.7 on 25 March 2008 with max. X-ray flux at 18:56UT or C7.6 & C7.2 on 18 & 22 December 2009, with max. X-ray flux at 18:55UT & 04:56UT respectively).

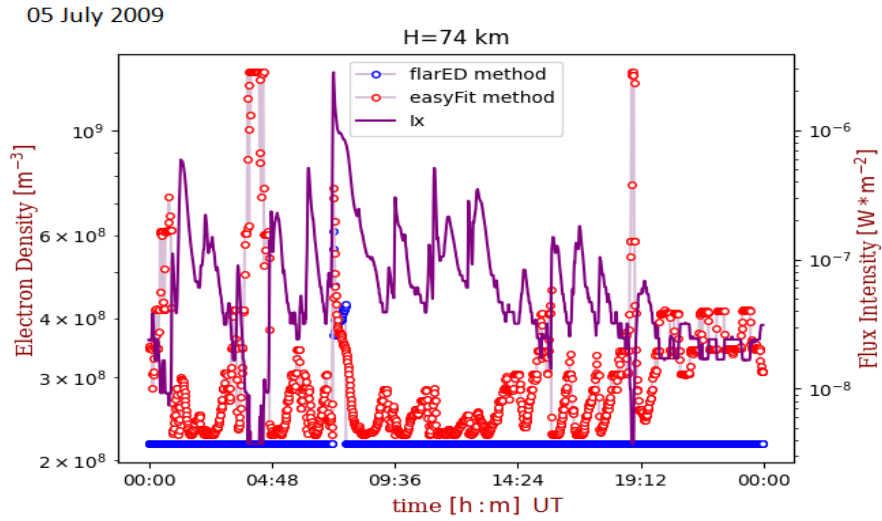


Figure 1. X-ray flux from GOES-15, and N_e ($h = 74$ km) as a function of UT during 05 July 2009. The dashed line with circles shows results acquired by FlarED and easyFit methods.

During 03 November 2008, there were only 3 SFs reported with 2 of them of B-class with intensities in the range B2.4-2.9 and 1 of C-class of intensity C1.6, background flux was of A0.0 and sunspot number was 18. During 05 July 2009, there were 14 SFs reported with 13 of them of B-class with intensities in the range B1-5.9 and 1 of C-class of intensity C2.7, background flux was of A2.2 and sunspot number was 26. During 27 October 2009, there were 23 SFs reported with 18 of them of B-class with intensities in the range B1-8.4 and 5 of C-class with intensities in the range C1.1-1.7, background flux was of A4.2 and sunspot number was 29. During 10 December 2009, there were only 3 SFs reported with 2 of them of B-class with intensities in the range B1.4-6 and 1 of C-class of intensity C3.4, background flux was of A0.0 and sunspot number was 13. During 16 December 2009, there were 9 SFs reported with 6 of them of B-class with intensities in the range B3-8.6 and 3 of C-class with intensities in

the range C1.4-5.3, background flux was of B1.2 and sunspot number was 30. During 21 December 2009, there were 12 SFs reported with 10 of them of B-class with intensities in the range B1.6-8.6 and 2 of C-class with intensities in the range C1-2.5, background flux was of A9.3 and sunspot number was 42. During 23 December 2009, there were 5 SFs reported with 4 of them of B-class with intensities in the range B1.6-8.6 and 1 of C-class of intensity C6.4, background flux was of A7.5 and sunspot number was 23.

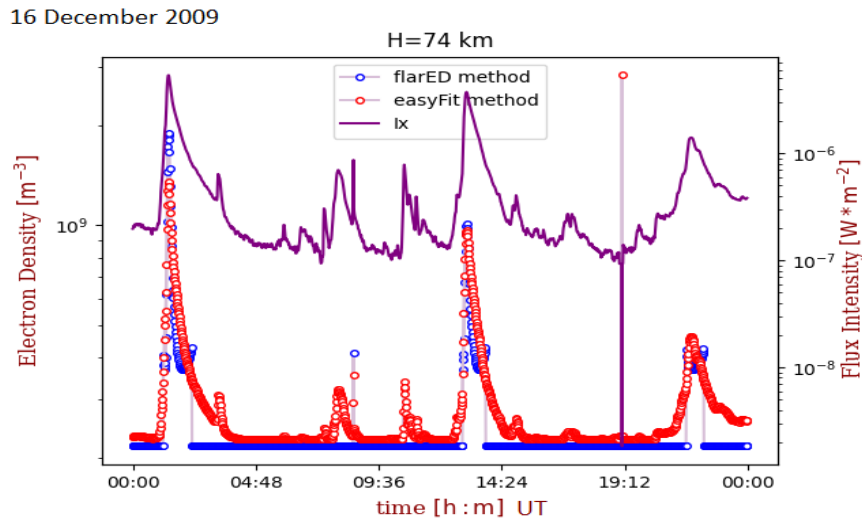


Figure 2. X-ray flux from GOES-15, and N_e ($h = 74$ km) as a function of UT during 16 December 2009. The dashed line with circles shows results acquired by FlarED and easyFit methods.

3. Results and discussion

As previously mentioned, the factor of flare induced perturbation “visibility” on BEL VLF recordings was of prime significance. However, numerical methods FlareED and easyFit were conducted for chosen days in 24h continuity manner, in order to explore these events that were not “visible” in Belgrade due to nocturnal conditions. Figures 1-5 give an overview of the obtained results by application of numerical methods FlareED (blue) and easyFit (red) to the cases chosen for analysis. Aside from all previously mentioned factors, difference between the X-ray flux (violet) characteristic in terms of standalone SF events separated by periods with no activity and numerous back-to-back SFs was subject of interest in the conducted analysis as well. Solar soft X-ray flux is taken from Geostationary Operational Environmental Satellite (GOES) archive database (<https://satdat.ngdc.noaa.gov/sem/goes/data/avg/>).

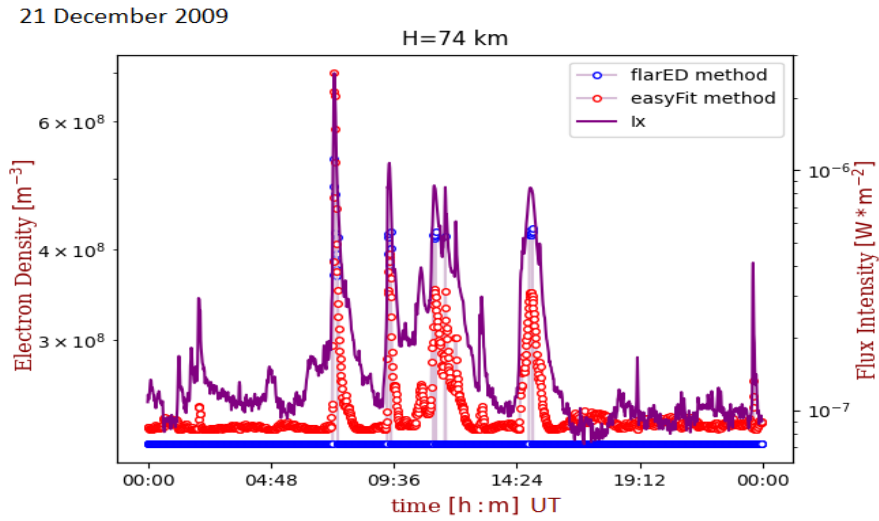


Figure 3. X-ray flux from GOES-15, and Ne ($h = 74$ km) as a function of UT during 21 December 2009. The dashed line with circles shows results acquired by FlarED and easyFit methods.

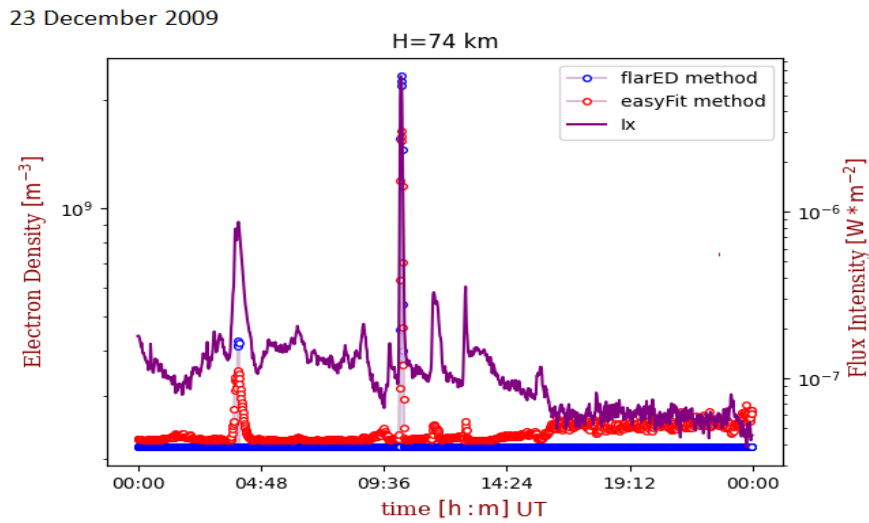


Figure 4. X-ray flux from GOES-15, and Ne ($h = 74$ km) as a function of UT during 23 December 2009. The dashed line with circles shows results acquired by FlarED and easyFit methods.

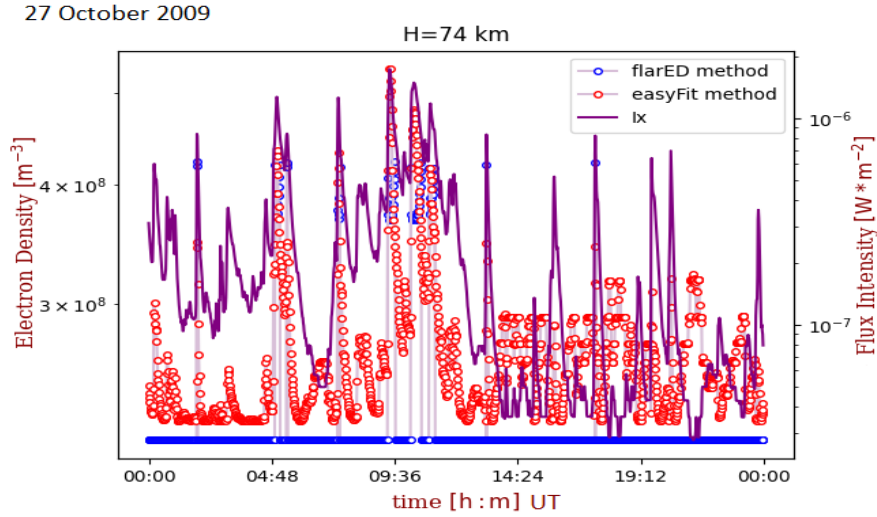


Figure 5. X-ray flux from GOES-15, and Ne ($h = 74$ km) as a function of UT during 27 October 2009. The dashed line with circles shows results acquired by FlareED and easyFit methods.

In general, the easyFit method provided significantly more detailed insight in processed data in terms of sensitivity when applied on cases of low intensity X-ray SFs, in range of low B-class SFs to moderate C-class SFs of B1-C6.4, in terms of mimicking solar soft X-ray flux variation as compared to the FlareED method. FlareED method proven not to be sensitive to very low intensity SFs, giving practically unperturbed parameters as output in such cases. However, in cases of higher intensity SFs, FlareED method gave better results in terms of reaching maximal values in relation to peaks of solar soft X-ray flux variation. On the other hand, both methods gave either some false and/or positive-exaggerated output results. Both methods have proven to be more efficient in cases of more standalone SFs than these of high back-to-back occurrence, with the easyFit method superior in efficiency of recognizing X-ray flux outages, compared to FlareED method that showed no ability at all.

Results regarding electron densities at a nominal height of 74 km, as obtained by the application of classical approach based on NAA/24.0 kHz signal recordings from the Belgrade VLF database in the case of three X-ray SF events of low intensity: a) low intensity C-class solar flare event C3.7 from the 16 December 2009 at 13:02UT, b) moderate intensity C-class solar flare event C6.5 from the 23 December 2009 at 10:17UT and c) moderate intensity B-class solar flare event B8.35 from the 27 October 2009 at 13:11UT (Kolarski *et al.*, 2022a) are compared with the results obtained by application of numerical methods the

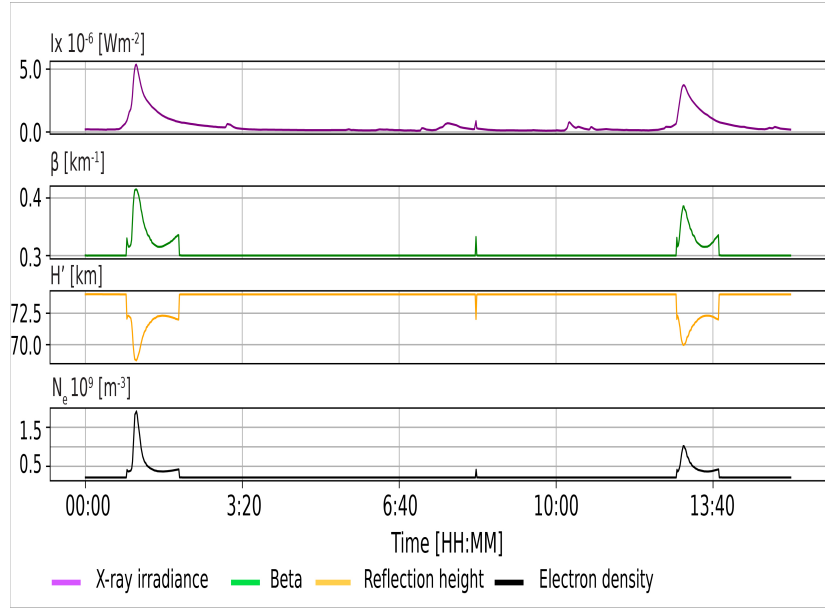


Figure 6. X-ray flux from GOES-15, β and H' parameters and N_e ($h = 74$ km) as a function of UT during 16 December 2009 as results acquired by method relying on ML.

FlareED and the easyFit in the case of these SF events and presented in Table 1. In the case of the B8.3 SF, the lowest in intensity of these three SFs, both methods gave electron density of the same order of magnitude, with the result from the easyFit method slightly overestimated and from the FlarEd method slightly underestimated compared to one from the classical approach, with the result from the FlarEd method closer in absolute value. In the case of C3.7 SF, medium in strength of these three SFs, both methods gave electron density underestimated compared to one from classical approach, with the result obtained from the easyFit method closer in absolute value. In the case of C6.4, the highest in strength of these three SFs, both methods gave electron density of the same order of magnitude, with the result from the easyFit method significantly overestimated and from the FlarEd method much overestimated compared to one from classical approach, with the result from the FlarEd method closer in absolute value. In all the cases, the FlareED method gave the results lower in absolute value compared to corresponding from the easyFit, for approximately 17.4%, 3.7% and 29.6% for these SFs, from lowest to highest in intensity, respectively. The results obtained based on the applied ML technique are in general very close to those obtained from the easyFit numerical method (Table 1), proving that this novel approach is as efficient as other two applied numerical

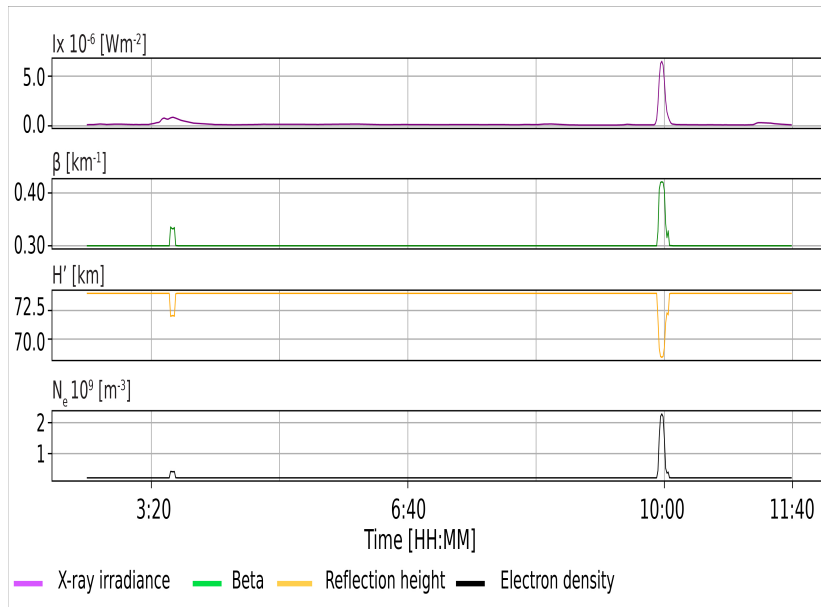


Figure 7. X-ray flux from GOES-15, β and H' parameters and N_e ($h = 74$ km) as a function of UT during 23 December 2009 as results acquired by method relying on ML.

methods in the cases of low intensity SFs. The ML based approach gave the same electron density value in the case of B8.3 SF, slightly higher (for 0.99%) value in the case of C3.7 SF and somewhat lower value (for 2.15%) in the case of C6.4 SF. Figures 6-8 present output results from the applied ML workflow for these three SFs, with X-ray flux in violet, sharpness in green, reflection height in orange and estimated electron density in black.

4. Summary

Overall, B - and C-class SFs occurring during solar minimum have a significant impact on the ionosphere. Impacts on the lower ionosphere of B - and C-class SFs during the 23/24 solar minimum has shed some light on the complex relationship between solar activity and Earth's upper atmosphere. Low class solar flares, although considered notably weaker in comparison to more powerful flares, still have the potential to significantly affect ionospheric dynamics especially during solar minimum periods. Understanding the ionospheric response to low intensity SFs during solar minimum is essential for understanding the potential adverse effects that such events can have on terrestrial communication systems and enhancing our overall space weather prediction capabilities.

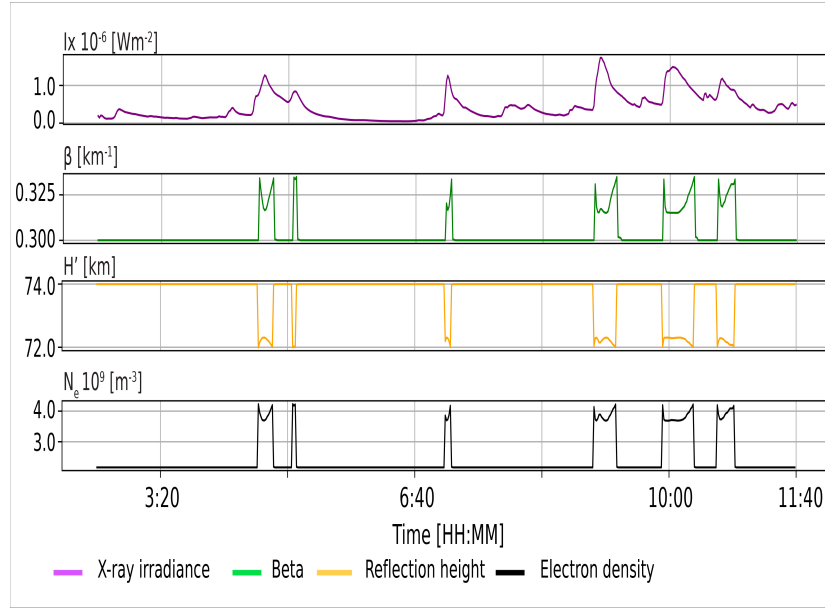


Figure 8. X-ray flux from GOES-15, β and H' parameters and N_e ($h = 74$ km) as a function of UT during 27 October 2009 as results acquired by method relying on ML.

Numerical methods FlareED and easyFit were applied to SFs of low intensity ranging from B1 to C6.4 during seven selected days from period covering solar minimum between SCs 23 and 24, with the aim to test these two methods' efficiency and sensitivity when going through calculations related to SFs of low intensity, mainly from upper B-class and lower C-class solar X-ray flux. Both methods gave satisfactory results in terms of efficiency in cases of events stronger in intensity, with the easyFit method absolutely superior in cases of X-ray flux variations of very low intensity. A novel ML based approach gave results very similar to these obtained by easyFit method, giving great potential for applications in future studies.

When compared with results obtained from the classical approach using the LWPC methodology, in case of three SFs that cover the entire analyzed soft X-ray flux range (B8.3, C3.7 and C6.4) and can be taken as typical examples from their range, results obtained from FlareED method are lower in value compared to ones from easyFit. Compared to classical approach, in case of a) B8.3 SF N_e (easyFit) and N_e (FlareED) are for approximately 11% higher and 9% lower compared to N_e (LWPC), respectively. In case of b) C3.7 N_e (easyFit) and N_e (FlareED) are for approximately 20% lower and 30% lower compared to N_e (LWPC), respectively. In case of c) C6.4 SF N_e (easyFit) and N_e (FlareED) are for approximately 99% higher (almost twice higher) and 40% higher compared to N_e (LWPC), respectively.

The findings from this research can potentially contribute to the development of improved forecasting models, enabling better prediction and preparedness for ionospheric disruptions caused by low class SFs. The study of low class SFs during solar minimum and their impact on the ionosphere highlights the need for continued research and monitoring of such events to enhance our understanding of space weather phenomena and protect technological infrastructure from potential disruptions.

Acknowledgements. This work was funded by the Institute of Physics Belgrade through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia. Authors appreciate comments expressed by referees, which improved this paper.

References

- Arnaut, F., Kolarski, A., & Srećković, V. A., Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. 2023, *Universe*, **9**, DOI: 10.3390/universe9100436
- Barta, V., Natras, R., Srećković, V., et al. 2022, Multi-instrumental investigation of the solar flares impact on the ionosphere occurring in December 2006, Tech. rep., Copernicus Meetings
- Bothmer, V., Daglis, I. A., & Bogdan, T. J., Space Weather: Physics and Effects. 2007, *Physics Today*, **60**, 59, DOI: 10.1063/1.2825074
- Breiman, L., Random forests. 2001, *Machine learning*, **45**, 5
- Chen, T. & Guestrin, C., Xgboost: A scalable tree boosting system. 2016, in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794
- Curto, J. J., Geomagnetic solar flare effects: a review. 2020, *Journal of Space Weather and Space Climate*, **10**, 27, DOI: 10.1051/swsc/2020027
- Grodji, O. D. F., Doumbia, V., Amaechi, P. O., et al., A Study of Solar Flare Effects on the Geomagnetic Field Components during Solar Cycles 23 and 24. 2022, *Atmosphere*, **13**, 69, DOI: 10.3390/atmos13010069
- Grubor, D., Šulić, D., & Žigman, V., Classification of X-ray solar flares regarding their effects on the lower ionosphere electron density profile. 2008, *Annales Geophysicae*, **26**, 1731, DOI: 10.5194/angeo-26-1731-2008
- Hayes, L. A., O'Hara, O. S. D., Murray, S. A., & Gallagher, P. T., Solar Flare Effects on the Earth's Lower Ionosphere. 2021, *Solar Physics*, **296**, 157, DOI: 10.1007/s11207-021-01898-y
- Kelly, M. C. 2009, *The Earth's Ionosphere: Plasma Physics and Electrodynamics, Second Edition*
- Kolarski, A., Grubor, D., & Šulić, D., Diagnostics of the Solar X-Flare Impact on Lower Ionosphere through the VLF-NAA Signal Recordings. 2011, *Open Astronomy*, **20**, 591, DOI: 10.1515/astro-2017-0342

- Kolarski, A., Srećković, V. A., & Mijić, Z. R., Monitoring solar activity during 23/24 solar cycle minimum through VLF radio signals. 2022a, *Contributions of the Astronomical Observatory Skalnaté Pleso*, **52**, 105, DOI: 10.31577/caosp.2022.52.3.105
- Kolarski, A., Srećković, V. A., & Mijić, Z. R., Response of the Earth's Lower Ionosphere to Solar Flares and Lightning-Induced Electron Precipitation Events by Analysis of VLF Signals: Similarities and Differences. 2022b, *Applied Sciences*, **12**, 582, DOI: 10.3390/app12020582
- Ratcliffe, J. A. et al. 1972, *An introduction to ionosphere and magnetosphere* (CUP Archive, Cambridge, UK)
- Silber, I. & Price, C., On the use of VLF narrowband measurements to study the lower ionosphere and the mesosphere-lower thermosphere. 2017, *Surveys in Geophysics*, **38**, 407, DOI: 10.1007/s10712-016-9396-9
- Srećković, V. A., Šulić, D. M., Ignjatović, L., & Vujčić, V., Low Ionosphere under Influence of Strong Solar Radiation: Diagnostics and Modeling. 2021a, *Applied Sciences*, **11**, 7194, DOI: 10.3390/app11167194
- Srećković, V. A., Šulić, D. M., Vujčić, V., Jevremović, D., & Vykyuk, Y., The effects of solar activity: Electrons in the terrestrial lower ionosphere. 2017, *Journal of the Geographical Institute "Jovan Cvijic", SASA*, **67**, 221, DOI: 10.2298/IJGI1703221S
- Srećković, V. A., Šulić, D. M., Vujčić, V., Mijić, Z. R., & Ignjatović, L. M., Novel Modelling Approach for Obtaining the Parameters of Low Ionosphere under Extreme Radiation in X-Spectral Range. 2021b, *Applied Sciences*, **11**, 11574, DOI: 10.3390/app112311574
- Šulić, D. M. & Srećković, V. A., A Comparative Study of Measured Amplitude and Phase Perturbations of VLF and LF Radio Signals Induced by Solar Flares. 2014, *Serbian Astronomical Journal*, **188**, 45, DOI: 10.2298/SAJ1488045S
- Šulić, D. M., Srećković, V. A., & Mihajlov, A. A., A study of VLF signals variations associated with the changes of ionization level in the D-region in consequence of solar conditions. 2016, *Advances in Space Research*, **57**, 1029, DOI: 10.1016/j.asr.2015.12.025
- Wait, J. R. & Spies, K. P. 1964, *Characteristics of the Earth-ionosphere waveguide for VLF radio waves* (US Department of Commerce, National Bureau of Standards, Boulder, CO, USA)
- Žigman, V., Grubor, D., & Šulić, D., D-region electron density evaluated from VLF amplitude time delay during X-ray solar flares. 2007, *Journal of Atmospheric and Solar-Terrestrial Physics*, **69**, 775, DOI: <https://doi.org/10.1016/j.jastp.2007.01.012>

Correlability of Solar Wind with Seismic Events in the Balkan Peninsula Zone

FILIP ARNAUT¹, DEJAN VUČKOVIĆ¹,
IVANA VASILJEVIĆ¹ & VESNA CVETKOV¹

Abstract. The Solar Heliospheric Observatory (SOHO) satellite was launched on the 2nd of December 1995 at L1 Lagrange point (1.5×10^6 km from Earth) with the purpose of gathering data for helioseismology, remote sensing of the solar atmosphere, and solar wind *in situ*. The satellite was positioned into orbit in early 1996, with data acquisition expected to commence on January 20th. The correlation between increased values of solar wind parameters and earthquakes in the Balkan peninsula zone between 1996 and 2018 was made possible by data obtained through continuous proton density and proton velocity monitoring. The assessment of the anomalous threshold was based on statistically determined parameters due to the huge fluctuation of solar wind over time and distinct value increases of proton density and speed. Visual representations of proton density and proton speed were created for the time window preceding each earthquake after defining the boundary between normal and anomalous values. According to the chart analysis, increased proton density occurred in 40 of the 50 cases observed, whereas increased proton velocity appeared in 28 of the 50 cases. Using hypergeometrical probability and an unbiased test with randomly generated parameters, the discovered correlation was statistically verified. A retrospective selection bias analysis is also provided in the research paper.

Key words:

*earthquake prediction,
solar wind, statistical validation
of earthquake precursors.*

Апстракт. Сателит соларне хелиосферске опсерваторије (SOHO) лансиран је ка L1 Лагранжовој тачки (1.5×10^6 km удаљена од Земље) 2. децембра 1995. године са циљем прикупљања података за потребе хелиосеизмологије, даљинске детекције соларне атмосфере и соларног ветра *in situ*. Сателит је доспео у орбиту почетком 1996. године са почетком аквизиције података 20. јануара. Подаци прикупљени континуалним мониторингом густине и брзине протона омогућили су корелацију појава повећаних вредности параметара соларног ветра и потреса на Балканском полуострву у периоду од 1996. до 2018. године. Због велике варијабилности соларног ветра у времену и изразитих скокова густине и брзине протона, одређивање аномалије густине и брзине протона рађено је на основу статистички срачунатих параметара. Након одређивања границе нормална вредност/аномалија, конструисани су графици густине и брзине протона за референтни период пред

¹ University of Belgrade, Faculty of Mining and Geology, Department of Geophysics, Đušina 7, 11000 Belgrade, Serbia.
E-mail: G601-20@rgf.bg.ac.rs

Кључне речи:

*предвиђање потреса,
соларни ветар, статистичка
валидација узрочника потреса.*

сваки потрес. Анализа графика показала је да се појава повећане густине протона десила у 40 од 50 случајева, а повећана вредност брзине протона догодила се 28 од 50 случајева. Добијена корелација је статистички верификована хипергеометријском вероватноћом и независним тестом са насумично генерисаним параметрима. У раду је приказана и анализа ретроспективне пристрасности избора која се манифестује приликом истраживања прекурсора потреса.

Introduction

The term “precursor” refers to a wide variety of physical phenomena that are used for earthquake prediction (CICERONE et al., 2009). Earthquake prediction is a branch of seismology that seeks to predict forthcoming earthquakes in the short, medium, and long term. Deterministic approach in earthquake prediction requires the prior knowledge of the earthquake epicenter’s geographical latitude and longitude, magnitude, and time, which can be seen as unrealistic compared to a probabilistic approach (similar to weather forecasting). The probabilistic approach yields a likelihood that an earthquake will occur in a given region at some time span. Every earthquake prediction method should be based solely on statistics i.e., probability (KAMER et al., 2021).

In order for earthquakes to be predicted with high certainty, the search for new precursors is ongoing still to this day. Today, one group of precursors are astronomical precursors. Modern research showed that there is a strong correlation between solar wind parameters (density, velocity, dynamic proton pressure, etc.) and global earthquakes with a minimum magnitude of M5.6 (MARCHITELLI et al., 2020). A positive correlation was also displayed with earthquakes that occur on a global scale and an increase in proton density (STRASER & CATALDI, 2014). Aside from proton density, vertical Z Earth’s magnetic field component was positively correlated with earthquakes with a magnitude M6.0 (STRASER & CATALDI, 2014).

It is also worth noting that increased solar wind parameter values in the days leading up to the earthquake can’t be considered as a precursor. Earthquake precursors are physical phenomena that occur as a result of an unstable subsurface state

in a particular region. As a result, precursors are the result of accumulated stress. Increases in solar wind parameters can only be considered possible triggers. Increased solar wind parameter values can act as a “straw that broke the camel’s back” by accelerating up an earthquake that would have happened anyhow and was tectonically controlled because of the tectonically accumulated stress (MULARGIA, 1997; MULARGIA, 2001)

Even if the increased solar wind parameters do have an effect on tectonically controlled earthquakes acceleration, it is crucial to understand the modulation they cause. Proposed mechanisms of the Sun’s influence on earthquakes should be discussed to get a better knowledge of those modulations.

SIMPSON (1967) presented the first mechanism, which states that the magnetohydrodynamical interaction of the solar and terrestrial magnetic fields can affect the Earth’s angular velocity. Within that mechanism SIMPSON (1967) states there are two sub-mechanisms that can alter the Earth’s angular velocity, causing earthquakes. The first sub-mechanism states that the continents and oceans are dynamically unstable. Variations in the rotating velocity of the Earth can produce changes in the subsurface stress level (SIMPSON, 1967). The second sub-mechanism states that the Earth’s viscosity prevents it from adapting to the angular velocity change. Inability to adapt to a new state might produce tension in the upper part of the Earth’s crust, causing an earthquake.

SIMPSON’S (1967) second mechanism is similar to MARCHITELLI et al. (2020) reported mechanism in certain ways. Electrical (telluric) currents passing through the subsurface can induce stress pulls due to the reverse piezoelectric effect, which, along with accumulated tectonic stress, can further destabilize

the fault and cause an earthquake. SIMPSON (1967) claimed that telluric currents raise subsurface temperatures, which destabilize faults. However, the subsequent study revealed that this claim was erroneous, as the subsurface temperature increase was two orders of magnitude lower than it should be to induce an unstable subsurface condition. MARCHITELLI et al. (2020) claimed that their mechanism was tectonically controlled, but that an external trigger was used (telluric currents increased by the solar wind).

During the construction of a precursory hypothesis, the main goal is to make sufficient advancement compared to the already existing precursory hypothesis. Because of that, a balance should be made between generalities and particularities (RHOADES & EVISON, 1989; MULARGIA, 1997). Prior mentioned research (particularly MARCHITELLI et al., 2020 and STRASER & CATALDI, 2016) gave a foundation for a statistical analysis of the correlation between solar wind parameters and earthquakes that occur in a region of complex (seismo)tectonic architecture, such as the Balkan peninsula.

Methodology

The CELIAS Proton Monitor (solar wind proton density and velocity data) and the United States Geological Survey (USGS) Earthquake catalog were used in this study. These two databases were used for a period of time ranging from 1996 to 2018.

The proton density and velocity datasets contain values measured every 30 seconds, i.e., a dataset of 1051200 datapoints is regarded full for a year with 365 days, while a dataset of 1054080 datapoints is considered full for a leap year (366) days. The true datapoint number in each dataset was computed, and the results ranged in the order of 90 percent capacity, except for 1998, where the datapoint capacity in the dataset is 59 percent. The explanation for the discrepancy is that the satellite lost control, lost power, and was no longer directed at the Sun from June 24 until October 29. The satellite was put back into operation towards the end of October 1998.

From 1996 to 2018, the USGS website provided an earthquake catalog at a worldwide scale with a minimum magnitude of M5.0. Datasets containing

information about earthquakes in the Balkan peninsula zone (mainly Dinarides, Carpatho-Balkanides, and the Pannonian Basin) were selected from the worldwide datasets. In the Balkan Peninsula region, 52 earthquakes with a minimum magnitude of M5.0 occurred over a 23 year period. Only 50 earthquakes were considered in the final analysis since two earthquakes that occurred on the territory of the Republic of Serbia occurred during a period when the satellite did not acquire any data. Three earthquakes that are thought to have occurred on the border between Albania and Greece were also taken into account. It should also be mentioned that the Republic of Hungary's territory was assumed to span a larger region of the Pannonian basin, despite the fact that no earthquakes with a minimum magnitude of M5.0 occurred on the territory from 1996 to 2018. Aside from the earthquakes already described, three more earthquakes in the Adriatic Sea region have been added.

It is worth noting that the earthquakes were simply filtered from the USGS database by the name of the country in which they occurred. The main goal was to cover as much of the Balkan peninsula as possible, including the Dinarides, Carpatho-Balkanides, and Pannonian Basin, which are three separate geological units in the Balkan Peninsula.

The digital elevation model (DEM: see references NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION, for the internet link) base can be used to map the earthquakes outlined previously (Fig. 1). The DEM utilized was GLOBE, which has a 1 km resolution. Figure 1a shows that the majority of earthquakes occur in Albania (Dinarides) and Romania (Carpatho-Balkanides), accounting for 68 percent of all earthquakes (34/50). The area of the researched surface in relation to the Earth's surface can be observed in figure 1b. The examined area is only 0.41 percent of the overall Earth's size, with a total area of 2 092 846 km².

The analysis was carried out with a total of six steps which will be discussed below:

Step 1: Basic statistical calculation for obtaining the information about the mean value, standard deviation, minimum, and maximum values for the yearly dataset. This type of data provides a foundational understanding of the dataset.

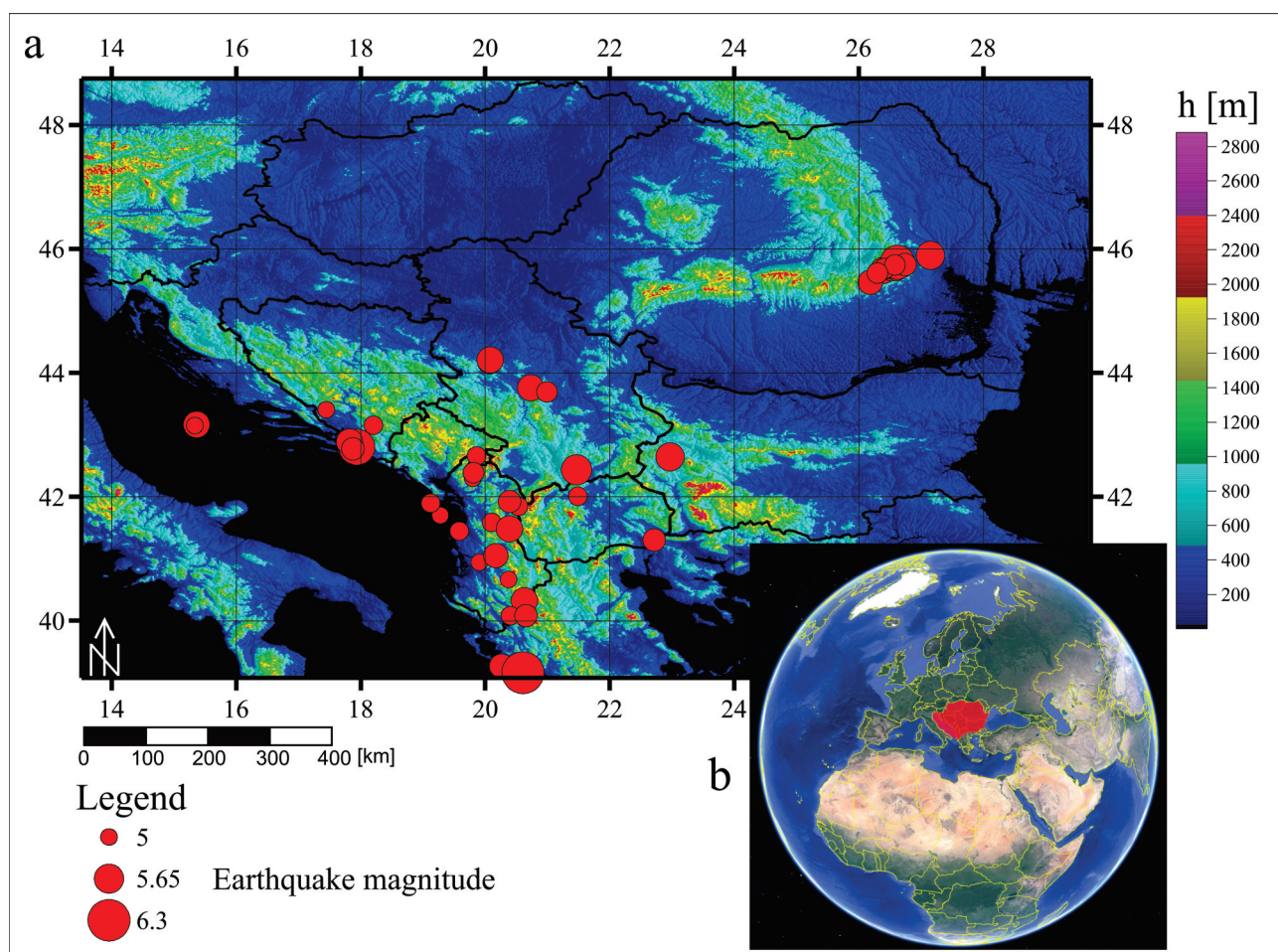


Fig. 1. a. Location of earthquakes with a DEM base; b. Size comparison of the investigated area and Earth's area (modified after: <https://earth.google.com/web/>)

Step 2: Determination of the anomaly threshold, which is the line between what is considered normal and what is considered anomalous. This is done by calculating skewness and kurtosis (and two standard errors of skewness and kurtosis), as well as basing the anomalous threshold on n standard deviations. According to the “33/67/99.7” rule (empirical rule used to estimate the percentage of data that falls within one, two, or three standard deviations), three standard deviations for normal distributions can be considered anomalous since they correspond with only 0.3 percent of data. The n is greater for distributions with larger skewness. This stage calculates the percentage of data that is equal to or greater than the anomaly threshold; this information is used to check the quality of the determined anomaly threshold. That value should not exceed 1% on average throughout the whole 23 year sample.

Step 3: Construction of graphic representation (charts) for the two weeks leading up to each earthquake.

Step 4: Identifying which earthquakes have a proton density (velocity) anomaly in the two-week, one-week, and four-day periods prior to the earthquake.

Step 5: Hypergeometrical probability based statistical significance test. In this stage, the number of unique anomalous days for each year is calculated, and the hypergeometrical probability is calculated to see if the anomaly occurred more frequently than might randomly. In addition, an independent test is run to corroborate the hypergeometrical probability as a statistical significance test.

Step 6: Analysis and interpretation of the results from the previous five steps. The correlation is thought not to be significant, i.e., can be thought to

be random, if the probability obtained by the hypergeometrical probability coincides in the range of 10% with the number of times analyzed earthquakes had an anomalous proton density (velocity) value in the time window chosen.

Results and Discussion

In a yearly format, Table 1 shows general statistical information of the proton density parameter. The mean value in the database (from 1996 to 2018) shows a decreasing trend until 2008, after which there is a modest increase until 2015, after which the mean yearly proton density decreases again (Fig. 2, middle). The average number of protons per cubic centimeter over the entire dataset is 5.3.

Table 1. General statistical information for the proton density dataset.

Year	Mean [cm^{-3}]	Standard deviation [cm^{-3}]	Minimum [cm^{-3}]	Maximum [cm^{-3}]
1996	7.82	4.43	-1	53.01
1997	8.65	4.96	-1	57.65
1998	7.67	4.61	-1	66.22
1999	6.35	4.52	-1	57.28
2000	6.31	4.82	-1	70.18
2001	6.50	5.38	-1	120.77
2002	6.96	5.30	-1	70.56
2003	5.38	3.98	-1	72.19
2004	4.97	3.87	-1	111.81
2005	4.95	4.58	-1	88.56
2006	4.68	4.24	-1	75.65
2007	4.06	3.99	-1	82.66
2008	3.74	3.08	-1	72.44
2009	4.03	3.14	-1	71.32
2010	4.35	3.53	-1	117.17
2011	4.33	3.87	0	90.81
2012	4.57	4.02	0	118.02
2013	4.45	4.00	0.01	126.92
2014	4.81	3.49	0.06	60.29
2015	5.27	4.43	0.07	77.81
2016	4.55	3.54	0.09	88.19
2017	3.94	3.14	0.06	63.2
2018	3.66	2.86	0.04	62.06
Average	5.30	4.08	/	81.51

With the exception of 2018, which has a standard deviation of 2.86 cm^{-3} , standard deviation values are consistent, ranging from 3 to 5 protons per cubic centimeter. The overall dataset's mean standard deviation is 4.08 cm^{-3} .

For the time span 1996 to 2010, the minimum values are -1 . This is not a representation of the measured proton density, but rather a lack of data (mainly of technical nature). Minimum values after 2010 are either 0 or a number greater than 0. Values less than 0 were filtered for the period up to 2010.

The maximum values in each dataset range from 53 (1996) to 127 (2013) protons per cubic centimeter. With two maximums (2003 and 2012) and three minimums (1996, 2007, and 2018), there is some type of pattern in the maximum levels (Fig. 2, bottom).

Because of the displayed periodicity of the mean and maximal proton density values, a correlation with the number of Sunspots in the same period was performed (Fig. 2, top). To visually represent the Sunspot number, a database of Sunspot numbers was taken (see references SILSO data, Royal Observatory of Belgium, Brussels for the internet link) and a local regression (LOESS) filter was applied.

For the period 1996 to 2018, the sunspot chart (Fig. 2, top) shows three minimums and two maximums, indicating two cycles (Solar cycles 23 and 24). Minimum values can be recorded in 1996, 2008, and 2018 (on average every 11 years), whereas maximum values are also located at an 11 year time period.

Two correlation coefficients, Pearson's and Spearman's, were calculated to assess the correlation between these three factors. Pearson's correlation coefficient is a traditional correlation coefficient that can occasionally reveal flaws, such as when one data point appears to be outside the general trend

of other data points. Pearson's correlation coefficient will yield a lower value than the true correlation coefficient (SCHÖBER et al., 2018). To acquire more detailed information, both correlation coefficients were calculated, and a description and a color based on the

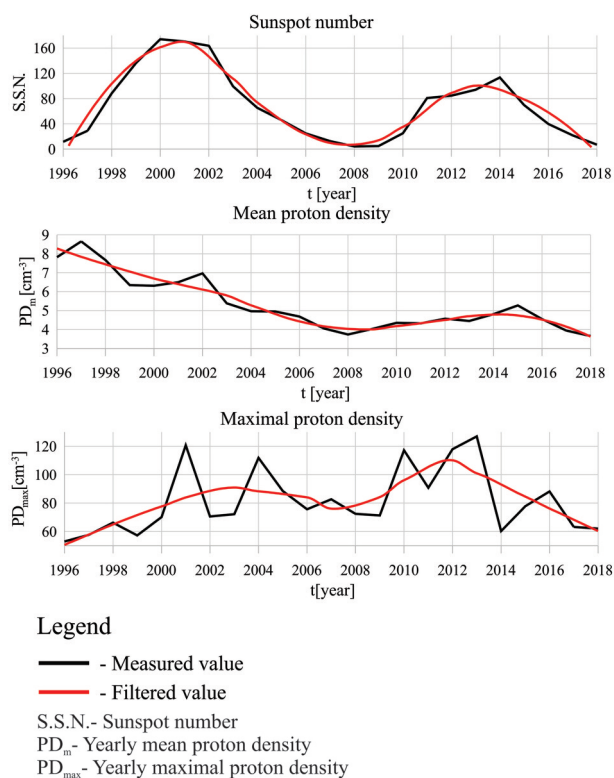


Fig. 2. Sunspot number (top) correlation with yearly mean proton density values (middle); Yearly maximal proton density values (bottom)

value range were assigned, as reported by SCHOBER et al. (2018).

Table 2 shows that maximum proton density values have a weak correlation with the Sunspot number calculated using both Pearson’s and Spearman’s correlation coefficients. On the other hand, the Pearson correlation coefficient shows a weak correlation with annual mean proton density values, whereas

Table 2. Correlation coefficients for the solar wind parameters and Sunspot number.

/	Pearson	Spearman	Description	Color	Range
				Negligible	Red
PD _m	0.39	0.57	Weak	Blue	0.1-0.39
			Moderate	Yellow	0.4-0.69
			Strong	Green	0.7-0.89
PD _{max}	0.16	0.12	Very strong	Dark Green	0.9-1.0

the Spearman correlation coefficient shows a moderate value.

To determine the anomalous threshold, the skewness, kurtosis, and two standard errors of skewness and kurtosis must be calculated first (table 3). From 1.76 in 1996 to 5.34 in 2007, the skewness parameter shows a wide range of fluctuation. Skewness takes an average 3.37 value over the course of the observed period. Because the skewness parameter informs us about the “length of the tail” of the distribution, the skewness is minimal for the lowest maximum value of yearly proton density. The contrary is not true, skewness is not maximal for the highest maximum values in the dataset (2013), but it does take values of around 4. The kurtosis parameter has a wide range of values, ranging from 5.32 in 1996 to 51.58 in 2007. Kurtosis has an average value of 24.31.

Table 3. Skewness, kurtosis and two standard errors of skewness and kurtosis for the investigation period.

Year	Skewness	Kurtosis	Year	Skewness	Kurtosis
1996	1.76	5.32	2008	3.73	25.63
1997	2.18	8.56	2009	3.36	23.79
1998	1.89	6.49	2010	4.36	50.29
1999	2.15	7.71	2011	4.30	34.65
2000	2.69	11.87	2012	4.09	40.47
2001	2.98	15.03	2013	3.99	44.35
2002	2.80	13.26	2014	2.52	12.12
2003	3.76	27.41	2015	3.53	20.75
2004	3.38	31.06	2016	3.36	19.82
2005	3.64	23.09	2017	4.14	32.74
2006	3.61	21.79	2018	3.89	31.25
2007	5.34	51.58	Average	3.37	24.31

Reference Skewness Kurtosis
1.02 2.04

Average values of skewness and kurtosis are higher than the approximated two standard error values of skewness and kurtosis for the investigated period. This information leads us to the conclusion that the proton density distribution is highly horizontally and vertically distorted, as was verified graphically in figure 3. Because of that, higher multiplicities of standard deviations can be used as the anomaly threshold for each year. For this research, five standard deviations will be used as the anomaly threshold. The anomaly threshold for each year will thus be calculated as:

$$Pd_{anomaly}(y) = PD_m(y) + 5SD(y) \quad (1)$$

where:

$PD_{\text{anomaly}}(y)$ - anomaly threshold for each year,

$PD_m(y)$ - mean proton density value for each year,

$SD(y)$ - standard deviation calculated for each year.

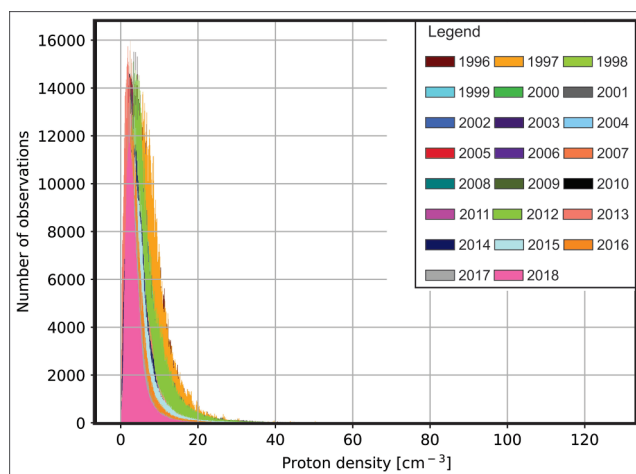


Fig. 3. Proton density histogram for the period from 1996 to 2018

It is possible to calculate the anomaly threshold and the percentage of data that is equal to or greater than the anomaly threshold using the previously displayed data. The maximal anomaly threshold can be obtained for the year with a comparatively high yearly mean proton density and high standard deviation, since the anomaly threshold is solely determined by the mean yearly proton density value and the standard deviation value. The anomaly threshold values range from 17.97 protons per cubic centimeter (2018) to 33.46 protons per cubic centimeter (2002), as shown in table 4. The overall dataset's mean anomaly threshold value is 25.69 protons per cubic centimeter. The percentage of data that is equal to or higher than the anomaly threshold varies between 0.24 percent (1996) to 0.67 percent (2015). The average percentage of data over the anomaly threshold is 0.49 percent, which is lower than the expected value of 1%.

Table 4. Anomaly threshold information for the proton density parameter.

Year	Mean [cm ⁻³]	Standard deviation [cm ³]	5 σ [cm ⁻³]	Maximum [cm ⁻³]	Percentage of data $\geq 5\sigma$ [%]
1996	7.82	4.43	29.98	53.01	0.24
1997	8.65	4.96	33.43	57.65	0.37
1998	7.67	4.61	30.72	66.22	0.26
1999	6.35	4.52	28.93	57.28	0.41
2000	6.31	4.82	30.42	70.18	0.45
2001	6.50	5.38	33.40	120.77	0.56
2002	6.96	5.30	33.46	70.56	0.52
2003	5.38	3.98	25.30	72.19	0.47
2004	4.97	3.87	24.31	111.81	0.42
2005	4.95	4.58	27.87	88.56	0.63
2006	4.68	4.24	25.87	75.65	0.58
2007	4.06	3.99	24.03	82.66	0.62
2008	3.74	3.08	19.14	72.44	0.61
2009	4.03	3.14	19.74	71.32	0.45
2010	4.35	3.53	21.98	117.17	0.48
2011	4.33	3.87	23.66	90.81	0.65
2012	4.57	4.02	24.65	118.02	0.52
2013	4.45	4.00	24.45	126.92	0.40
2014	4.81	3.49	22.26	60.29	0.47
2015	5.27	4.43	27.43	77.81	0.67
2016	4.55	3.54	22.26	88.19	0.58
2017	3.94	3.14	19.66	63.2	0.54
2018	3.66	2.86	17.97	62.06	0.48
Average	5.30	4.08	25.69	81.51	0.49

Table 5. Anomalous proton density occurrences in the time period before every earthquake.

Earthquake location	Year	Month	Day	Anomaly[14 days]	Anomaly [7 days]	Anomaly [4 days]
Serbia	2010	November	3	•	•	•
	2002	April	24	•	•	
	1999	July	1	•	•	
	1999	April	30	•		
Bosnia and Herzegovina	2005	September	27	•	•	•
	2004	May	23	•	•	•
	1996	September	5			
Croatia	1996	September	17	•	•	
	1996	September	9			
Montenegro	2018	January	4	•	•	•
	2014	December	29	•	•	•
Albania	2018	August	11	•	•	•
	2018	July	4	•	•	•
	2014	May	19	•		
	2009	September	6	•	•	•
	2009	August	21	•	•	•
	2007	June	29	•	•	•
	2006	October	19	•		
	2006	June	21			
	2005	July	10	•	•	•
	2004	November	23	•	•	•
	2004	April	7	•	•	•
	2003	August	14			
	2001	April	9	•		
	1999	December	22			
	1997	May	16	•	•	•
1996	August	20	•			
1996	August	5	•			
1996	July	26				
North Macedonia	2016	September	11			
	2009	May	24	•	•	•
Bulgaria	2012	May	22	•	•	•
Romania	2018	October	28	•	•	
	2016	December	27	•		
	2016	September	23	•	•	
	2014	November	22	•		
	2013	October	6	•	•	•
	2009	April	25	•	•	•
	2005	June	18	•	•	
	2005	May	14			
	2004	October	27	•	•	•
	2002	November	30	•		
	2001	July	20			
	2001	May	24	•	•	•
	2000	April	6			
	1999	April	28	•		
1998	March	13	•	•	•	
Adriatic sea	2004	November	25	•	•	
	2003	March	29	•	•	•
	2003	March	27	•	•	•
Sum •/50 earthquakes				40	31	23
Sum • [%]				80	62	46

After calculating the anomaly threshold, determining whether earthquakes had anomalous proton density values in the two weeks, one-week, and four-day periods leading up to it was done. Table 5 shows the individual earthquakes examined in this study, as well as the occurrence of the anomalous proton density value in the time interval before it denoted by the symbol „●“.

Table 5 shows that in the two-week period leading up to an earthquake, 80 percent of earthquakes (40/50) exhibit anomalous proton density values. This number lowers to 62 percent, or 31 out of 50 earthquakes, in the week leading up to an earthquake. For the four days leading up to each analyzed earthquake, 23/50 (46 percent) showed a proton density anomaly.

The first step in determining the statistical significance of these findings is to establish the number of individual anomalous days in a calendar year. To begin, a definition of an anomalous day should be established:

30 seconds and another with a duration of several hours are not the same. In keeping with the survey's generalities, it is considered that every anomaly has the same relevance, regardless of its duration.

After determining the number of individual anomalous days for each year in the dataset, the hypergeometric probability can be calculated, which is the probability that in a year with 365 days and, for example, 22 individual anomalous days, at least one anomalous day can be found in the chosen time window (in this case 14, 7 and 4 days prior to each earthquake). This probability shows the likelihood of finding a proton density anomaly in the time window chosen before each earthquake at random.

To validate the hypergeometric probability accuracy, or whether it is a good probability model for probability computation, an independent test was created with the same input parameters as the hypergeometric probability, with the exception that all of the parameters are picked at random. Figure 4 depicts the flowchart for such a test.

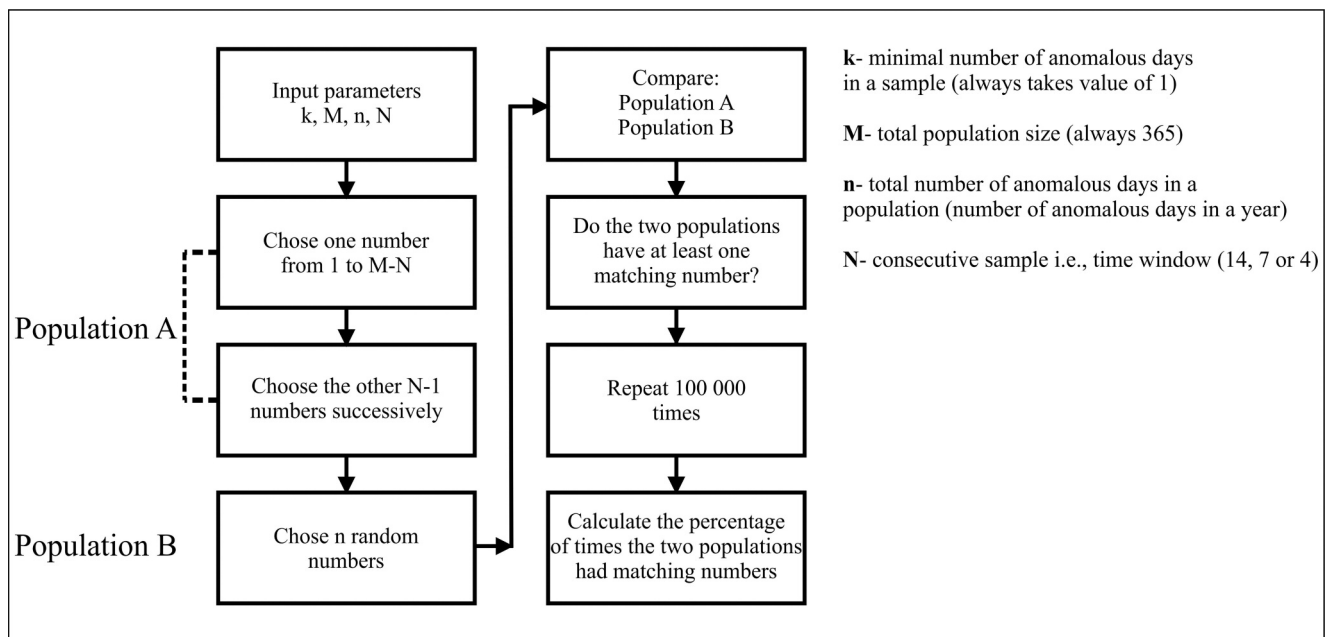


Fig 4. Flowchart of the independent test

“Every day with at least one proton density measurement higher than the anomaly threshold value for a particular year is an anomalous day.”

This notion of an anomalous day has some disadvantages. The first is that it ignores the duration of the anomaly, i.e., one anomaly with a duration of

There are eight steps in the flow of the test probability. The first step is to choose a number at random from 1 to M-N (i.e., 365-14/7/4). The remaining N-1 numbers are calculated by adding to the first number +1, +2, +3, etc. This group of numbers is referred to as “Population A,” and it repre-

sents the time period preceding the earthquake. The number of anomalous days in a year is represented by Population B, which is picked at random from 1 to 365 integers. After both Population A and Population B have been determined, the two populations can be compared to see if they have the same numbers. After iterating the method 100 000 times, the percentage of times the two populations had at least one matching number was calculated.

This percentage is an approximation of the hypergeometric probability and is used to check the quality of the probabilities obtained. Table 6 shows the hypergeometric probability, the test probability based on the previously shown algorithm, and the difference between these two probabilities for the time windows of 14, 7, and 4 days before the earthquake.

Table 6 shows that the average hypergeometric probability values for the time windows of 14, 7 and 4 days are 82.52 percent, 59.11 percent, and 40.32 percent, respectively. These probabilities are in line with the algorithm’s probability, which is in the 2 percent range (maximal discrepancy being 2.26 percent for the time window of 7 days). Even if the difference between the two probabilities is small, it can be explained in two ways.

The first reason for the disparity is that as the number of algorithm iterations approaches infinity, the algorithm probability tends to hypergeometric probability. By increasing the number of iterations, this effect can be reduced.

The hypergeometric probability is a cumulative probability of having all anomalous days in a time window sample, which is the second cause of the tiny disparity between the hypergeometric probability and the algorithm probability. Even though the chance of having more than two anomalous days in the time sample is small, it is accounted for in the hypergeometric probability.

With the two explanations in mind, it is easy to see why there’s such a modest difference between the hypergeometric and algorithm probabilities. As a result, the hypergeometric probability model is considered to be a satisfactory probability model for this study.

The obtained hypergeometric probability was compared to the calculated percentage of proton density anomaly occurrence for the time window chosen before the analyzed earthquakes on the Balkan peninsula zone (table 7). If the two values are within 10% of each other, any occurrence of proton density anomaly in the time window preceding an earthquake is regarded to be random. As a result, there is no statistically significant correlation between the two events.

The same method was applied for the proton velocity parameter, with anomalous proton velocity values found in 56 percent (28/50) of earthquakes in the two-week time frame prior to it, 32 percent (16/50) in the one-week time window, and 18 percent (9/50) in the four-day time window. Similar to

Table 6. Hypergeometric probability, algorithm probability and difference for the proton density parameter.

General		14- Days			7- Days			4- Days		
Year	Number of anomalous days	H. Probability [%]	A. Probability [%]	Difference [%]	H. Probability [%]	A. Probability [%]	Difference [%]	H. Probability [%]	A. Probability [%]	Difference [%]
1996	22	58.80	57.79	1.01	35.53	34.71	0.82	22.10	21.54	0.56
1997	49	87.23	85.23	1.99	63.87	61.34	2.53	43.96	41.62	2.34
1998	21	57.05	56.00	1.05	34.19	33.61	0.58	21.18	20.69	0.49
1999	29	69.31	67.731	1.57	44.26	42.96	1.30	28.29	27.46	0.84
2000	37	78.24	76.70	1.54	52.99	51.01	1.98	34.91	33.81	1.10
2001	42	82.53	80.80	1.73	57.82	55.65	2.18	38.81	37.17	1.64
2002	38	79.17	77.31	1.86	53.99	52.27	1.73	35.70	34.46	1.24
2003	31	71.81	70.74	1.07	46.56	45.25	1.31	29.99	29.04	0.95
2004	44	84.01	81.99	2.02	59.63	57.10	2.54	40.32	38.56	1.75
2005	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2006	42	82.53	80.80	1.73	57.82	55.65	2.18	38.81	37.17	1.64
2007	46	85.38	83.41	1.97	61.38	58.82	2.56	41.80	39.87	1.93
2008	51	88.34	86.50	1.84	65.46	62.85	2.61	45.38	43.31	2.06
2009	56	90.73	88.97	1.76	69.16	66.23	2.94	48.79	46.11	2.68
2010	44	84.01	81.99	2.02	59.63	57.10	2.54	40.32	38.56	1.75
2011	46	85.38	83.41	1.97	61.38	58.82	2.56	41.80	39.87	1.93
2012	60	92.31	90.38	1.93	71.88	68.81	3.07	51.40	48.27	3.13
2013	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2014	47	86.02	84.33	1.69	62.23	60.09	2.14	42.53	40.51	2.01
2015	46	85.38	83.41	1.97	61.38	58.82	2.56	41.80	39.87	1.93
2016	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2017	52	88.86	86.78	2.08	66.23	63.63	2.60	46.07	43.74	2.33
2018	66	94.21	92.59	1.62	75.56	72.21	3.36	55.13	51.99	3.15
Average	44.57	82.52	80.75	1.77	59.11	56.86	2.26	40.32	38.47	1.85

Table 7. Comparison of the hypergeometric probability and the results obtained by the number of proton density occurrences before earthquakes in the Balkan peninsula zone.

Year	Anomaly probability [14 days] [%]	Anomaly probability [7 days] [%]	Anomaly probability [4 days] [%]
1996	58.80	35.53	22.10
1997	87.23	63.87	43.96
1998	57.05	34.19	21.18
1999	69.31	44.26	28.29
2000	78.24	52.99	34.91
2001	82.53	57.82	38.81
2002	79.17	53.99	35.70
2003	71.81	46.56	29.99
2004	84.01	59.63	40.32
2005	88.86	66.23	46.07
2006	82.53	57.82	38.81
2007	85.38	61.38	41.80
2008	88.34	65.46	45.38
2009	90.73	69.16	48.79
2010	84.01	59.63	40.32
2011	85.38	61.38	41.80
2012	92.31	71.88	51.40
2013	88.86	66.23	46.07
2014	86.02	62.23	42.53
2015	85.38	61.38	41.80
2016	88.86	66.23	46.07
2017	88.86	66.23	46.07
2018	94.21	75.56	55.13
Average:	82.52	59.11	40.32
Calculated:	80	62	46

the proton density case, no statistically significant correlation was discovered. The last possibility for correlation remaining is to check if the proton velocity parameter was a small subset of days with both the density and velocity anomalies. Only 66 days (0.7 percent of the dataset) showed both density and velocity anomalies across the 23 year period. Only 12% (6/50) had an anomaly over the two-week period, 4% (2/50) during the one week period, and 2% (1/50) during the four-day period preceding the earthquake. There was no statistically significant correlation found.

Increased Anomaly Threshold Test

Because of the skewness of the proton density distribution, bigger n standard deviation multipli-

ties could be used. The increase from five to nine standard deviations resulted in an average increase to 42 protons per cubic centimeter throughout the entire sample. This increase is 40% greater than the prior anomaly threshold, however, it has resulted in a 75% reduction in the number of anomalous days. An average of 11 days each year were marked as anomalous during the investigation period.

Only 28% (14/50) had an anomaly in the two-week period, 22% (11/50) in the one-week period, and 12% (6/50) in the four-day period before the earthquake when the entire sequence was repeated as previously shown for the proton density parameter. The predicted values were derived by recalculating the hypergeometric probability with the updated values (33 percent for the two-week pe-

riod, 18.67 percent for the one-week period, and 11 percent for the four-day period). As can be observed, the expected values derived using hypergeometric probability and the calculated values are within 10% of one another, which is enough to rule out any statistically significant correlation even with higher proton density values.

The proton velocity parameter was treated in the same way, with the n standard deviations increased from three to three and a half. This was a 6 percent increase, but it resulted in a 47 percent reduction in the number of anomalous days. In the two-week period, 28 percent (14/50) had an anomalous value, 14 percent (7/50) had an anomalous value in the one-week period, and 6 percent (3/50) had an anomalous value in the four-day period before an earthquake. There was no statistically significant correlation found.

It is worth noting that even when the dynamical anomaly threshold is disregarded (the threshold is based on the mean yearly proton density/velocity value and the standard deviation for the year) and an exceptionally high and constant anomaly threshold is applied (100 cm^{-3} for proton density and 1000 km/s for proton velocity), only six individual anomalous days for the proton density are observed and 14 for the proton velocity parameter. Those two figures, respectively, represent 0.07 percent and 0.1 percent of the overall data set. Only one earthquake (Albania, April 9th, 2001) displayed a proton density anomaly prior to the earthquake, while none had a proton velocity anomaly. There was no statistically significant association, as with the other examples previously shown.

Conclusion

A statistical correlation between increased solar wind parameters (proton density and velocity) and earthquakes on the Balkan Peninsula zone between 1996 and 2018 was presented in this study paper.

In the two weeks leading up to an earthquake, the presented increased proton density parameters show an 80% (40/50) correlation. This is the highest correlation found in this study. There is no substantial statistical correlation between these two occurrences on the Balkan Peninsula zone, according to statistical verification.

Proton velocity, like proton density, had no statistically significant correlation with earthquakes on the Balkan Peninsula, even when they occurred 56 percent of the time prior to earthquakes. A small group of days with both proton density and velocity anomalies produced no statistically significant correlation when studied.

Increasing the anomaly threshold for the proton density parameter from five to nine standard deviations and for the proton velocity parameter from three to three and a half standard deviations reduced the number of anomalous days by 75 and 47 percent, respectively. This reduction in the number of anomalous days did not result in a statistically meaningful improvement in the association. Furthermore, ignoring a dynamical anomaly threshold and concentrat-

ing on rare occurrences (100 cm^{-3} and 1000 km/s) did not produce any significant results.

In summarizing the subjectivities in this study, the primary focus must be on the selection of the research area. MARCHITELLI et al. (2020) found a positive association on a worldwide scale, however, this is of no practical utility because knowing when an earthquake will occur without knowing where it will occur is useless. The focus of this research article was on a smaller scale with a complex geological setting, however, the subjectivity involved in selecting such a location could not be adequately quantified.

The second subjectivity involves expanding the time span of 6.9 days stated by STRASER & CATALDI (2014). As the time window is extended, days that are more anomalous will surely occur before an earthquake. This subjectivity, like the preceding one, cannot be measured, but it is thought that statistical verification will eliminate it.

The last subjectivity is that when kurtosis increases, the value of variance decreases (TABACHNIK et al. 2007), and so the value of standard deviation decreases. If the standard deviation is less than the true value, the anomalous threshold will be less than the true value as well. Although this subjectivity is not quantifiable, it is eliminated by the use of the test of increasing anomalous thresholds.

Even when viewing this research topic with a critical mindset, it is believed that subjectivities in this study did not have a substantial, if any, influence. GELLER et al. (1997) also express skepticism regarding the discovery of a new, major precursor, stating that the probability of discovering the next one decreases with each new attempt, and that the possibility of discovering a novel precursor is now exceedingly minimal.

To locate a novel precursor candidate, you must first have a well-understood and well-established mechanism for such an occurrence. There is no documented and verified mechanism that could explain the solar wind generating earthquakes phenomenon, for example. The earthquakes are assumed to be tectonically controlled, but the increase in solar wind parameters acts as a trigger.

Even though the presented research found no statistically significant association for the Balkan Peninsula, it does not rule out the positive correlation found by MARCHITELLI et al. (2020) on a worldwide

scale. To establish it as a viable method, first, the mechanism should be established, then the usage of a statistical method to demonstrate the statistical significance of such a method with high confidences and low false alarm values should be undertaken. This approach has yet to be established for any of the known precursors.

References

- CATALDI, G., CATALDI, D. & STRASER, V. 2016. Solar activity correlated to the M7.0 Japan earthquake occurred on april 15, 2016. *New Concepts in Global Tectonics Journal*, 4 (2): 279–285.
- CHARGE, ELEMENT, AND ISOTOPE ANALYSIS SYSTEM, Retrieved from <https://11.umd.edu/>, [1. 12. 2020]
- CICERONE, R.D., EBEL, J.E. & BRITTON, J. 2009. A systematic compilation of earthquake precursors. *Tectonophysics*. 476 (3–4): 371–396.
- GELLER, R.J., JACKSON, D.D., KAGAN, Y.Y. & MULARGIA, F. 1997. Earthquakes cannot be predicted. *Science*. 275(5306): 616–1616.
- GOOGLE EARTH, Retrieved from <https://earth.google.com/web/>, [7. 4. 2021]
- KAMER, Y., NANDAN, S., OUILLO, G., HIEMER, S. & SORNETTE, D. 2021. Democratizing earthquake predictability research: introducing the RichterX platform. *The European Physical Journal Special Topics*, 230 (1): 451–471.
- MARCHITELLI, V., HARABAGLIA, P., TROISE, C. & DE NATALE, G. 2020. On the correlation between solar activity and large earthquakes worldwide. *Scientific reports*, 10 (1): 1–10.
- MULARGIA, F. 1997. Retrospective validation of the time association of precursors. *Geophysical Journal International*, 131 (3): 500–504.
- MULARGIA, F. 2001. Retrospective selection bias (or the benefit of hindsight). *Geophysical Journal International*. 146 (2): 489–496.
- RHOADES, D.A. & EVISON, F.F. 1989. On the reliability of precursors. *Physics of the Earth and Planetary interiors*, 58 (2–3): 137–140.
- NATIONAL CENTERS FOR ENVIRONMENTAL INFORMATION, Retrieved from <https://www.ngdc.noaa.gov/mgg/topo/gltiles.html>, [22. 2. 2021]
- SCHOBER, P., BOER, C. & SCHWARTE, L.A. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126 (5): 1763–1768.
- SILSO DATA, ROYAL OBSERVATORY OF BELGIUM, BRUSSELS, Retrieved from <http://www.sidc.be/silso/infosnmtot>, [26. 2. 2021]
- SIMPSON, J.F. 1967. Solar activity as a triggering mechanism for earthquakes. *Earth and Planetary Science Letters*, 3: 417–425.
- STRASER, V. & CATALDI, G. 2014. *Solar wind proton density increase and geomagnetic background anomalies before strong M6+ earthquakes*. Space Research Institute of Moscow, Russian Academy of Sciences, MSS-14.
- TABACHNICK, B.G. FIDELL, L.S. & ULLMAN, J.B., 2007. *Using multivariate statistics Vol. 5*. Boston, MA: Pearson.
- UNITED STATES GEOLOGICAL SURVEY (USGS), Retrieved from <https://earthquake.usgs.gov/earthquakes/search/>, [13. 11. 2020]

Резиме

Корелабилност соларног ветра са сеизмичким догађајима у зони Балканског полуострва

Термин прекурсор потреса описује велики број различитих физичких феномена који се користе за предвиђање потреса (CICERONE et al., 2009). Предвиђање потреса представља правац сеизмологије који се бави краткорочним, средњерочним и дугорочним прогнозирањем потреса. Детерминистички приступ предвиђању потреса захтева познавање географске ширине и дужине епицентра потреса, магнитуду и време потреса, што се сматра нереалним за разлику од пробабилистичног приступа (врло слично временској прогнози), који даје вероватноћу да ће се догодити потрес за дати регион у одређеном временском периоду. Исто тако, свака метода предвиђања потреса треба да буде искључиво базирана на статистици, односно вероватноћи (KAMER et al., 2021).

Како би се предвиђање потреса извело са што већом сигурношћу, потрага за новим прекурсорима је веома актуелна. Данас, једну занимљиву групу прекурсора чине астрономски прекурсори.

У раду је приказана статистичка корелација између повећаних вредности параметара соларног ветра (густина и брзина протона) и појаве потреса у региону Балканског полуострва у периоду од 1996. до 2018. године.

Приказани резултати за параметар густине протона показују да је у периоду од 14 дана пред потрес аномалија густине протона била је при-

сутна у 80 % случајева (40/50). Оваква вредност корелације је уједно и највећа добијена корелација у истраживању. Статистичка верификација значајности добијених резултата показала је да не постоји никаква значајна статистичка корелација између повећане вредности густине протона и потреса у региону Балканског полуострва.

Приказани резултати за параметар брзине протона највећу корелацију показују у вредности од 56 % (28/50) за референтни период од 14 дана пред потрес. Као и за параметар густине протона, статистичка верификација добијених резултата није показала значајну повезаност између две појаве.

Вршена је провера мањег подкупа аномалних појава соларног ветра, тј. појава аномалије густине и аномалије брзине протона истовремено (у истом дану). За цео истражни сет података постоји укупно 66 таквих дана, односно такви дани чине само 0.7 % истражног сета података. Корелација такве појаве са потресима приказала је да 12 % (6/50) потреса има аномалију и густине и брзине протона у референтном периоду од 14 дана. Статистичка провера овакве корелације није прешла постављени праг значајности појаве од 10 %, односно статистички значајна корелација није пронађена.

Пошто приказане појаве нису показале статистичку значајност са границама аномалије у вредности од пет и три стандардне девијације, вршено је повећање вредности границе аномалије ради провере статистичке корелације са мање аномалних дана. Иако се број аномалних дана за параметар густине протона смањило за око 75 %, а за параметар брзине протона се број аномалних дана смањило за 47 %, статистички значајна корелација није пронађена. Даљим занемаривањем динамичке границе аномалије (зависност границе аномалије од средње вредности и стандардне девијације за дату годину) и постављањем константне границе за аномалије за цео истражни сет података од 100 cm^{-3} за густину протона и 1000 km/s за брзину протона, такође није пронађена значајна статистичка корелација.

Највећа субјективност у овом истраживању своди се, пре свега, на одабир истражног простора, односно на ослањање на тектонске карактери-

стике. Пошто је приказана корелација у MARTICHELLI et al. (2020) на глобалном величинском подручју, те било каква спознаја да ће се потрес догодити нема практичну примену, јер није могуће тачно одредити географску ширину и дужину епицентра потреса, сматрало се да је потребно смањити истражно подручје ради ужег одређивања епицентра потреса. Сматра се да није могуће квантификовати овако унешену субјективност.

Друга значајна субјективност представља повећање референтног периода пред потрес у односу на пријављени референтни период према STRASER & CATALDI (2014). Повећавањем референтног периода повећава се и број потреса који има аномалију. Као и претходна субјективност, сматра се да се ни ова субјективност не може квантификовати.

Последњу субјективност у овом истраживању представља чињеница да повећане вредности куртозе смањују вредност варијанце (ТАВАСНИК et al. 2007), а тиме утичу и на смањење вредности стандардне девијације. Ако су вредности стандардне девијације умањене у односу на њихову истиниту вредност, онда је и граница аномалије умањена за исту ту вредност. Овакву субјективност није потребно квантификовати, већ се повећавањем границе аномалије и одређивањем статистичке значајности такве појаве она и уклонила.

Иако се овом истраживању приступило са скептичног аспекта проналаска позитивне и статистички значајне корелације између повећаних вредности параметара соларног ветра и потреса, сматра се да су у овом истраживању субјективности максимално смањење. GELLER et al. (1997) такође приказује скептичан приступ проналаска поузданог прекурсора потреса, где се наводи да се сваким новим покушајем *a priori* вероватноћа проналаска наредног прекурсора смањује, као и да је тренутна вероватноћа проналаска поузданог прекурсора екстремно мала.

Да би одређена појава била добар кандидат за прекурсор потребно је најпре имати утемељен механизам који доводи до потреса (или бар убрзава потрес). И даље није познат, нити квантификован утицај соларног ветра на потресе, сматра да су потреси и даље тектонски кон-

тролисани, а само евентуално „убрзани“ повећаним вредностима параметара соларног ветра, ако такав механизам постоји.



Приказано истраживање, иако је показало мањак статистичке корелације повећаних вредности параметара соларног ветра и потреса на Балканском полуострву, није искључило постојање позитивне корелације на глобалном величинском подручју. Да би се прецизно одредио утицај соларног ветра на потресе потребно је прво одредити механизам утицаја који ће бити

одређен мереним подацима. Након тога, потребно је одредити статистички значајну методу која предвиђа географску ширину и дужину епицентра, као и време потреса, са високом поузданошћу и малим вредностима стопе лажних аларма. Оваква метода за сада није поуздано откривена ни за један пријављен прекурсор (или потенцијални узрочник).

Manuscript received October 29, 2021

Revised manuscript accepted November 22, 2021

Sustainable Transportation Characteristics Diary—Example of Older (50+) Cyclists

Sreten Jevremović¹ , Carol Kachadoorian², Filip Arnaut³, Aleksandra Kolarski^{1,3} and Vladimir A. Srećković^{1,3,*} 

¹ Scientific Society “Isaac Newton”, Volgina 7, 11160 Belgrade, Serbia; jevremovic749@gmail.com (S.J.); kolarski@ipb.ac.rs (A.K.)

² dbiTilde CORE, Inc., 105 Jacks Point Road, Oxford, MD 21654, USA; carol@dbltildedecore.org

³ Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; arnaut@ipb.ac.rs

* Correspondence: vlada@ipb.ac.rs

Abstract: Cycling is a sustainable and healthy form of transportation that is gradually becoming the primary means of transportation over shorter distances in many countries. This paper describes the dataset used to determine the cycling characteristics of seniors in the USA and Canada. For these purposes, a specially created questionnaire was used in a survey conducted from August 2021 to July 2022. The questionnaire contained sections related to the general socio-demographic characteristics of the respondents, general characteristics of cycling (type of bicycle, cycle time, mileage, etc.), and specific characteristics of cycling (riding in night conditions, termination of cycling, motivating and demotivating factors for cycling, etc.). The total sample consisted of 5096 respondents (50+ years old). This database is particularly significant because it represents the first set of publicly available data related to the cycling characteristics of older adults. The database can be used by various researchers dealing with this topic, but also by the decision-makers who want to design a sustainable and accessible cycling infrastructure, respecting the requirements of this category of users. Finally, this dataset can serve as an adequate basis in the process of determining the specificities and understanding the needs of older cyclists in traffic.

Dataset: <https://zenodo.org/doi/10.5281/zenodo.13908441>, accessed on 24 October 2024.

Dataset License: CC BY 4.0.

Keywords: cycling; seniors; cycling infrastructure; cycling characteristics; traffic; older adults; dataset



Citation: Jevremović, S.; Kachadoorian, C.; Arnaut, F.; Kolarski, A.; Srećković, V.A. Sustainable Transportation Characteristics Diary—Example of Older (50+) Cyclists. *Data* **2024**, *9*, 123. <https://doi.org/10.3390/data9110123>

Academic Editor: Han Woo Park

Received: 2 September 2024

Revised: 9 October 2024

Accepted: 24 October 2024

Published: 25 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Summary

Cycling is gaining prevalence as a competitive mode of transportation in urban areas for shorter distances. In bicycle-centric nations, including the Netherlands, Denmark, and Germany, the proportion of bicycle usage is markedly elevated and exhibits a distinct upward trajectory relative to other countries [1]. However, a comparable tendency is evident in other nations, particularly across the American continent, where cycling is witnessing a revival post-COVID-19, even in relation to the pre-pandemic period. For instance, ridership in the USA has increased by a remarkable 37% since 2019. Every large city with a population of approximately 5 million or more experienced at least a 25% rise in biking activity during that period [2].

The popularity of cycling remains consistent with advancing age, particularly in highly developed nations like the Netherlands and Denmark. Senior bicycle users constitute a vulnerable demographic, necessitating particular considerations about interactions with other road users. In 2019, nearly half (47%) of cycling fatalities in the European Union involved individuals aged 65 or older [3], which illustrates some of the challenges associated with this mode of transportation, alongside its significant appeal.

This research presents a database with a sample of 5096 cyclists aged over 50 from the United States and Canada. The database was compiled with a specifically designed questionnaire over the timeframe from August 2021 to July 2022, encompassing the region of North America (USA and Canada). The survey consists of inquiries that document an individual's riding activities throughout their lifetime, in addition to their present cycling styles, habits, and preferences.

The questionnaire was promoted through many methods, including social media, websites, conference and webinar presentations, and email outreach utilizing a contact list compiled during the three years in which the surveys were conducted.

This research presents a dataset that continues a long-term investigation on the characteristics of senior cyclists in North America. The research was initially carried out in 2018, leading to the establishment of a multi-year best practice that resulted in several significant publications, as displayed in Table 1. The Year 5, 50+ Cycling Survey is now being implemented and is accessible via the following website: <https://survey.alchemer.com/s3/7706636/50-Cycling-Survey-Year-5> (accessed on 9 October 2024).

Table 1. A summary of the most important publications.

No.	Publication Name	Type	Reference
1.	Cycling Past 50: A Closer Look Into the World of Older Cyclists (Year 3)	Report	[4]
2.	Cycling Part 50: A Closer Look Into the World of Older Cyclists (Year 4)	Report	[5]
3.	Ageing and Mobility: A Look at How Ageing Affects Driving and Cycling	Report/ Article	[6]
4.	It's About the Bike: What we Learned about eBikes from the 50+ Cycling Survey, Year 4 Survey	Report/ Article	[7]
5.	Cycling and Nature: The Connection between Cycling and Nature for Older Adults	Report/ Article	[8]

The primary advantage of this database is that it constitutes a freely accessible dataset on the riding characteristics of senior cyclists. This database serves as a great resource for decision-makers aiming to establish a unique, accessible, and sustainable cycling infrastructure while addressing the requirements of this vulnerable demographic. This dataset can provide a sufficient foundation for identifying the specificities and comprehending the demands of older bikers within the transportation and traffic systems of contemporary countries.

2. Data Description

The dataset comprises two sheets: the first, titled "Raw data", contains the data gathered during the survey, and the second, titled "Code book", provides details on all coded variables. If the question is absent from the second sheet, the responses are documented as text (sentences) rather than numerical values and can be located within the dataset.

The questionnaire had 26 questions presented via an algorithm, illustrated in Figure 1. All questions were of the closed-ended type, with responses selected from the provided options. Additionally, respondents could select "other" if their response was not included among the supplied options. In that instance, the respondent possessed the autonomy to compose the answer independently, as evidenced in the database.

Considering that the questionnaire was completed online, the initial page of the survey presented the study's objective and offered information on data protection, researcher contact details, and the anticipated duration for survey completion. The participants were advised that the survey was anonymous and voluntary, and that they could withdraw from the study at any time. Upon providing their informed consent to participate in the study, the respondents were forwarded to the initial set of questions.

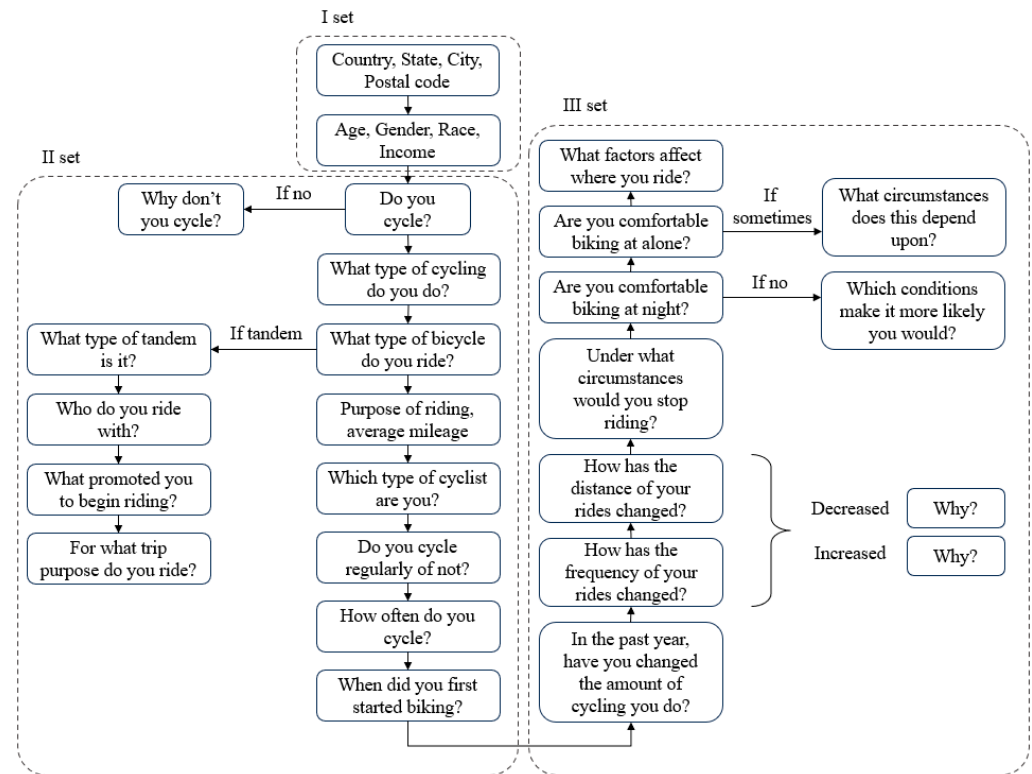


Figure 1. Algorithm of questions in the questionnaire.

Figure 1 illustrates that the questionnaire comprised three sets of questions:

Set I: General socio-demographic questions related to the place of residence and characteristics of the respondents such as gender, age, and economic status.

Set II: Questions such as the type of bicycle used, purpose of travel, mileage, and frequency of cycling.

Set III: Specifics related to the use of the bicycle, such as the conditions under which the bicycle would be used and the use of the bicycle at night or alone (without the presence of another person).

The primary objective of the questions structured in this manner and the organized questionnaire was to ascertain the cycling characteristics of older individuals, as well as to identify their preferences, such as situations in which they feel at ease or uncomfortable, their preferred times for cycling, and their inclination to use a bicycle for transportation, recreation, or other purposes.

The data were gathered and stored in an Excel database, where they were prepared and checked, which will be elaborated upon in the subsequent section.

The Excel file containing the data is structured such that each column header corresponds to a question or a component of a question from the questionnaire. The following is a concise overview of the data contained in the Excel file:

The initial four columns (A, B, C, and D) pertain to generic information regarding the respondents: country, abbreviated name, city, and postal code. No data coding occurred; all information was documented exactly as provided by the respondents.

Column E documents the responses to the inquiry “do you cycle?”. The responses are encoded in binary (see the codebook).

Columns F–M provide the reasons for respondents’ reluctance to cycle, with each column documenting the corresponding response. This question allows multiple answers.

Columns N–V correspond to the specific form of riding practiced by the respondents. Multiple answers could also be selected.

Columns W–AE denote the type of bicycle used.

Columns AF-AR delineate the attributes of the tandem bicycle: its purpose, rationale for use, and intended users.

Columns AS-AX shows the general socio-economic characteristics of the respondents, including age, gender, ethnicity, and income, as detailed in Table 2.

Table 2. Socio-demographic characteristics of the sample.

Variable	N	[%]
Country		
USA	4324	84.8
Canada	772	15.2
Age		
50–59	1401	27.5
60–69	2092	41.1
70–79	1025	20.1
More than 80	249	4.9
No answer	329	6.4
Gender		
Female	1665	32.7
Male	3142	61.6
Other	9	0.2
No answer	280	5.5
Race		
Asian	128	2.5
Native Hawaiian or Other Pacific Islander	7	0.1
Black/African American	76	1.5
White	4328	84.9
Hispanic/Latino	64	1.3
American Indian/Alaska Native	11	0.2
At least two	39	0.8
Other	58	1.1
Prefer not to answer	385	7.6
Annual Income		
Less than USD 25,000	109	2.1
USD 25,000 to USD 34,999	182	3.6
USD 35,000 to USD 49,999	322	6.3
USD 50,000 to USD 74,999	722	14.2
USD 75,000 to USD 99,999	775	15.2
USD 100,000 to USD 124,999	727	14.3
USD 125,000 to USD 149,999	452	8.9
USD 150,000 or more	1272	25.0
Prefer not to say	535	10.4

Columns AY-BJ present data regarding the frequency of usage and the exact conditions under which the bicycle is used.

Columns BK-CQ correspond to a series of inquiries regarding the initiation of cycling, variations in the frequency and distance of cycling over the past year, and the rationale behind these alterations.

Columns CR-DD present information regarding the circumstances that would prompt respondents to resume cycling after having previously ceased.

Columns DE-DM present information regarding the user's solitary cycling habits and the settings under which such activity occurs.

The DN-EF columns present information regarding the causes and conditions that affect the user's cycling preferences.

Columns EG-GA present information regarding the purpose of travel and average mileage.

The final column, GB, presents information regarding the type of cyclist users identify themselves as.

3. Methods

The data included in this study were gathered in the USA and Canada from August 2021 to July 2022. Participants engaged in an online survey conducted via the Alchemer Survey platform, which was distributed over all of North America. The poll was disseminated through multiple methods, including social media, websites, conference and webinar

presentations, and email outreach utilizing a contact list compiled over the three years during which the surveys were conducted.

The survey was conducted in accordance with the Declaration of Helsinki of 1975, revised in 2008, and was approved by the Institutional Review Board of San Jose State University and the Board of Directors of dbiTilde CORE, Inc. (No. 029/24).

By applying one of the most commonly used sampling methods, the snowball technique [9–11], participants were enlisted to disseminate the survey link across their networks. No inclusion criteria were established other than age, as the intended demographic for this dataset comprised riders aged 50 and above. Microsoft Excel was employed to recode, calculate, and analyze the obtained data as necessary.

3.1. Data Validation

After data collection, validation and verification was carried out through several steps:

- Logical validation (consistency check)—ensuring that all answers make sense in relation to each other. For example, if a person indicated that the cycling distance decreased during the review period, and the average mileage for the same period is more than 40 but fewer than 50 miles, that might indicate an inconsistency. Such responses were thoroughly checked and excluded from the analysis if inconsistencies were found. We also removed responses registered from countries other than the USA and Canada. Also, we removed nearly incomplete responses, as well as other responses determined to be not suitable for inclusion.
- Range validation—verifying that numerical answers fall within the expected ranges. For example, if asking about age, the age range is between 50 and 100.
- Format validation—verifying that there is no placeholder text (e.g., “gfdgda”, or “0000”) in open-ended questions.
- Outlier Detection—for numerical data, we used min and max functions to identify responses that deviate significantly from the norm, mainly for the age question, since all of the other questions were prepared in such a way that the respondents could choose one of the offered answers.
- Cross-validation—checking for multiple submissions from the same respondent, and removing them. Also, for the sake of reliability, a brief cross-verification of the obtained results with existing and available socio-economic (e.g., income) and traffic data (for which it was possible to obtain the data, e.g., average cycling mileage, cycling purpose) at the state level was performed. Based on this comparison, the obtained data (in the survey) corresponds to the real data with a deviation of up to 5%.

In this way, 404 respondents were excluded from the sample. A summary of the final sample's ($N = 5096$; MeanAge = 65, StDevAge = 33.2) socio-demographic characteristics is provided in Table 2.

3.2. Data Usability

The provided dataset can be utilized in two principal ways.

Researchers can examine the various factors and characteristics of cycling among older users. It would be crucial to examine the reasons why respondents refrain from cycling and identify the necessary conditions to encourage these individuals to commence cycling [12–14]. Moreover, analyzing the sorts of bicycles utilized and their respective uses, particularly focusing on cargo and tandem bicycles, would be useful. Another approach is to examine the reasons for the alteration in cycling frequency (either a drop or increase) during the preceding period. An interesting direction for future research involves examining the characteristics of cycling under night conditions, along with the factors that influence bicycle usage in specific areas.

Decision makers: this dataset serves as a suitable foundation for the prospective planning of new infrastructure. For instance, the type of bicycle utilized can inform the planning of appropriate infrastructure for the specified categories [15]. Data on the predominant routes is required for this purpose. Conversely, alterations in the frequency and distance of cycling,

along with the motivations behind these changes [16], may indicate potential deficiencies in the current infrastructure, necessitating modifications to accommodate older users.

4. Limitations of the Study and Future Research

It is important to note some limitations and restrictions that should be considered by researchers who wish to use the current dataset. The survey was initially conducted online, potentially introducing sampling bias by excluding elderly persons without internet access. Future research of this nature, commencing with data collection in early 2025, also intends to conduct live surveys of older users to enhance the quality of the data acquired. Also, response bias should be taken into account due to the nature of data collection that involves surveying; that is, it should be taken into account that the answers are self-reported and may reflect a certain subjectivity on the matter. A concise cross-verification of the acquired results with available socio-economic (e.g., income) and traffic data (for which data were accessible, such as average mileage and cycling purposes) at the state level was conducted. The comparative analysis indicates that the survey data align with the actual data, exhibiting a variance of up to 5%.

The quality of the acquired data can be enhanced by gathering further information, including on health problems and environmental factors, which is scheduled for the next research phase and will be incorporated into subsequent questionnaires.

Author Contributions: Conceptualization, S.J., C.K., A.K., V.A.S. and F.A.; methodology, S.J.; validation, S.J. and C.K.; resources, A.K., V.A.S. and F.A.; data curation, S.J. and C.K.; writing—original draft preparation, S.J.; writing—review and editing, S.J., C.K., A.K., V.A.S. and F.A.; visualization, S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The survey was conducted in accordance with the Declaration of Helsinki of 1975, revised in 2008, and was approved by the Institutional Review Board of San Jose State University and the Board of Directors of dbiTilde CORE, Inc. (No. 029/24).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data described in this paper can be found at the following link: <https://zenodo.org/doi/10.5281/zenodo.13908441> (accessed on 9 October 2024).

Conflicts of Interest: Author Carol Kachadoorian was employed by the company dbiTilde CORE, Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Buehler, R.; Goel, R. Chapter Seven—A Global Overview of Cycling Trends. *Adv. Transp. Policy Plan.* **2022**, *10*, 137–158. [CrossRef]
2. Hurford, M. Good News, Everyone! People Are Riding More than Ever. Available online: <https://www.bicycling.com/culture/a45325157/good-news-everyone-people-are-riding-more-than-ever/> (accessed on 9 September 2024).
3. European Commission Facts and Figures Cyclists. *European Road Safety Observatory*; European Commission Facts and Figures Cyclists: Brussels, Belgium, 2021.
4. Kachadoorian, C. *Cycling Past 50: A Closer Look into the World of Older Cyclists*; Mineta Transportation Institute: San Jose, CA, USA, 2021.
5. Kachadoorian, C. *Cycling Past 50: A Closer Look into the World of Older Cyclists, Year 4 Survey*; Mineta Transportation Institute: San Jose, CA, USA, 2023.
6. Kachadoorian, C. *Ageing and Mobility: A Look at How Ageing Impacts Driving and Cycling*; Mineta Transportation Institute: San Jose, CA, USA, 2023.
7. Kachadoorian, C. *It's About the Bike: What We Learned About EBikes from the 50+ Cycling Survey, Year 4*; Mineta Transportation Institute: San Jose, CA, USA, 2023.
8. Kachadoorian, C. *Cycling and Nature: A Look at the Connection Between Cycling and Nature for Older Adults*; Mineta Transportation Institute: San Jose, CA, USA, 2024.
9. Wright, R.; Stein, M. Snowball Sampling. *Encycl. Soc. Meas.* **2005**, *1*, 495–500. [CrossRef]
10. Stivala, A.; Koskinen, J.; Rolls, D.; Wang, P.; Robins, G. Snowball Sampling for Estimating Exponential Random Graph Models for Large Networks. *Soc. Netw.* **2016**, *47*, 167–188. [CrossRef]
11. Leighton, K.; Kardong-Edgren, S.; Schneidreith, T.; Foisy-Doll, C. Using Social Media and Snowball Sampling as an Alternative Recruitment Strategy for Research. *Clin. Simul. Nurs.* **2021**, *55*, 37–42. [CrossRef]

12. Tarkkala, K.; Vosough, S.; West, J.; Roncoli, C. Evaluating the influence of cyclists' route choices incorporation into travel demand modelling: A case study in greater Helsinki. *Trans. Res. Int. Pers.* **2024**, *27*, 101224. [[CrossRef](#)]
13. Uijtdewilligen, T.; Baran Ulak, M.; Jan Wijlhuizen, G.; Geurs, K. Effects of crowding on route preferences and perceived safety of urban cyclists in the Netherlands. *Trans. Res. Part A.* **2024**, *183*, 104030.
14. Zeng, Q.; Gong, Z.; Wu, S.; Zhuang, C.; Li, S. Measuring cyclists' subjective perceptions of the street riding environment using K-means SMOTE-RF model and street view imagery. *Int. Jour. App. Earth Obs. Geoinf.* **2024**, *128*, 103739. [[CrossRef](#)]
15. von Stülpnagel, R.; Rintelen, H. A matter of space and perspective—Cyclists', car drivers', and pedestrians' assumptions about subjective safety in shared traffic situations. *Transp. Res. Part A* **2024**, *179*, 103941. [[CrossRef](#)]
16. Hagemester, C.; Bertram, L. Reported pushy driving against cyclists in Germany. *Jour. Saf. Res.* **2024**, *88*, 395–405. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

КАТЕГОРИЈА М30

Analysis of possible relationship between earthquakes and Solar flare events during Solar cycle 24

Filip Arnaut^{1*}

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Republic of Serbia

*Correspondence: Filip Arnaut, filip.arnaut@ipb.ac.rs

Abstract: Earthquakes pose a significant threat to the population, causing harm to individuals and also impacting modern infrastructure. Prediction of earthquakes as a subfield of seismology aims to address questions and problems related to forecasting of earthquakes. An area of research that could contribute to this matter is the field of space weather specifically related to the impact of the sun on Earth. This brief communication presents a statistical analysis of occurrences of solar flares and earthquakes during the 24th solar cycle. Overall, the findings indicate that it is unlikely to attribute solar flares as a trigger factor for earthquake events; however, the analysis reveals a limited number of intriguing instances that warrant further research.

Keywords: space weather, earthquake prediction, statistical analysis.

Introduction

Earthquakes (EQs) are natural occurrences in which seismic waves emitted from a hypocenter propagate outward through surrounding geological structures. During an EQ, a significant amount of energy is released, resulting in potentially catastrophic effects on the population. Consequently, significant endeavors have been undertaken throughout history to assess potential precursors of EQs and to statistically ascertain them. However, predictions of EQs are still with minimal success.

This brief communication presents a statistical analysis of solar flares (SFs) as a potential trigger of EQs on the Balkan peninsula during the 24th solar cycle (SC). A multitude of literature studies have proposed correlations between different space weather parameters and EQs. For instance, solar wind (Straset et al. 2014; Cataldi et al. 2016; Marchitelli et al. 2020) and SFs (Novikov et al. 2020; Takla, Samwel, 2023) have been linked to EQs. However, there are also literature findings that raise serious questions about the predictability of EQs as a whole (Rhoades, Evison, 1987; Geller et al. 1997; Kagan, 1997; Hough, 2010). However, conducting a statistical analysis of the strongest (X-class) SFs and EQs registered in the region of the Balkan Peninsula is advantageous in order to verify possible the correlation between SFs and EQs, despite the low likelihood to accurately predict future EQs based on reported SF events.

Methods and data

The EQ data was acquired from the Earthquake Catalog of the United States Geological Survey (<https://earthquake.usgs.gov/earthquakes/search/>). The dataset used in this research was collected for the Balkan peninsula during the 24th SC, covering a 11-year time period with EQs of at

least magnitude $M \geq 4.0$. Further data processing was conducted in QGIS to eliminate any EQs that are not exclusively located within the Balkan Peninsula.

The algorithm developed for this study was thoroughly presented in Arnaut et al. (2021). Utilized algorithm relies on the use of 5 parameters in total. The parameters used in this study include the time period (measured in days), the number of SFs of a specific class (X- class), the number of EQs, the look-back parameter that defines the duration of days before the EQ that are considered, and the number of iterations conducted by the algorithm.

The algorithm's workflow involves selecting a random sample from a specific time period for two separate groups of data. The first group, referred to as Group A, selects random numbers from a range of 1 to the value of the time period. This group is alternatively referred to as the SF group, as it employs a random distribution of SFs within a given time period. The second group, referred to as Group B, is designated as the EQ group, and values for this group are selected at random from a range of 1 to the value of the time period. This group emulates the random distribution of EQs throughout the specified time period. The initial iteration of the workflow examines each value in group B to determine if there are any values in group A prior to the given value of group B in the length of the look-back parameter. In this way, SFs and EQs are randomly distributed throughout the duration of the analyzed period. Each EQ was assessed to determine if there was a SF in the N days (lookback value) before it. The process was iterated 10,000 times, and the final outcomes were examined. Prior to executing the simulation, we conducted a count of the number of EQs that had an X-class SF(s) for the period of 14, 7, 4, and 1 day(s) before the EQ itself. If the simulated count was greater than the actual count, the results cannot be interpreted as indicating that SFs trigger EQs

in general, as they do not differ from a random chance draw. In order that any correlation between these two phenomena can be considered valid, a higher occurrence rate compared to one obtained in the case of random chance alone is expected.

Results and discussion

The simulation was initially conducted for a look-back value of 14 days, which was an extremely generous duration, as there is currently no evidence to support a two-week effect following an SF. Based on random chance, the simulation predicted that 139 EQs should have an SF in the two weeks prior to the EQ. This value ranged from 102 to 176 out of the 633 EQs in total (16-27%). The actual count of the number of EQs that were affected by SF in the two weeks prior to the current data was 62 (Table 1). This number is insufficient to be considered an actual correlation for all of the instances that were displayed.

The simulation was replicated for a 7-day period prior to the occurrence of the EQ as well. The actual count of EQs that experienced an X-class SF in the week preceding the EQ was 30. However, the analysis indicates that the average count is 74, with a minimum of 46 and a maximum of 101, solely due to random chance. Just like in the previous example, this cannot be considered a correlation, since for establishing the correlation, it would be expected that the number of positive cases is much higher compared to the case of a random distribution.

The simulation was also conducted for a retrospective period of 4 and 1 day(s), yielding results for the 4-day period relatively similar to those previously discussed. However, results obtained for the 1-day period are quite intersecting. In period one day prior to an EQ itself, there were a total of 8 earthquakes with an X-class SF occurring the day before the EQ. The simulated values ranged from a minimum of 1 to a maximum of 23, with an

average value of 11. The simulation was replicated with an increased number of 250,000 iterations, yielding comparable overall results, albeit with a higher maximum value of 27.

Table 1. Obtained results of the simulation

Simulation 1	Mean	Min	Max
Look-back parameter [days]	14		/
True number of EQs with a SF before it	62		/
Simulated expected occurrences EQs with SF affected days before	139	102	176
Simulation 2			
Look-back parameter [days]	7		/
True number of EQs with a SF before it	30		/
Simulated expected occurrences EQs with SF affected days before	74	46	101
Simulation 3			
Look-back parameter [days]	4		/
True number of EQs with a SF before it	21		/
Simulated expected occurrences EQs with SF affected days before	43	24	72
Simulation 4			
Look-back parameter [days]	1		/
True number of EQs with a SF before it	8		/
Simulated expected occurrences EQs with SF affected days before	11	1	23

SF- Solar flare; EQ- Earthquake; Number of iterations 10,000; Total number of days in the 24th Solar cycle 4379; Number of X-class SFs in the 24th Solar cycle 77; Total number of EQs in the 24th Solar cycle 633

An in-depth analysis was conducted for cases of these 8 EQs that experienced an X-class SF during the day prior. Among

these 8 cases, two were noteworthy as they were involved in multiple X-class SFs in the days leading up to the EQ. The first EQ analyzed in detail was one that took place on September 25, 2011, and was “associated” with several SFs. Specifically, an X-class SF with a magnitude of X2.14 occurred on September 22, two X-class SFs with magnitudes of X2.74 and X1.02 occurred on September 24, and one X-class SF with a magnitude of X1.07 occurred on the day of the EQ. The second EQ analyzed in detail occurred on May 15th, 2013, and was “associated” with multiple X-class SFs. Specifically, there were two X-class SFs on May 13th, 2013 (X4.11 and X2.51), one on May 14th, 2013 (X4.64), and one on the day of the EQ itself (X1.85). The initial EQ occurred within the borders of Albania (city of Durres), while the subsequent one took place in Montenegro (11 km from Herceg Novi).

Based on the obtained results, SFs cannot serve as a trigger for EQs, mainly due to the complex nature of EQs, which involve numerous parameters that cannot be measured or modeled in an adequate manner. The sample size of 8 earthquakes is insufficient to draw any statistically reliable conclusions about the correlation between analyzed SFs and EQs, specifically the triggering effect of SFs on EQs. However, it remains an intriguing avenue of research that can hold significance if approached appropriately and with statistical rigor. It should be kept in mind that much previous research attempted to establish correlations between various phenomena and EQs but ultimately failed to do so, leaving the identification of a potential trigger for EQs as a still-open and very interesting topic of research.

Conclusions

This communication presents a statistical analysis of reported SFs and registered EQ events conducted to investigate whether SFs

have a triggering effect on EQs in the Balkan Peninsula zone during the 24th SC. The analysis determined that although there are certain instances that warrant additional research, it is not possible to definitively establish a causal relationship between SFs and EQs or to utilize SFs as reliable predictors of EQs. In general, the evidence suggests that SFs are unlikely to be a factor in causing EQs.

References

- Arnaut, F., Vučković, D., Vasiljević, I. and Cvetkov, V., 2021. Correlability of solar wind with seismic events in the Balkan peninsula zone. *Geoloski anali Balkanskoga poluostrva*, 82(2), pp.69-83.
- Cataldi, G., Cataldi, D. and Straser, V., 2016. Solar activity correlated to the M7. 0 Japan earthquake occurred on April 15, 2016. *New Concepts in Global Tectonics Journal*, 4(2), pp.202-208.
- Geller, R.J., Jackson, D.D., Kagan, Y.Y. and Mulargia, F., 1997. Earthquakes cannot be predicted. *Science*, 275(5306), pp.1616-1616.
- Hough, S.E., 2010. *Predicting the unpredictable: the tumultuous science of earthquake prediction*. Princeton University Press.
- Kagan, Y.Y., 1997. Are earthquakes predictable?. *Geophysical Journal International*, 131(3), pp.505-525.
- Marchitelli, V., Harabaglia, P., Troise, C. and De Natale, G., 2020. On the correlation between solar activity and large earthquakes worldwide. *Scientific reports*, 10(1), p.11495.
- Novikov, V., Ruzhin, Y., Sorokin, V. and Yaschenko, A., 2020. Space weather and earthquakes: possible triggering of seismic activity by strong solar flares. *Annals of Geophysics*, 63(5), pp.PA554-PA554.

- Rhoades, D.A. and Evison, F.F., 1989. On the reliability of precursors. *Physics of the Earth and Planetary interiors*, 58(2-3), pp.137-140.
- Straser, V., Cataldi, G. and Cataldi, D., 2014. Solar wind proton density increase and geomagnetic background anomalies before strong M6+ earthquakes. *Space Research Institute of Moscow, Russian Academy of Sciences, MSS-14*.
- Takla, E.M. and Samwel, S.W., 2023. Possible connection between solar activity and local seismicity. *Terrestrial, Atmospheric and Oceanic Sciences*, 34(1), p.9.

[<https://doi.org/10.69646/aob241203>]

[*Invited Lecture*]

Daytime and nighttime VLF signal classification utilizing machine learning methods

Filip Arnaut^{1*}

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Republic of Serbia

*Correspondence: filip.arnaut@ipb.ac.rs

Abstract: The automatic classification of ionospheric very low frequency (VLF) signals is a current research endeavor aimed at creating a machine-learning (ML) methodology capable of differentiating among various influences on VLF signals, including solar flares, VLF receiver malfunctions, nighttime VLF signals, and other factors. This communication discusses the enhancement in ML classification of daytime and nighttime ionospheric VLF signals, including the different methodologies, data processing, and various processes that demonstrated improvement over prior research.

Keywords: Machine learning, data-driven modelling, anomaly detection, geophysics.

Introduction

The automatic classification of very low frequency (VLF) signal disturbances and characteristics through machine learning (ML) techniques is a subject previously presented (Arnaud et al., 2023; Arnaut, Kolarski, 2023; Arnaut et al. 2024) and remains an area of active research. VLF signals generally capture the impacts of solar flares, instrument malfunctions, erroneous measurements and other, as well as the variation between daytime and nighttime signals.

Typically, the nighttime signal exceeds the daytime signal, with the exception of the terminator, which is characterized by a minor decline in the signal prior to a pronounced ascent, ultimately stabilizing at a level exceeding that of the daytime signal. The ongoing research focuses on the automatic classification of diverse effects on the VLF signal, aiming to establish a method that minimizes error in this classification process. This research demonstrates the advancements in the classification of daytime and nighttime VLF signals, the modifications implemented during this period, and the resultant outcomes.

Methods and data

The primary distinction in the methodology between prior research and the current study is that data labeling in the latter was conducted on an individual case basis. Daytime and nighttime conditions were previously established through conditional labeling based on the local receiver time. The transition from imprecise conditional labeling to manual data labeling enhanced data quality, thereby improving the predictive capabilities of models and facilitating a more accurate differentiation between daytime and nighttime conditions.

Secondly, the feature list was expanded to include weighted moving averages, in addition to the previously employed statistical features, whereby data points nearer to the instance under analysis are assigned a greater weight coefficient. Furthermore, the class balancing approach of random undersampling was replaced with the Synthetic Minority Oversampling Technique (SMOTE), which ensures that no data is discarded, as occurs in random undersampling. The random forest model was ultimately replaced by the extreme gradient boosting (XGB) model, utilizing the random search hyperparameter tuning method for tuning of the number of estimators and the learning rate.

The model's predictions underwent cluster analysis, as the nighttime signal extends over a longer duration; clusters of relatively few consecutive nighttime labels were reclassified using the cluster analysis. The outcomes for both raw model predictions and predictions reclassified through cluster analysis are presented in this study.

Results and discussion

As previously stated, the training dataset was balanced using the SMOTE technique, which ensured that no data points were omitted, while the minority class (specifically, the nighttime signal class) was oversampled. The original distribution in the training dataset was 70-30, favoring the daytime signal; therefore, SMOTE was employed to oversample the minority class.

The hyperparameter tuning method chosen was random search for both the number of estimators (ranging from 100 to 1000 in increments of 20) and the learning rate (ranging from 0.01 to 0.2 in increments of 0.01). The most effective model exhibited 840 estimators and a high learning rate of 0.2. The initial model exhibited accuracy, precision, and F1-score values of 0.79, 0.71, and 0.78, respectively, while the AUC value was 0.8, indicating an acceptable capacity of the model to differentiate between the classes.

Figure 1 illustrates a dataset comprising 1000 data points, specifically in minute intervals, pertaining to the NLK-Sheridan transmitter-receiver pair, where the nighttime signal is represented (true class labels are shown in the upper panel and predicted labels in the middle panel). The model's raw output exhibited relatively satisfactory classifications, demonstrating moderate predictive capability to differentiate between nighttime and daytime VLF conditions in the provided example. The classification of the daytime-to-nighttime terminator and the segment of the signal immediately following the nighttime-to-daytime terminator could be further improved as to align more with the true classifications.

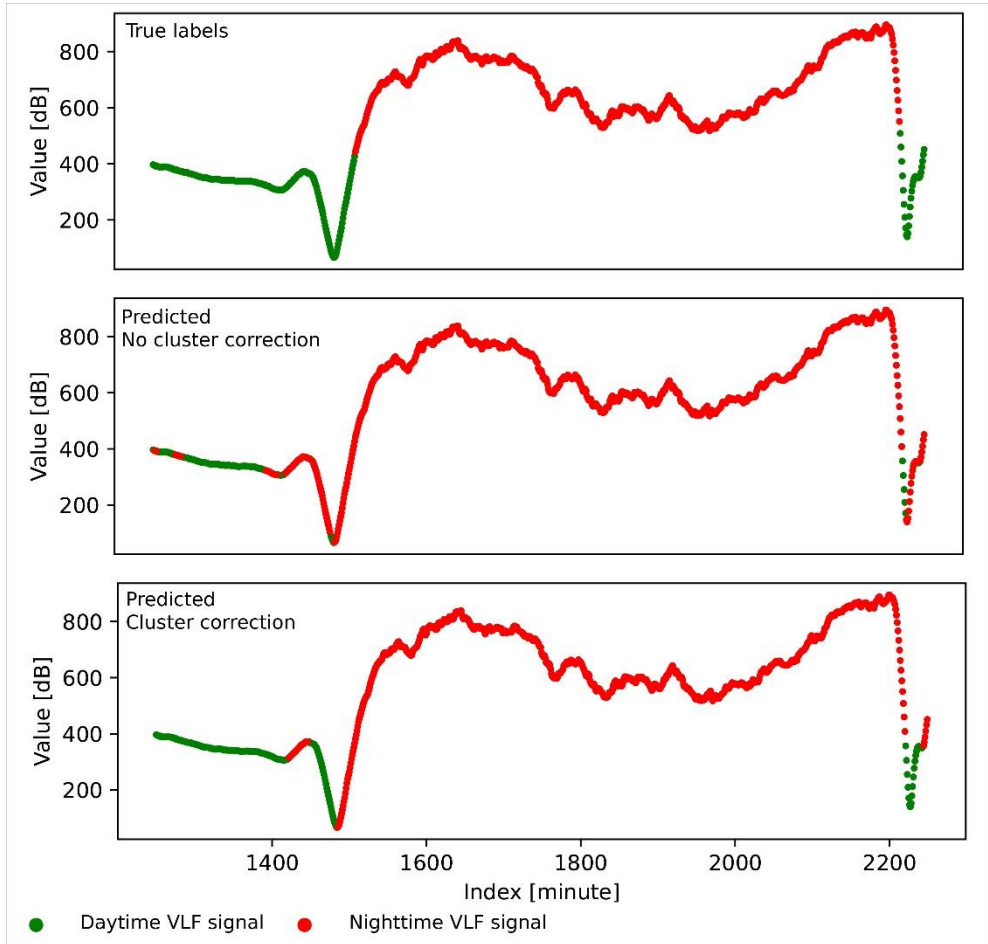


Figure 1. True daytime and nighttime labels (upper panel); Predicted labels without the cluster correction (middle panel); Predicted labels with the cluster correction (lower panel)

Cluster analysis was utilized to correct the minor groups of nighttime labels in the predicted signal. The cluster analysis statistics revealed a significant positive skew in the distribution of cluster lengths, suggesting that the model generates a considerable number of brief nighttime predictions. Conversely, the nighttime signal is expected to be present for a longer duration; consequently, all clusters with a consecutive prediction length of fewer than 25 were

reclassified as daytime signals. Figure 1 (bottom panel) illustrates the reclassification of small clusters previously identified as nighttime signals to daytime signals. The cluster analysis functioned as a corrective measure in this instance and exhibited satisfactory characteristics.

The comparison of evaluation metrics between cluster-corrected and non-cluster-corrected values reveals relatively similar results, occasionally favoring the cluster-corrected values (Table 1). The precision and F1-score for the nighttime class exhibit increased values, whereas the recall parameter shows a lower value for the nighttime class in the cluster-corrected predictions.

Table 1. Comparison between evaluation metrics for the non-cluster corrected and the cluster corrected predictions

	Prec. NC	Prec. C	Rec. NC	Rec. C	F1 NC	F1 C
Daytime class	0.88	0.86	0.74	0.82	0.8	0.84
Nighttime class	0.71	0.77	0.86	0.82	0.78	0.8
Macro averaged	0.8	0.82	0.8	0.82	0.79	0.82
Weighted averaged	0.81	0.82	0.79	0.82	0.79	0.82

NC- No cluster C- Cluster

The integration of the cluster correction yielded satisfactory results; however, it requires additional refinement to fully automate the process and produce improved, more precise predictions.

Conclusions

The complete automation of ionospheric VLF signal classification will require considerable time and extensive research effort. This communication presents the enhancement of ML classification for daytime and nighttime VLF signal conditions. The improvement was achieved through case-by-case manual labeling, replacing random undersampling with SMOTE for training dataset

balancing, substituting the random forest model with the XGB model, expanding the feature list and employing a cluster correction after the classification process. The results are promising; however, additional refinement and improvement is necessary, which will be the focus of subsequent research.

Acknowledgement

VLF data are provided by the WALDO database (<https://waldo.world>, accessed on 1 January 2023), operated jointly by the Georgia Institute of Technology and the University of Colorado Denver, using data collected from those institutions as well as Stanford University, and has been supported by various US government grants from the NSF, NASA, and the Department of Defense.

References

- Arnaud, F. & Kolarski, A., 2023. Machine learning approach for distinguishing daytime and nighttime ionospheric conditions on VLF signals related to solar flares during 2011. In XX Serbian Astronomical Conference, 16-20 October 2023, Belgrade, Serbia. Book of Abstracts, Astronomical Observatory of Belgrade and Faculty of Mathematics, pp. 79
- Arnaud, F., Kolarski, A. and Srećković, V.A., 2023. Random forest classification and ionospheric response to solar flares: Analysis and validation. *Universe*, 9(10), p.436.
- Arnaud, F., Kolarski, A. and Srećković, V.A., 2024. Machine Learning Classification Workflow and Datasets for Ionospheric VLF Data Exclusion. *Data*, 9(1), p.17.

Multilayer Perception Hyperparameter Fine-Tuning for Ionospheric VLF Amplitude Data Exclusion

Filip Arnaut and Aleksandra Kolarski

*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080
Belgrade, Serbia
E-mail: filip.arnaut@ipb.ac.rs*

Abstract

The analysis of ionospheric amplitude data is affected by various factors, including the influence of solar flare events, instrument malfunction, and other sources of error. These factors collectively contribute to a decrease in the overall quality of the data. The removal of such data is performed manually by researchers, a process that is characterized by its labor-intensive nature and time-consuming requirements. This research paper presents a procedure for fine-tuning of the learning rate (LR), number of epochs and the momentum in Multilayer Perception (MLP) classification models. The proposed method can be utilized as a benchmark for optimizing other hyperparameters in the future.

Introduction

Very low frequency (VLF) ionospheric amplitude data is adversely affected by solar flare events and instrumental errors, including malfunctions. In order to render the VLF amplitude data applicable for further analyses by researchers, it is imperative to eliminate those effects. The process of manually removing these effects is known to be time-consuming, tedious, and labor-intensive. Consequently, there is an advantage in automating this process using machine learning (ML) classification techniques.

In order to accomplish this task, a range of ML techniques can be utilized, including Artificial Neural Networks (ANN) such as Multilayer Perception (MLP) models, as well as traditional ML models like Random Forests (RF) and Support Vector Machines (SVM). The distinction between MLP models and classic models, such as RF models, becomes evident when considering the quantity of hyperparameters that require configuration. In the case of RF models, the researcher is only required to determine the number of trees. Conversely, MLP models necessitate the fine-tuning of multiple hyperparameters, including the learning rate (LR), momentum, and the number of epochs among others. This study

aims to demonstrate the process of fine-tuning the LR, number of epochs and the momentum. This initial step will serve as a starting point in identifying the most suitable MLP model for the intended research objective, which involves the automated identification of erroneous data points in ionospheric VLF amplitude data.

Methods and data

The data employed for this study consists of VLF amplitude measurements obtained during the months of September and October of 2011. These measurements capture solar flare events falling within the range of C2.5 to X2.1 in terms of their class. The training dataset for September 2011 consists of 59,344 datapoints after being balanced as to remove any bias (Prusa et al. 2015; Kulkartni et al. 2020; Devi et al. 2020), while the testing dataset for October 2011 comprises 180,071 data points. The training and testing datasets consist of 40 features, encompassing the original VLF amplitude data, X-ray irradiance data, encoded values of the transmitter and receiver, and the local receiver time expressed in decimal points. Statistical features encompass additional features, including rolling window statistics that employ diverse window lengths (5, 20, and 180 minutes) and varying parameters such as mean, median, and standard deviation. The features utilized in the analysis also consisted of the first and second differential data of both VLF amplitude and X-ray measurements.

The process of searching and optimizing hyperparameters was conducted in a two-step manner. The initial step consisted of determining the optimal combination of the LR and the number of epochs, while maintaining a constant momentum hyperparameter of 0.2. The second phase of the procedure involved determining the optimal momentum value, while keeping the LR and number of epochs constant. In the initial phase, the LR and the number of epochs were subject to variation within the ranges of 0.1 to 0.85 and 100 to 500, respectively. The LR varied by increments of 0.1, while the number of epochs varied by increments of 200. The initial iteration of the modeling process consisted of a total of 18 models, whereas the second stage of modeling involved a total of 9 models. The Weka software package (Frank et al. 2016) included several noteworthy parameters, one of which was the number of nodes. In the case of this MLP, this value was determined by dividing the sum of the number of features and classes by 2, resulting in a total of 22 nodes.

The evaluation of all models involved the utilization of standard ML classification methods, including the F-measure, the area under the receiver operating characteristic curve (ROC), the number of Correctly Classified Instances (CCI), the percentage of Incorrectly Classified Instances (ICI), Cohen's kappa statistics, true positive (TP) and false positive (FP) rates, as well as the Matthew's correlation coefficient (MCC). Due to the inherent imbalance in the ML problem at

hand, the analysis primarily emphasized the evaluation of the MCC and kappa values. These metrics were chosen as they offer a less biased assessment of the model compared to other measures like the F-measure for imbalanced datasets (Chicco, Jurman, 2020). Furthermore, the F-measure, TP and FP rates were examined subsequent to the evaluation of MCC and kappa statistics. This analysis included an examination of per-class statistics. The ROC value was additionally analyzed in order to ascertain the models' ability to distinguish between different classes.

Results and Discussion

The preliminary analysis involved conducting modeling tests with different LR values and varying the number of epochs. The resulting MCC values ranged from 0.34 to 0.42. Notably, the model trained with an LR value of 0.85 and 500 epochs achieved the highest MCC value among all the models tested. Furthermore, it is worth noting that the aforementioned model exhibited the highest Kappa coefficient, which was measured at 0.36. The weighted TP rates for all models ranged from 0.659 to 0.745. In contrast, the TP rate for the anomalous data class ranged from 0.94 to 0.73, with the best model achieving a TP rate of 0.843. The average weighted F-measure for all 18 models was found to be 0.753. However, the model that was considered the best overall exhibited a higher weighted F-measure value of 0.774, surpassing the mean value. In general, the model with a LR of 0.85 and 500 epochs exhibited satisfactory evaluation metric statistics. It demonstrated a higher F-measure compared to the average F-measure of all the models, as well as higher values for the MCC and Kappa coefficient.

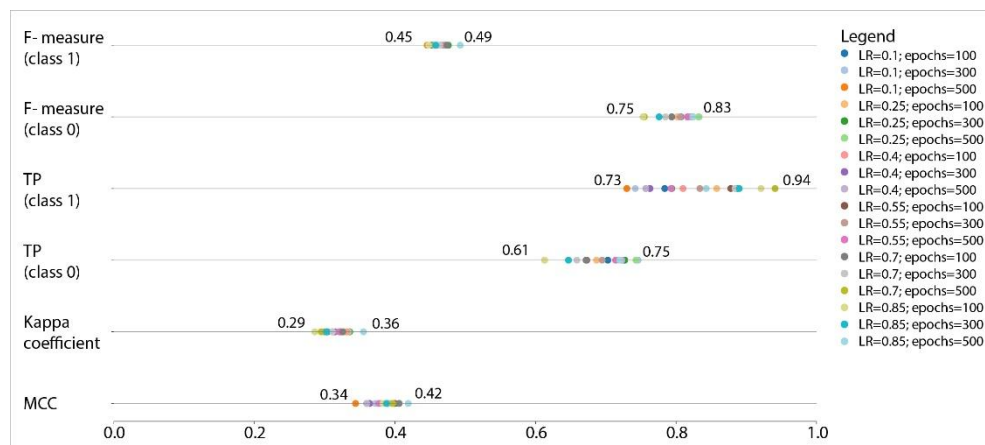


Fig. 1. Evaluation metrics for the first phase of the hyperparameter tuning.

The second iteration of modeling was conducted using a fixed LR and epoch values of 0.85 and 500, respectively. However, the momentum was not constant and varied from 0.1 to 0.9 in increments of 0.1. The observed CCI values exhibited a greater upper limit of 84%. However, the model responsible for generating these outcomes demonstrated inaccurate MCC and Kappa values, rendering it unsuitable for further analysis. As a consequence, the model was excluded from consideration. In contrast, among the models considered in the second round, the model with a momentum value of 0.2, which was also employed in the initial modeling phase, emerged as the best model overall.

The figure depicting the performance of the best overall model (LR= 0.85, trained for 500 epochs and with a momentum of 0.2) was created to illustrate both successful and unsuccessful classifications. Figures 2a and 2b in the upper panels depict the pre-processing outcomes conducted by the researchers, whereas the bottom panels illustrate the display generated by the MLP classification. In the case of the satisfactory classification example, it is evident that the model successfully classified erroneous data points accurately. In the case of the poorer classification example, the model misclassified a significant number of data points that were determined to be non-anomalous and also inaccurately interpreted the duration of the anomaly.

It is important to note that the time intervals depicted in Figure 2 may not be continuous due to the removal of erroneous data points during the pre-processing stage for MLP classification. Additionally, the figure is presented in instances representing minutes of the day rather than in units of time.

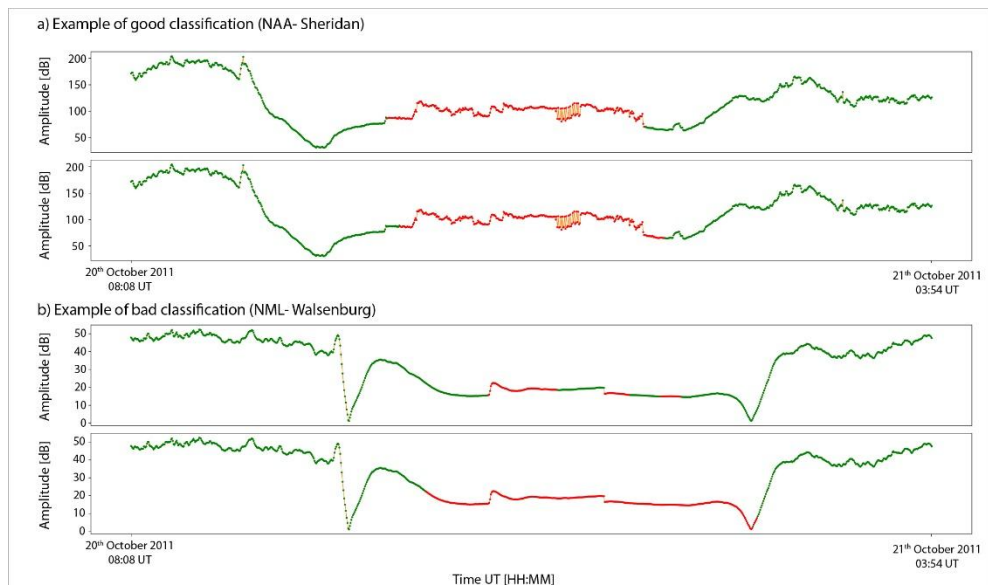


Fig. 2. Examples of good and bad classifications made by the model.

When compared to other ML methods, such as the RF model, the MLP model requires more extensive fine-tuning and computational resources. Consequently, the RF model can be considered a more favorable initial alternative for assessing the suitability and solvability of a given task using ML methods, given the available and used features and data.

Conclusions

Multilayer Perceptron models exhibit the potential for further refinement and future investigation in optimizing additional parameters. The findings suggest that the model with a LR of 0.85, 500 epochs, and a momentum value of 0.2 performed the best among all MLP models evaluated. The model chosen as best and described in the study exhibited the highest MCC and kappa values, as well as F-measure values that surpassed the average value for all models developed.

Once further research has been conducted and fine-tuning have been made, the model can be employed for the purpose of classifying ionospheric VLF amplitudes with supposed high quality. This model can also be used for future comparisons with other models and for the exclusion of future data in the analysis of ionospheric VLF amplitudes.

Acknowledgments

This work was funded by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

References

- Chicco, D. and Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics*, 21(1), pp.1-13.
- Devi, D., Biswas, S. K. and Purkayastha, B., 2020, July. A review on solution to class imbalance problem: Undersampling approaches. In 2020 international conference on computational performance evaluation (ComPE) (pp. 626-631). IEEE.
- Frank, E., Hall, M. A., Witten, I. H., 2016. The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
- Kulkarni, A., Chong, D. and Batarseh, F. A., 2020. Foundations of data imbalance and solutions for a data democracy. In *Data democracy* (pp. 83-106). Academic Press.

Prusa, J., Khoshgoftaar, T. M., Dittman, D. J. and Napolitano, A., 2015, August. Using random undersampling to alleviate class imbalance on tweet sentiment data. In 2015 IEEE international conference on information reuse and integration (pp. 197-202). IEEE.

FEATURE IMPORTANCE ANALYSIS IN RANDOM FOREST REGRESSION FOR AIR QUALITY FORECASTING IN BELGRADE, SERBIA

Filip Arnaut^{1*}, Aleksandra Kolarški¹

¹Institute of Physics Belgrade, University of Belgrade, Belgrade, Serbia; email: filip.arnaut@ipb.ac.rs

Abstract: The significance of feature importance is crucial in determining the effectiveness of a machine learning algorithm during both its training and testing stages, especially with regards to the duration of the model's training process. This brief communication conducted feature importance analysis using a Random Forest regression algorithm on air quality data collected from the measuring station Mostar, Belgrade, Serbia. The input features utilized in this study consist of PM₁₀, CO, NO₂, and SO₂ air quality parameters, whereas the target was set to the PM_{2.5} parameter. The utilized dataset consists of data collected over a three-year period (2019–2021), in which all five parameters were measured (23660 instances). The findings indicate that the model utilizing all four parameters as features exhibits the lowest root-mean-square-error (5.093 $\mu\text{g}\cdot\text{m}^{-3}$), whereas the model using only PM₁₀ and CO as input features demonstrates the highest root-mean-square error (5.509 $\mu\text{g}\cdot\text{m}^{-3}$). The differences in errors are indicative of a 7.5% disparity, whereas the duration of model training exhibited a decline of 39.6%. The analysis of feature importance revealed that a viable compromise could be achieved between the quantity of features utilized for model training and the duration of the training process, with a slight reduction in model accuracy. The conducted correlation analysis revealed that the PM_{2.5} parameter exhibited strong to very strong correlation coefficients with the PM₁₀ ($\rho_s=0.91$) and CO ($\rho_s=0.71$) parameters, while displaying weak correlation coefficients with the SO₂ ($\rho_s=0.25$) and NO₂ ($\rho_s=0.39$) parameters, which can serve as a potential explanation for the feature importance analysis results.

Keywords: Machine learning; Network analysis; Air quality data analysis; PM_{2.5}

1. Introduction

Air quality has emerged as a significant environmental and public health concern, particularly in urban centers, industrialized regions, and areas with high traffic and industrial activity. The Republic of Serbia's cities with their frequent appearance on the list of top polluted cities globally, has one possible prevalent factor which is the increased use of personalized motor vehicles. Between 2011 and 2022, a 30% increase in the total number of registered passenger vehicles can be observed (Statistical Office of the Republic of Serbia, 2011; 2022). This pattern has the potential to contribute to increased air pollution levels.

Environmental or atmospheric air pollution forecasting research is considered advantageous due to the potential to develop a reliable, stable, and accurate method for forecasting air pollution. This could provide valuable information for vulnerable populations, including individuals with chronic respiratory illnesses and elderly individuals.

*Corresponding author, e-mail: filip.arnaut@ipb.ac.rs

Recent approaches to air pollution forecasting have utilized time series forecasting (Samal et al., 2019; Shen et al., 2020; Ye et al., 2019; Zhao et al., 2018) or a combination of time series forecasting and machine learning techniques i.e., hybrid methods (Ejohwomu et al., 2022).

In the realm of feature importance or feature selection research, Sethi and Mittal (2019) have highlighted CO and NO₂ as significant parameters for air quality pollution concerning PM_{2.5}. Conversely, Mahajan et al. (2021) have emphasized that NO₂ and particulate matter are the primary contributors to the air quality index (AQI). This study utilizes a commonly employed machine learning algorithm to analyze data gathered from a high-traffic location. The objective is to predict future values and implement feature importance and selection techniques to isolate the most significant features for subsequent research.

2. Methods and data

The Random Forest (RF) algorithm was first introduced by Breiman (2001), and since has become a widely used machine learning classification and regression algorithm. A notable advantage of the RF algorithm is its minimal number of hyperparameters that necessitate tuning, with the sole hyperparameter being the number of trees. In machine learning research utilizing the RF algorithm, the determination of the number of trees can typically be achieved through iterative modeling and the minimization of a loss function in this example the out-of-bag mean squared error. Once the loss function begins to converge, the effective number of trees can be established. The present study employed JASP software (JASP Team, 2023) to automatically determine the number of trees, with the restriction that the maximum number of trees could not surpass 500.

The RF modeling procedure was conducted utilizing data that was partitioned into three distinct groups, namely the training, validation, and testing sets. The allocation for this study was divided into a ratio of 64-16-20. During the training phase, the RF algorithm utilizes input features comprising of PM₁₀, CO, NO₂, and SO₂ to “learn” of the patterns in the data with respect to the PM_{2.5} parameter, which serves as the target variable. During the testing phase, the algorithm is provided with input features and attempts to forecast the value of the target variable. Upon utilizing the RF algorithm to forecast the target variable and comparing the forecasted values with the actual values, a variety of assessment measures can be computed, including the root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2). In the context of comparing multiple RF models that were trained and tested on identical data, the evaluation metric parameters play a critical role in model interpretation. In this study, where various models were trained on the same data but with different features (some of which were excluded), the evaluation metric parameters are essential in identifying the optimal model.

The process of feature importance or feature selection involves assigning a score to each utilized feature based on a given metric. In this study, the two metrics employed were Mean Decrease in Node Accuracy (MDNA) and Total Increase in Node Purity (TINP). The straightforward interpretation of both parameters is that features with higher scores hold greater significance in effectively modeling the target variable. Regarding MDNA, the score denotes the significance of a particular feature in relation to the model's overall accuracy. Conversely, the TINP score reflects the improvement in the quality of tree splits that result from utilizing a specific feature. Modelling was undertaken with respect to the MDNA and TINP parameters. Each

successive model excluded the least informative feature from the previous model. As a cost effective or efficiency parameter the time taken to train and test the model was used.

The data utilized in this study was obtained from the Environmental Protection Agency of the Republic of Serbia (n.d.) for the air quality monitoring station Mostar, situated in Belgrade, Serbia. The monitoring station is in a heavily transport-oriented part of the city where the number of vehicles is great. The dataset utilized in this study encompasses a three-year timeframe spanning from 2019 to 2021, during which air quality was monitored. It is noteworthy that solely those instances in which all five parameters were measured were considered for the purposes of the RF modelling. A total of 23660 data points (instances) were used.

3. Results and discussion

The three RF regression models were trained with the same input data, but different features used to forecast $PM_{2.5}$ values. The initial model (RF_1) was trained utilizing all input features (SO_2 , NO_2 , PM_{10} , and CO) with a total data count of 23660 instances per feature. Subsequently, the RF_2 and RF_3 models were trained with one less feature, contingent on the feature importance analysis of the preceding model.

The RF_1 model exhibits the lowest RMSE value of $5.093 \mu\text{g}\cdot\text{m}^{-3}$ (Table 1), while the succeeding models (RF_2 and RF_3) manifest an increase in the RMSE value. It is noteworthy, however, that the most significant disparity in the RMSE value is observed between models RF_1 and RF_3, with a mere 7.5% variation. The MAE exhibits a comparable pattern to that of the RMSE, whereby the error parameter values demonstrate an increase as the quantity of features decreases. Notably, the most significant deviation is observed between the RF_1 and RF_3 models, with a discrepancy of approximately 8.8%.

The analysis of feature importance using the MDNA and the TINP for the initial model indicates that PM_{10} and CO are the most significant features, whereas NO_2 and SO_2 may be deemed insignificant. Feature importance analysis that was performed on the RF_2 model, wherein the NO_2 parameter was selected for exclusion. According to the RF_3 model, the features of CO and PM_{10} hold similar levels of significance in predicting $PM_{2.5}$ values.

Conversely, there is a notable reduction in the duration required for model training, with the RF_3 model taking approximately 39.5% less time to train than the RF_1 model. A feasible compromise may be achieved by reducing the number of features, shortening the model training time, and accepting a relatively minor rise in the model's error metrics. In this scenario, the compromise would entail accepting a model with a 7.5% reduction in accuracy while simultaneously reducing the time required to train the model by approximately 40%. The case presented in Table 1 serves as a mere illustration of a trade-off. Given that the discrepancies in the maximum time required are approximately one minute, the trade-off does not appear to confer significant advantages. Conversely, if the analysis were to incorporate a broader range of features, such as meteorological data comprising precipitation, temperature, humidity, among others, as well as other pertinent factors that could impact air quality, such as traffic information, and if the analysis were to span a time frame, the potential trade-off could prove to be highly advantageous in terms of the overall time saved for the analysis.

Table 1. Random Forest modelling results

Model	Feature	MDNA	TINP	RMSE [$\mu\text{g}\cdot\text{m}^{-3}$]	MAE [$\mu\text{g}\cdot\text{m}^{-3}$]	MAPE [%]	R^2	Time [sec]
RF_1	PM ₁₀	■	■	5.093	3.371	22.25	0.93	159
	CO	■	■					
	NO ₂	■	■					
	SO ₂	■	■					
RF_2	PM ₁₀	■	■	5.25	3.515	23.49	0.92	130
	CO	■	■					
	NO ₂	■	■					
RF_3	PM ₁₀	■	■	5.509	3.697	23.33	0.91	96
	CO	■	■					

Note. MDNA = Mean Decrease in Node Accuracy; TINP = Total Increase in Node Purity; RMSE = Root Mean Square Error; MAE = Mean Absolute Error; MAPE = Mean Absolute Percentage Error; R^2 = Coefficient of determination.

To potentially clarify the reasons behind the observed outcomes of the feature importance analysis, a correlation-based network analysis was conducted (Figure 1). The Spearman's correlation coefficient was utilized for the network analysis due to the non-normal distribution of the data, specifically its skewed nature (the skewness parameter displays values from 1.305 for the NO₂ feature to 7.354 for the SO₂ feature). The results of the network analysis indicate that the PM_{2.5}, PM₁₀, and CO parameters exhibit a significant correlation. Specifically, the PM₁₀ and PM_{2.5} parameters exhibit a very strong correlation with a coefficient of 0.91, while the PM_{2.5} and CO parameters exhibit a strong correlation with a coefficient of 0.71. In contrast, it can be observed that the parameters of SO₂ and NO₂ exhibit a limited correlation with PM_{2.5}, with coefficients of only 0.25 and 0.39, respectively.

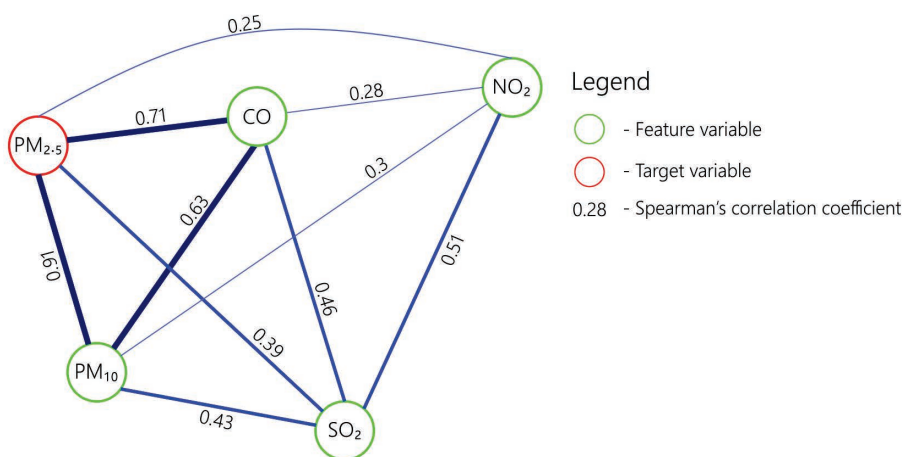


Figure 1. Modified network analysis utilizing Spearman's correlation coefficient for air quality data from the measuring station Mostar, Belgrade, Serbia (2019 to 2021).

It is noteworthy that the analysis was repeated utilizing supplementary data, including the month of the year, day of the month, and hour of measurement. The feature importance analysis, incorporating MDNA and TINP, revealed that PM₁₀ and CO, consistent with the prior example, were the top two most significant features. The third most informative feature was identified as the month of the year. The features of SO₂ and NO₂ were found to be followed by the hour of measurement and the day of the week on which the measurement was taken, which were deemed to be the least informative features added. Subsequent research should incorporate supplementary data that delineates the time of measurement, specifically the month, as demonstrated in this instance. This information can potentially provide greater insight compared to the features of SO₂ and NO₂. The model incorporating supplementary data exhibited better results in terms of lower RMSE values of 4.929 $\mu\text{g}\cdot\text{m}^{-3}$, MAE values of 2.946 $\mu\text{g}\cdot\text{m}^{-3}$, and MAPE values of 20.11%. Additionally, this model demonstrated a higher R² score of 0.934.

4. Conclusion

The utilization of machine learning algorithms for air quality modelling and forecasting has been demonstrated to be a reliable, precise, and effective approach for predicting air quality data. The findings of this study demonstrate that effective modelling can be achieved through the displacement of redundant features, such as SO₂ and NO₂ in this case, resulting in a time reduction of approximately 40%. The process of feature selection plays a crucial role in enhancing the efficiency of RF modelling when dealing with extensive datasets that may encompass variables such as meteorological and transport data. Subsequent research should acquire additional features, including the previously mentioned meteorological and transportation information, and conduct a thorough analysis of feature selection.

References

- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. <http://dx.doi.org/10.1023/A:1010933404324>
- Ejohwomu, O. A., Shamsideen Oshodi, O., Oladokun, M., Bukoye, O. T., Emekwuru, N., Sotunbo, A., & Adenuga, O. (2022). Modelling and Forecasting Temporal PM_{2.5} Concentration Using Ensemble Machine Learning Methods. *Buildings*, 12, Article 46. <https://doi.org/10.3390/buildings12010046>
- Environmental Protection Agency of the Republic of Serbia. *Air quality- Unverified hourly data in real time* [Data set]. Environmental Protection Agency of the Republic of Serbia. <http://data.sepa.gov.rs/dataset/kvalitet-vazduha>
- JASP Team. (2023). JASP (Version 0.17.1) [Computer software]. <https://jasp-stats.org/download/>
- Mahajan, M., Kumar, S., Pant, B., Tiwari, U.K. & Khan, R. (2021). *Feature selection and analysis in air quality data*. In 11th International Conference on Cloud Computing, Data Science & Engineering (pp. 280–285). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/Confluence51648.2021.9376882>
- Samal, K. K. R., Babu, K. S., Das, S. K., & Acharaya, A. (2019). Time series based air pollution forecasting using SARIMA and prophet model. In *Proceedings of the 2019 international conference on information technology and computer communications* (pp. 80–85). Association for Computing Machinery. <https://doi.org/10.1145/3355402.3355417>
- Sethi, J. K., & Mittal, M. (2019). A new feature selection method based on machine learning technique for air quality dataset. *Journal of Statistics and Management Systems*, 22(4), 697–705. <https://doi.org/10.1080/09720510.2019.1609726>

- Shen, J., Valagolam, D., & McCalla, S. (2020). Prophet forecasting model: a machine learning approach to predict the concentration of air pollutants (PM_{2.5}, PM₁₀, O₃, NO₂, SO₂, CO) in Seoul, South Korea. *PeerJ*, 8, Article e9961. <https://doi.org/10.7717/peerj.9961>
- Statistical Office of the Republic of Serbia. (2011). *Statistical Yearbook of Serbia 2011*. Retrieved May 19, 2023 from <https://publikacije.stat.gov.rs/G2011/PdfE/G20112004.pdf>
- Statistical Office of the Republic of Serbia. (2022). *Statistical Yearbook of the Republic of Serbia, 2022*. Retrieved May 19, 2023 from <https://publikacije.stat.gov.rs/G2022/PdfE/G20222055.pdf>
- Ye, Z. (2019). Air pollutants prediction in Shenzhen based on ARIMA and Prophet method. *E3S Web of Conferences*, 136, Article 05001. <https://doi.org/10.1051/e3sconf/201913605001>
- Zhao, N., Liu, Y., Vanos, J. K., & Cao, G. (2018). Day-of-week and seasonal patterns of PM_{2.5} concentrations over the United States: Time-series analyses using the Prophet procedure. *Atmospheric Environment*, 192, 116–127. <https://doi.org/10.1016/j.atmosenv.2018.08.050>

Solar Wind and Seismic Activity in the Balkan Peninsula: a 2019- 2023 Progress Report

Filip Arnaut^{1*}

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Republic of Serbia

*Correspondence: Filip Arnaut, filip.arnaut@ipb.ac.rs

Abstract: This brief progress report provides an update on the ongoing research that aims to test the statistical significance and explore the potential causal relationship between earthquakes and increased solar wind parameters, distinguishing whether these events occur randomly or are interconnected. The updated research confirmed that heightened solar wind parameters, such as proton density and velocity, cannot be differentiated from the random distribution of these events prior to earthquakes in the Balkan Peninsula region.

Keywords: proton density, proton velocity, earthquakes

Introduction

Previous research correlating solar wind parameters to global increased seismic activity displayed a positive correlation between the two (Marchitelli et al. 2020). In the study conducted by Arnaut et al. (2021), the researchers examined the correlation between heightened solar wind parameters, specifically proton density and velocity, and their potential to trigger earthquakes (EQs) in the Balkan Peninsula region. The research findings

indicate that there is no causal relationship that can be distinguished from random chance.

Inferring a causative relationship for EQs and any causing external factor is a difficult task due to the complex subsurface activities that cause them. However, several mechanisms have been proposed. An example of such a mechanism is the magnetohydrodynamical interaction between the solar wind and the Earth's magnetic field, which can influence the rotational velocity of the Earth (Simpson, 1967). This, in turn, can impact the timing of EQs and potentially lead to a shorter timeframe for their occurrence. Marchitelli et al. (2020) proposed a mechanism in which a reverse piezoelectric effect initiates subsurface currents that subsequently destabilize fault zones and induce EQs. The solar wind induces heightened subsurface currents, while the EQ is primarily tectonically driven, although its initiation is influenced by external factors.

Both mechanisms are rooted in physical principles. However, the objective of previous and subsequent studies is to determine the statistical significance of these events. The goal is to determine whether the anomalies can be differentiated from random chance when given the same input parameters. This brief progress report presents the continuation of analysis conducted in the original research paper (Arnaut et al. 2021) covering the period 1996-2018, incorporating updated data with a time span of 2019 to 2023.

Methods and data

The dataset used in this research employed EQ data from the EQ catalog of the United States Geological Survey (<https://earthquake.usgs.gov/earthquakes/search/>). The retrieved data comprised of EQs that occurred between 2019 and 2023 in the region encompassing the Republic of Serbia and its

neighboring countries. The data specifically included EQs with a magnitude of $M \geq 5.0$, which aligns with the methodology presented in Arnaut et al. (2021).

The proton density and velocity of the solar wind were acquired from OMNIWeb (<https://omniweb.gsfc.nasa.gov/form/dx1.html>), which collects data from the Solar Heliospheric Observatory (SOHO). The SOHO satellite positioned at the L1 Lagrange point (beyond Earth's magnetosphere) measures various parameters, including in situ measurements of the solar wind (Ipavich et al., 1998).

An anomalous day, as was defined in Arnaut et al. (2021), refers to any day in a given year where at least one value for proton density or velocity exceeds the mean for that year by three standard deviations. Further statistical analysis was conducted to determine which EQs exhibited an anomalous day within 14, 7, 4, or 1 day(s) prior to the EQ itself.

Results and discussion

The analysis results revealed that all EQs had an anomalous value for the proton density parameter in the 14 and 7 days leading up to it (Table 1). By contrast, the percentages from the previous analysis conducted between 1996 and 2018 were 80 and 62 percent, respectively. Regarding the four days leading up to the EQ, the findings remain relatively consistent with the previous analysis. In the sample of EQs from 2019 to 2023, 7 out of 20 EQs (35%) had an anomalous proton density day, which is in congruence to the 46% observed in the period from 1996 to 2018. Additionally, only 6 out of 20 EQs (30%) displayed an anomalous proton density value in the day preceding the EQ.

Table 1. Results of the conducted analysis for period 2019-2023

EQ date-time	Two-weeks		One-week		4-days		1-day	
	PD	PV	PD	PV	PD	PV	PD	PV
2/14/2023 13:16	•		•					
2/13/2023 14:58	•		•					
11/3/2022 4:50	•		•					
4/22/2022 21:07	•		•		•			
3/27/2021 13:47	•		•		•		•	
12/29/2020 11:19	•	•	•	•				
11/11/2020 3:54	•		•		•		•	
3/22/2020 5:24	•		•					
1/28/2020 20:15	•		•		•		•	
11/28/2019 10:52	•		•		•		•	
11/27/2019 14:45	•		•					
11/26/2019 9:19	•		•					
11/26/2019 6:08	•		•					
11/26/2019 3:02	•		•					
11/26/2019 2:59	•		•					
11/26/2019 2:54	•		•					
9/21/2019 14:15	•		•		•		•	
9/21/2019 14:04	•		•		•		•	
6/1/2019 7:00	•		•					
6/1/2019 4:26	•		•					

EQ- Earthquake; PD- Proton density; PV- Proton velocity;
 •- confirmed anomaly

From the previous research, the proton velocity parameter displayed anomalous values of 56%, 32%, and 18% for the periods 14, 7, and 4 days prior to the EQs, respectively. Only one EQ exhibited an anomalous proton velocity value during the 14 and 7 days leading up to the EQ.

Applying the same simulation procedure as conducted in Arnaut et al. (2021), it was determined that the anticipated number of EQs exhibiting an anomalous proton density within the two weeks preceding them is 19.58. This finding aligns with the observed occurrence of 20 out of 20 EQs. During the one-week timeframe, the value is 17.2, which is relatively close to the value obtained, which is also 20. The expected value for the 4-day period is 13, while the obtained value is 7. The close proximity between the expected/simulated values and the true values are so minimal that it is impossible to distinguish a random occurrence from any other possible correlation. This conclusion aligns with findings from previous research (Arnaut et al., 2021).

Conclusions

This brief progress report presents ongoing research on identifying any possible correlation between solar wind parameters (such as proton density and velocity) and EQs that have taken place in the Balkan Peninsula region. The examination of EQ and solar wind data spanning from 2019 to 2023 has reaffirmed previous obtaining that differentiating between the presence of heightened solar wind parameters and registered EQs is exceedingly challenging, as it is difficult to ascertain whether their co-occurrence is merely a result of random chance or possible causal relationship indicator. The most likely explanation is that solar wind cannot be definitively identified as a trigger of EQs registered on the Balkan Peninsula.

References

- Arnaut, F., Vučković, D., Vasiljević, I. and Cvetkov, V., 2021. Correlability of solar wind with seismic events in the Balkan peninsula zone. *Geoloski anali Balkanskoga poluostrva*, 82(2), pp.69-83.
- Ipavich, F.M., Galvin, A.B., Lasley, S.E., Paquette, J.A., Hefti, S., Reiche, K.U., Coplan, M.A., Gloeckler, G., Bochsler, P., Hovestadt, D. and Grünwaldt, H., 1998. Solar wind measurements with SOHO: The CELIAS/MTOF proton monitor. *Journal of Geophysical Research: Space Physics*, 103(A8), pp.17205-17213.
- Marchitelli, V., Harabaglia, P., Troise, C. and De Natale, G., 2020. On the correlation between solar activity and large earthquakes worldwide. *Scientific reports*, 10(1), pp.1-10.
- Simpson, J.F., 1967. Solar activity as a triggering mechanism for earthquakes. *Earth and Planetary Science Letters*, 3, pp.417-425.

Alternative Evaluation Metrics for Machine Learning Model Selection in Ionospheric VLF Amplitude Data Exclusion

Filip Arnaut and Aleksandra Kolarski

*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080
Belgrade, Serbia
E-mail: filip.arnaut@ipb.ac.rs*

When applying machine learning (ML) methods to classify ionospheric VLF amplitude data for data exclusion, there is a significant imbalance in the ML task. Specifically, the proportion of non-anomalous data to anomalous data is 85- 15% in our example. Commonly used ML evaluation metrics include accuracy, precision, F-measure, recall (Powers, 2020) among others. Standard evaluation metrics for imbalanced ML tasks can yield subpar results, requiring careful interpretation in relation to the distribution of the test dataset. This communication attests to the selection of the Random Forest (RF) model and discusses the inclusion of additional evaluation metrics, including Youden's J statistic, Markedness, General Performance Score (GPS) (De Diego et al. 2022), and Unified Performance Measure (UPM) (Redondo et al. 2020). According to Youden's J statistic, Markedness, GPS, and UPM, the previously selected model with 100 trees is the best overall model, with values of 0.692, 0.673, 0.776, and 0.833, respectively. Furthermore, the interpretation suggested that there was little difference between the models, which is supported by the additional evaluation metrics (the biggest discrepancy can be seen in Markedness at 2.1%). However, the model with 100 trees had the highest evaluation metric values and the fewest (hyper)parameters, making it the most preferable option. Additional evaluation metrics should be incorporated in further research on the utilization of ML methods for automating data exclusion, which will provide a more comprehensive understanding of the model.

References

- De Diego, I. M., Redondo, A. R., Fernández, R. R., Navarro, J. and Moguerza, J. M., 2022. General Performance Score for classification problems. *Applied Intelligence*, 52(10), pp.12049-12063.
- Powers, D. M., 2020. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), pp. 37-63.

Redondo, A. R., Navarro, J., Fernández, R. R., de Diego, I. M., Moguerza, J. M. and Fernández-Muñoz, J. J., 2020. Unified performance measure for binary classification problems. In *International Conference on Intelligent Data Engineering and Automated Learning* (pp. 104-112). Cham: Springer International Publishing.

MACHINE LEARNING APPROACH FOR DISTINGUISHING DAYTIME AND NIGHTTIME IONOSPHERIC CONDITIONS ON VLF SIGNALS RELATED TO SOLAR FLARES DURING 2011

Filip Arnaut, Aleksandra Kolarski

Institute of Physics, Pregrevica 118, 11080, Belgrade, Serbia

E-mail: filip.arnaut@ipb.ac.rs

The automated classification or labeling of ionospheric very low frequency (VLF) signals into daytime and nighttime condition categories can offer significant advantages by reducing the amount of time and effort required by researchers to manually label VLF data for further analysis and processing, e.g., in cases of studying the impacts of solar flare events on VLF data. This study employed the Random Forest (RF) classification algorithm to analyze ionospheric VLF signals during 2011, as related to several solar flare events of classes from C2.5 to X2.1 during September that were used as a training dataset (balanced) and several solar flare events of classes from C5.5 to M1.5 during October that were used as a testing dataset. The objective was to evaluate the algorithm's efficacy in automatically classifying daytime and nighttime ionospheric conditions on VLF signals. The dataset utilized in this study consisted of various combinations of transmitters (NAA, NAU, NLK, NPM, and NLM) and receivers (Oklahoma East and South, Sheridan, and Walsenburg). The data encompassed a period of three days prior to the occurrence of the solar flare, the day of the solar flare event itself, and two days following the solar flare. The study incorporated three main features: the transmitter, receiver, and VLF amplitude data. Additionally, it employed 17 secondary features, which were statistical in nature and derived from the VLF amplitude data. These secondary features primarily consisted of rolling statistics with varying window lengths. The training dataset consisted of 113926 datapoints, while the testing dataset contained 170662 datapoints. The findings derived from the complete testing dataset reveal an area under the receiver operating characteristic curve (AUC) of 0.802 and an aggregate F1 score of 0.764. Upon analyzing the data obtained from each individual transmitter-receiver pair, it is possible to draw more detailed conclusions. In accordance with the best outcome, the achieved accuracy is 0.92, with an F1 score of 0.91 and an AUC value of 0.93. The precision and Matthew's correlation coefficient (MCC) exhibit values of 0.87 and 0.85, respectively, indicating a highly effective model that can successfully be employed for the automated classification of VLF signals during both daytime and nighttime periods. Conversely, the testing data also exhibited the most unfavorable outcome, as indicated by the AUC value, which indicated the model's inability to differentiate between daytime and nighttime conditions on VLF signals. Additionally, the model demonstrated an MCC of 0, signifying a lack of correlation between the

predicted and actual binary classes. The utilization of the RF algorithm has shown potential in the automatic determination of daytime and nighttime conditions on VLF signals. However, to improve the predictive power of the model, further research and the inclusion of additional and expanded datasets are necessary.

EGU24-2046, updated on 03 Oct 2024

<https://doi.org/10.5194/egusphere-egu24-2046>

EGU General Assembly 2024

© Author(s) 2024. This work is distributed under the Creative Commons Attribution 4.0 License.



Comparative Analysis of Random Forest and XGBoost in Classifying Ionospheric Signal Disturbances During Solar Flares

Filip Arnaut, Aleksandra Kolarski, and Vladimir Srećković

Institute of Physics Belgrade, University of Belgrade, Laboratory for astrophysics and physics of ionosphere, Belgrade, Serbia (filip.arnaut@ipb.ac.rs)

In our previous publication (Arnaut et al. 2023), we demonstrated the application of the Random Forest (RF) algorithm for classifying disturbances associated with solar flares (SF), erroneous signals, and measurement errors in VLF amplitude data i.e., anomaly detection in VLF amplitude data. The RF algorithm is widely regarded as a preferred option for conducting research in novel domains. Its advantages, such as its ability to avoid overfitting data and its simplicity, make it particularly valuable in these situations. Nevertheless, it is imperative to conduct thorough testing and evaluation of alternative algorithms and methods to ascertain their potential advantages and enhance the overall efficiency of the method. This brief communication demonstrates the application of the XGBoost (XGB) method on the exact dataset previously used for the RF algorithm, along with a comparative analysis between the two algorithms. Given that the problem is framed as a machine learning (ML) problem with a focus on the minority class, the comparative analysis is exclusively conducted using the minority (anomalous) data class. The data pre-processing methodology can be found in Arnaut et al. (2023). The XGB tuning process involved using a grid search method to optimize the hyperparameters of the model. The number of estimators (trees) was varied from 25 to 500 in increments of 25, and the learning rate was varied from 0.02 to 0.4 in increments of 0.02. The F1-Score for the anomalous data class is similar for both models, with a value of 0.508 for the RF model and 0.51 for the XGB model. These scores were calculated using the entire test dataset, which consists of 19 transmitter-receiver pairs. Upon closer examination, it becomes evident that the RF model exhibits a higher precision metric (0.488) than the XGB model (0.37), while the XGB model demonstrates a higher recall metric (0.84) compared to the RF model (0.53). Upon examining each individual transmitter-receiver pair, it was found that XGB outperformed RF in terms of F1-Scores in 10 out of 19 cases. The most significant disparities are observed in cases where the XGB model outperformed by a margin of 0.15 in terms of F1-Score, but conversely performed worse by approximately -0.16 in another instance for the anomalous data class. The XGB models outperformed the RF model by approximately 6.72% in terms of the F1-score for the anomalous data class when averaging all the 19 transmitter-receiver pairs. When utilizing a point-based evaluation metric that assigns rewards or penalties for each entry in the confusion matrix, the RF model demonstrates an overall improvement of approximately 5% compared to the XGB model. Overall, the comparison between the RF and XGB models is ambiguous. Both models have instances where one is superior to the other. Further research is necessary to fully optimize the method, which has benefits in automatically classifying

VLF amplitude anomalous signals caused by SF effects, erroneous measurements, and other factors.

References:

Arnaut, F., Kolarski, A. and Srećković, V.A., 2023. Random Forest Classification and Ionospheric Response to Solar Flares: Analysis and Validation. *Universe*, 9(10), p.436.

X-RAY-INDUCED CHANGES IN NEAR-EARTH PLASMA: A MACHINE LEARNING PERSPECTIVE

F. Arnaut[✉], A. Kolarski[✉] and V. A. Srećković[✉]

*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
11080 Belgrade, Serbia*

E-mail: filip.arnaut@ipb.ac.rs, aleksandra.kolarski@ipb.ac.rs

We explored the feasibility of utilizing a multi-output machine learning algorithm to estimate ionospheric plasma parameters (sharpness and reflection height). The ionospheric plasma parameters are crucial for determining the properties of ionospheric plasma, such as electron density, rate coefficients, and cross sections for ionization/recombination processes. We examined the feasibility of employing two single-output algorithms, such as a combination of Random Forest (RF) and XGB, for the target variables. The findings revealed that during the in-sample testing phase, the multi-output model (XGB-XGB) consistently yielded the most favorable outcomes in terms of the RMSE. However, a close alternative was observed in the combination of RF and XGB models, where RF was employed for the sharpness parameter and the XGB algorithm was utilized for the reflection height parameter. During the out-of-sample validation, there was minimal variation observed among the four algorithm combinations. The most significant difference was observed between the RF-XGB and RF-RF combinations (7.6 percent decrease in RMSE).

The utilization of different algorithms and combinations of algorithms may yield marginal improvements, suggesting that the most significant improvement can be achieved through the expansion of the database.

Acknowledgments

The authors acknowledge the support from the Institute of Physics Belgrade which was made possible by grants from the Ministry of Science, Technological Development and Innovation of the Republic of Serbia.

MONITORING SOLAR FLARES AND GAMMA RAY BURSTS: MULTI-INSTRUMENTAL APPROACH INVESTIGATION

Aleksandra Kolarski[✉], Aleksandra Nina[✉], Vladimir A.
Srećković[✉], Filip Arnaut[✉]

*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
11080 Belgrade, Serbia*

*E-mail: aleksandra.kolarski@ipb.ac.rs, vlada@ipb.ac.rs,
arnaut@ipb.ac.rs*

In this paper ionospheric phenomena were studied by utilization of VLF (Very Low Frequency radio signals, 3-30 kHz) technology for monitoring of lower ionospheric plasma response to driving agents originating both within our solar system, e.g. on Sun like solar flare (SF) events, and beyond our solar system, e.g. like gamma ray bursts (GRBs), with time span that encompasses period of several years. The data on X-ray fluxes measured by Geostationary Operational Environmental Satellites (GOES) were taken from NOAA National Centers for Environmental Information database, while data on GRBs were taken from Swift satellite database. VLF signal amplitude and phase data were monitored along multiple radio signal paths using network of VLF transmitters globally positioned worldwide and received in Belgrade, Serbia by BEL receiving system, covering mainly European subcontinent. Short-termed ionospheric perturbation associated with GRBs and relatively longer-lasting ionospheric perturbations associated with

SFs were compared in order to study similarities and differences in ionospheric plasma responses to these fundamentally different driving mechanisms.

Acknowledgments

The authors acknowledge the support from the Institute of Physics Belgrade which was made possible by grants from the Ministry of Science, Technological Development and Innovation of the Republic of Serbia.

NEW MOLECULAR DATA FOR ASTROCHEMICAL MODELLING

Vladimir A. Srećković¹, Aleksandra Kolarski¹, Filip
Arnaut¹, Milan S. Dimitrijević² and Magdalena D.
Christova³ and Nikolai N. Bezuglov⁴

¹*Institute of Physics Belgrade, Pregrevica 118,
11080, Belgrade, Serbia*

²*Astronomical Observatory, Volgina 7, 11060 Belgrade 38, Serbia*

³*Department of Applied Physics, Technical University of Sofia,
1000 Sofia, Bulgaria*

⁴*Saint Petersburg State University, 7/9 Universitetskaya nab.,
St. Petersburg, Russia 199034*

E-mail: vlada@ipb.ac.rs, mdimitrijevic@aob.rs

Atomic and molecular (A&M) data and databases, which contain information about species, their identities, and processes, are critical and useful tools used in many fields of astrophysics, chemistry, and astro-informatics. Moreover methods of computational astrochemistry have become increasingly important in the last decades for the investigation of interaction and dynamics of small molecules enclosed in larger structures (Albert et al 2020, Srećković et al. 2020). In this contribution the role of some A&M processes has been studied.

Acknowledgments

The article is based upon work from COST Action CA21101, Confined molecular systems: from a new generation of materials to the

stars (COSY) and Science Fund of the Republic Serbia [Grant no. 3108/2021, NOVA2LIBS4fusion]. Authors thank N. Pop for fruitful discussions.

References

Albert, Damien, et al. (2020). A decade with VAMDC: Results and ambitions. *Atoms* 8.4, 76.

Srećković, Ignjatović, L. M., & Dimitrijević, M. S. (2020). Photodestruction of diatomic molecular ions: Laboratory and astrophysical application. *Molecules*, 26(1), 151.

Novel Research in Astrophysics and Geophysics

Vladimir A. Srećković^{1,2*}, Aleksandra Kolarski¹, Milica Langović¹,
Filip Arnaut¹, Zoran Mijić¹, Sreten Jevremović², Jelena Barović³
and Ognyan Kounchev⁴

¹ Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
Belgrade 11000, Serbia

² Scientific Society "Isaac Newton", Volgina 7, 11160 Belgrade, Serbia

³ University of Montenegro, Podgorica, Montenegro

⁴ Institute of Mathematics and Informatics, Bulgarian Academy of
Sciences, Sofia, Bulgaria

*Correspondence: Vladimir A. Srećković, vlada@ipb.ac.rs

Abstract: In the past few decades innovative approach is to foster collaboration and effective synergies among disciplines such as space exploration, atmospheric and Earth observations, laboratory and field experiments, and numerical modeling, with a high potential for direct application in Earth and other planetary research. Modeling various atmospheres with supercomputer capability, as well as diagnosing astrophysical and laboratory plasma using atomic and molecular datasets, relies on the creation and improvement of theoretical techniques and data computation methods (see e.g. Srećković et al. 2024). Multi-instrument and multi-disciplinary competence are needed to solve complicated climate concerns and its repercussions. Moreover, the growing amount of data suggests a rise in the use of automated tools and retrieval techniques (see e.g. Škoda and Adam 2020 and references therein). Model evaluation, data assimilation, satellite validation, and investigations of diverse

processes in the atmosphere and on Earth can all easily make use of the new information and retrieval products.

This contribution is progress report of work on a common topic within the bilateral project "The analysis of big data related to earth and sky observation: environmental applications and influence on life sciences" between the Bulgarian Academy of Sciences and the Serbian Academy of Sciences and Art.

Keywords: modeling, climate, multi-disciplinary investigation

Acknowledgement

This work was funded by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development, and Innovations of the Republic of Serbia. We acknowledge the support COST Action CA22162 A transdisciplinary network to bridge climate science and impacts on society (FutureMed), supported by COST.

References

- Srećković, V.A.; Dimitrijević, M.S.; Mijić, Z.R. Data in Astrophysics and Geophysics: Novel Research and Applications. *Data* 2024, 9, 32
- Škoda, P.; Adam, F. Knowledge Discovery in Big Data from Astronomy and Earth Observation; Elsevier: Amsterdam, The Netherlands, 2020

Dataset for low ionosphere modeling

Vladimir A. Srećković^{1,2*}, Aleksandra Kolarski¹, Milica Langović¹,
Filip Arnaut¹ and Sreten Jevremović²

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
Belgrade 11000, Serbia

²Scientific Society "Isaac Newton", Volgina 7, 11160 Belgrade, Serbia

*Correspondence: Vladimir A. Srećković, vlada@ipb.ac.rs

Abstract: Strong radiation has the potential to alter the structure of the Earth's atmosphere by causing more ionization. These kinds of solar activity and radiation cause sudden ionospheric disturbances (SIDs), interfere with communications from space and electrical devices on Earth, and may cause a variety of natural disasters. This work focuses on the analysis of ionosphere plasmas and their properties, as well as the investigation of SIDs utilizing very low frequency (VLF) radio signals to forecast the impact of strong radiation on Earth. The daylight atmospheric characteristics caused by this severe radiation are obtained using the model computation, which is based on all data gathered by VLF BEL stations (see Srećković et al. 2021; Šulić et al. 2016). We offer a straightforward approximation formula for electron density and an empirical model of the D-region plasma density.

Keywords: modeling, climate, VLF

Acknowledgement

This work was funded by the Institute of Physics Belgrade, University of Belgrade, through a grant by the Ministry of Science, Technological Development, and Innovations of the

Republic of Serbia. We acknowledge the support COST Action CA22162 A transdisciplinary network to bridge climate science and impacts on society (FutureMed), supported by COST.

References

- Srećković, V. A., Šulić, D. M., Ignjatović, L., & Vujčić, V. (2021). Low Ionosphere under Influence of Strong Solar Radiation: Diagnostics and Modeling. *Applied Sciences*, 11(16), 7194.
- Šulić, D. M., Srećković, V. A., & Mihajlov, A. A. (2016). A study of VLF signals variations associated with the changes of ionization level in the D-region in consequence of solar conditions. *Advances in Space Research*, 57(4), 1029-1043.

Modern society and climate change: oil and space industries perspectives

Aleksandra Kolarski^{1*}, Filip Arnaut¹, Sreten Jevremović¹,
Milica Langović¹ and Snežana Komatina²

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
Belgrade 11000, Serbia

²Technical faculty "Mihailo Pupin", University of Novi Sad, Đure
Đakovića bb, Zrenjanin 23000, Serbia

*Correspondence: Aleksandra Kolarski, aleksandra.kolarski@ipb.ac.rs

Abstract: Modern society is most often evaluated according to the presence and development stage of the space industry in a country on one hand and of the oil industry on the other. These two completely different industry branches are closely related in sense of prosperity and comfortable living style conditions of a population. However, fundamental differences between these two are most prominent in sense that space industry, primarily with infrastructure of space station and satellites, also including rocket probes and diverse space missions, is strongly influenced by and directly under space weather and space climate conditions as "external" factors, while oil industry, with its up-mid- and down-stream sectors and their infrastructures including all of the specificities and diversities, is directly impacted by social components as "internal" factors, with wide range of aspects, such as from e.g. war unstable regions of the Earth, across monetary and market trends to geopolitics. As the oil industry is most often and primarily recognized as the main culprit of accelerated and intensified climate changes, it is often forgotten that climate change is a natural phenomenon in the

geological evolution of the Earth. Although historically, oil industry with all related industries had indeed strongly impacted our living conditions, and in continual manners still does in sense of enormous pollution of all spheres of the Earth, including soil, water and air, space industry is often neglected as source of pollution even though produces massive pollutions primarily of the air and also of the soil at the place of launch in first place, but also in the space as well. As two sides of the same coin, global influences of these two industries on climate through pollution processes and related living conditions in modern societies are examined and analyzed, with main findings presented here.

Keywords: climate change, pollution, space industry, oil industry

[<https://doi.org/10.69646/aob241205>]

[*Invited Lecture*]

Annual PM_{2.5} and PM₁₀ Variations on Belgrade's Mostar Interchange – Traffic Impact

Sreten Jevremović^{1*}, Milica Langović¹, Aleksandra Kolarski¹ and Filip Arnaut¹

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Republic of Serbia

*Correspondence: jevremovic749@gmail.com

Abstract: Air pollution is becoming an increasingly pronounced problem globally, while Belgrade is often one of the most polluted cities in Europe. Some of the most common emitters and air pollutants are individual combustion plants, factory plants and traffic. In this paper, the variations of PM_{2.5} and PM₁₀ particles were analyzed in the four-year period (from 2018 to 2021) in the area of Mostar interchange in Belgrade. The results showed that there is a statistically significant difference in the levels of PM_{2.5} and PM₁₀ particles by year, which is why the influence of traffic on the increase in the level of harmful substances in the air was analyzed. The analysis showed that traffic has no influence on the change of PM_{2.5} and PM₁₀ particles at the analyzed location. The obtained results indicate that traffic, in this particular case, is not a central source of pollution, but of course it can have an impact on reducing air quality. Accordingly, the impacts of other emitters, such as individual combustion plants and factory plants, on the territory of Belgrade should be analyzed and determined in more detail, which will be the subject of future research.

Keywords: Air pollutants, air quality, PM_{2.5}, PM₁₀, traffic

[<https://doi.org/10.69646/aob241221>]

[Poster]

Novel Research in Astronomy & Earth Observation

Vladimir A. Srećković^{1*}, Zoran Mijić¹, Aleksandra Kolarski¹, Milica Langović¹, Filip Arnaut¹, Sreten Jevremović², Jelena Barović³, Ognyan Kounchev⁴ and Georgi Simeonov⁴

¹Institute of Physics Belgrade, UB, 57, 11001, Belgrade, Serbia

²Scientific Society "Isaac Newton", Volgina 7, 11160 Belgrade, Serbia

³University of Montenegro, Podgorica, Montenegro

⁴Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria

*Correspondence: vlada@ipb.ac.rs

Abstract: With a high potential for direct application in Earth and other planetary research, the innovative approach of the past few decades has promoted cooperation and productive synergies among disciplines like space exploration, atmospheric and Earth observations, laboratory and field experiments, and numerical modeling. To model atmospheres with supercomputers and diagnose astrophysical and laboratory plasma using atomic and molecular datasets, theoretical methodologies and data computation methods must be developed and improved (see e.g. Srećković et al., 2024 and references therein). To address complex climate issues and their consequences, multi-instrumental and multi-disciplinary expertise is required. As data grows, automated tools and retrieval approaches are increasingly being used (e.g., Škoda and Adam 2020). We participated in this research with our contribution.

Keywords: AstroGeoInformatics, modeling, climate, multi-disciplinary investigation

[<https://doi.org/10.69646/aob241213>]

[Lecture]

Impacts of seismic activity on human mobility

Milica Langović^{1*}, Vladimir A. Srećković¹, Aleksandra Kolarski¹, Filip Arnaut¹
and Sreten Jevremović¹

¹Institute of Physics Belgrade, University of Belgrade, Pregrevica 118,
Belgrade, Republic of Serbia

*Correspondence: milica.langovic@ipb.ac.rs

Abstract: Earthquakes are natural disasters that have led to the displacement of large number of people in different parts of the world. These geophysical events, often destructive in scale, lead to short- or long-term displacement from vulnerable areas. The aim of this research is to analyse the impact of earthquakes on the intensity and dynamics of human mobility in the Balkan region. For this purpose, data on internal displacement under the influence of earthquakes in the Balkan countries in the period 2008-2023 were analysed and processed. The results show that the earthquake is a serious push factor for human mobility and has a major influence on the decision of the population in the Balkans to migrate. It was found that the largest number of internal displacements under the influence of earthquakes was recorded in Croatia and Albania. However, it is important to point out that the intensity of displacement is directly related to the magnitude of the earthquake and the resulting living conditions and environmental damage. In order to provide a detailed overview of the different aspects of this topic, this paper emphasises the need for more detailed research on this topic in the area mentioned.

Keywords: human mobility, earthquakes, Balkan countries

<http://doi.org/10.69646/14sbac47a>

EXPLORING SOLAR RADIATION: INFLUENCE, DIAGNOSTICS, PREDICTION

VLADIMIR A. SREČKOVIĆ¹ , ALEKSANDRA KOLARSKI¹ , FILIP ARNAUT¹ 

¹*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Serbia*
E-mail: vlada@ipb.ac.rs

Since solar-influenced space weather is known to be crucial for sustainable development, researchers are particularly interested in studying extreme weather occurrences, climate change, protection, and preservation (Srećković et al. 2021). Therefore, one of the most important questions is whether we can estimate the effects of solar radiation and determine the consequences for the Earth, human population, electronics, and telecommunications (Kolarski et al. 2022). This two-part question has a very complicated answer that is not at all obvious. As a result, the study necessitates a multidisciplinary methodology in addition to the deployment of several models from diverse scientific and industrial domains. In this contribution, we analyze this thematic.





Acknowledgments: This research was supported by the Science Fund of the Republic Serbia [Grant no. 3108/2021, NOVA2LIBS4fusion].

References:

- [1] Kolarski, A., Srećković, V. A., & Mijić, Z. R. (2022). *Applied Sciences*, 12(2), 582.
- [2] Srećković, V. A., Šulic, D. M., Vujcic, V., Mijic, Z. R., & Ignjatovic, L. M. (2021). *Appl. Sci*, 11, 11574.

<https://doi.org/10.69646/aob241116>

Earth's Lower Ionosphere under energetic events: solar flares and gamma ray bursts as drivers for VLF signal perturbations

Kolarski, A.,¹  Nina, A.,¹  Srećković, V.A.¹  Arnaut, A.¹  and Langović, M.¹ 






¹*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11000 Belgrade, Serbia*

E-mail: aleksandra.kolarski@ipb.ac.rs

Earth's lower ionospheric perturbations induced by different phenomena, such as energetic solar flare events of high class on one hand and gamma ray bursts on the other, were examined through a remote sensing approach by utilization of VLF radio signals (Very Low Frequency, 3-30 kHz) propagating globally within Earth-ionosphere waveguide. Solar X-ray radiation data were taken from Geostationary Operational Environmental Satellites (GOES) database, while data related to gamma ray bursts were taken from Swift satellite database. VLF data used in this research are from Belgrade database, obtained from receiver system located at the Institute of Physics in Belgrade, Serbia, cover period of several consecutive years during solar cycle 24. Based on VLF registrations from European transmitters, temporal evolution and general patterns in VLF signal perturbations related to these events were examined, with main results related to D-region plasma behavior under such energetic events presented here.

Key words: X-ray flares, gamma ray bursts, ionosphere, VLF signal perturbation

Machine Learning Classification Difficulties of VLF Amplitude Variations Around the Terminator

Arnaud, F.,¹  Kolarski, A.,¹  Langović, M.,¹  Srećković, V.A.¹  and Jevremović, S.² 

¹*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Republic of Serbia*

²*Scientific Society Isaac Newton Belgrade, Volgina 7, Belgrade, Republic of Serbia*
E-mail: filip.arnaut@ipb.ac.rs

Previously, very low frequency (VLF) ionospheric amplitude data was employed in various machine learning (ML) classification efforts (Arnaud, Kolarski, 2023; Arnaut et al., 2023; Arnaut et al., 2024), where the problem was defined as a binary classification task, determining whether a given data point in the test dataset was anomalous (e.g., solar flare events, instrument errors, nighttime signals, outliers) or a normal daytime signal. In a multiclass classification problem, each class must be clearly defined to prevent overlap and minimize the risk of false positives and/or false negatives. The analysis for this brief communication displays that the challenges in classifying nighttime versus daytime VLF amplitude signals occur at the terminator, a transitional zone between daytime-nighttime and nighttime-daytime signals. The majority of false positives and/or false negatives occur in this zone. Possible strategies for enhancing predictive power in this area may include statistical techniques such as cluster analysis or other methods that can be employed post-classification to further augment the overall efficacy of the approach.

References

- Arnaud, F. & Kolarski, A., 2023. Machine learning approach for distinguishing daytime and nighttime ionospheric conditions on VLF signals related to solar flares during 2011. In XX Serbian Astronomical Conference, 16-20 October 2023, Belgrade, Serbia. Book of Abstracts, Astronomical Observatory of Belgrade and Faculty of Mathematics, pp. 79
- Arnaud, F., Kolarski, A. and Srećković, V.A., 2023. Random forest classification and ionospheric response to solar flares: Analysis and validation. *Universe*, 9(10), p.436.
- Arnaud, F., Kolarski, A. and Srećković, V.A., 2024. Machine Learning Classification Workflow and Datasets for Ionospheric VLF Data Exclusion. *Data*, 9(1), p.17.

PROJECT REASONING: CHARACTERIZATION AND TECHNOLOGICAL PROCEDURES FOR RECYCLING AND REUSING OF THE RUDNIK MINE FLOTATION TAILINGS

Vesna Cvetkov¹, Vladimir Simić¹, Stefan Petrović¹, Filip Arnaut², Milena Kostović¹, Dragan Radulović³, Jovica Stojanović³, Vladimir Jovanović³, Dejan Todorović³, Nina Nikolić⁴, Jelena Senćanski⁵, Grozdanka Bogdanović⁶, Dragana Marilović⁶

¹University of Belgrade, Faculty of Mining and Geology, Djusina 7, 11000 Belgrade, Serbia;
vesna.cvetkov@rgf.bg.ac.rs

²UB, Institute of Physics, Belgrade, Serbia

³UB, Institute for Technology of Nuclear and Other Mineral Raw Materials, Belgrade, Serbia

⁴UB, Institute for Multidisciplinary Research, Belgrade, Serbia

⁵UB, Institute of General and Physical Chemistry, Belgrade, Serbia

⁶UB, Technical Faculty in Bor, Serbia

Key Words: Prizma project, REASONING, Rudnik, flotation tailings

INTRODUCTION

The project "Characterization and technological procedures for recycling and reusing of the Rudnik mine flotation tailings" (REASONING) is implemented through the "Prism" program of the Science Fund of the Republic of Serbia. The primary objectives of this project are the mineralogical-geochemical characterization of the flotation tailings of the mine "Rudnik", the development of technological methods for the efficient extraction of useful metals and other elements from the tailings, and the assessment of the possibility of using this tailings as raw material for the construction industry.

The project will be implemented using geological, mineralogical, geochemical, geophysical, chemical, metallurgical and technological investigation. Researchers from six scientific research institutions from Serbia and one from Germany will participate in these complex multidisciplinary investigations.

METHODOLOGY

Tailings flotation represent material of different composition that is created in the process of extracting useful elements from the ore. This mining by-product accumulates in flotation tailings which, due to the nature of the material and the method of disposal, are designated as a potential environmental hazard. The tailings created by the processing of polymetallic Pb, Zn, Cu and Ag ore from the "Rudnik" mine is one of the most famous tailings ponds in Serbia. It is composed of fine-grained (-0.40+0.00 mm) aluminosilicate minerals quartz and feldspar (>60%), smaller

amounts of clay minerals (5-10%), while the rest is represented by Pb, Cu, Zn, Fe minerals with the presence of heavy metals up to 1%. According to the REASONING project, the measurement and tracking of changes in the physical and chemical conditions along the vertical profile of the tailings body will be carried out, the distribution of mineral associations will be determined, the magnetic susceptibility of the tailings will be measured and correlated with the content of heavy metals, along with the definition of enrichment zones, variations and trends in the composition of pore water, heavy metals and accompanying elements will be tracked. The distribution of stable mineral association and elements will be determined through theoretical geochemical modeling. An optimal scheme for the extraction of elements using a theoretical approach, electrochemical separation of elements from synthetic and leached solutions, as well as a technological procedure for the valorization of useful components will be developed.

Research will be carried out through three work packages:

1. *Mineralogical and geochemical characterization of the tailings* - physical and chemical parameters will be measured using a multi-parameter instrument and different analyzes will be applied that will provide a comprehensive understanding of changes in physical and chemical properties at different depths of the tailings. The obtained information will be important for determining the stability of mineral associations, the distribution

of metals and other elements. The research will contribute to the correlation between magnetic susceptibility and heavy metal concentration, and measurements will focus on 10 cm intervals to identify areas of enrichment in high resolution. The high resolution data obtained from this survey will provide information for subsequent analyzes (carried out at a lower resolution), related to the enriched zones. The obtained results will serve as input parameters for theoretical modeling of stable mineral associations under normal conditions.

2. Electrochemical separation and extraction of elements - includes theoretical modeling of extraction of elements from leached solution, as well as testing of electrochemical separation and extraction. The goal of theoretical modeling of elements from leached solutions is to determine the most effective methods for the extraction of these elements. The process of electrochemical separation and extraction will be carried out on prepared solutions containing the desired elements in the form of sulfates (e.g. Pb, Zn, Cu, Fe, Mn, Al, Cd, Ni) in ratios determined from chemical analyzes of pore fluids. The results of this work package will make it possible to determine the most efficient methods for the separation and extraction of elements.

3. Technological research and procedures for valorization of useful components from tailings - consist of three separate subgroups, each focused on specific procedures related to valorization of PGE (platinum group of elements) and non-ferrous metals, as well as recycling and application of flotation tailings in the road industry. Samples of 150 kg will be collected for PGE evaluation and subsequent magnetic separation, grinding, and laboratory shaking table processes. Procedures for evaluating the valorization of non-ferrous metals, will be performed on samples of 350 kg, include leaching that is carried out under natural atmospheric climatic conditions. During the experiments, the pH, salinity, temperature and concentration of elements in the solution will be monitored over time. The separation of clay minerals will be done using a sieve. The desilted sample will go through magnetic separation using different parameters of magnetic separators. Finally, physical and mechanical testing of the recycled tailings will be performed to determine its suitability for the road industry for road construction.

CONCLUSION

The proposed methodology of the REASONING project has not been previously utilized in the study of flotation tailings in Serbia. The project aims to analyze the vertical and horizontal distribution of major and microelements, as well as their geochemical association, in the flotation tailings of the "Rudnik" mine. The findings will serve as a foundation for implementing suitable technological methods to process the tailings and ensure their safe utilization in road industries.

ACKNOWLEDGEMENT

This research was supported by the Science Fund of the Republic of Serbia, 7522, – REASONING

REFERENCES

1. Abramović, F., Cvetkov, V., Ilić, A., Životić, D. (2022). Korelacija magnetnog susceptibiliteta i sadržaja metala u flotacijskom jalovištu. XVIII Kongres geologa Srbije, Divčibare, 1-4 jun 2022.
2. Gómez-García, C., Martín-Hernández, F., López-García, J.Á., Martínez-Pagán, P., Manteca, J.I., Carmona, C. (2015). Rock magnetic characterization of the mine tailings in Portman Bay (Murcia, Spain) and its contribution to the understanding of the bay infilling process. *J. Appl. Geophys.*, 120, 48-59.
3. Lam, E.J., Carle, R., González, R., Montofré, I.L., Veloso, E.A., Bernardo, A., Cánovas, M., Álvarez, F.A. 2020. A Methodology Based on Magnetic Susceptibility to Characterize Copper Mine Tailings. *Minerals*, 10, 939.
4. Lottermoser, Bernd. (2010). *Mine Wastes: Characterization, Treatment and Environmental Impacts*. Springer-Verlag, Berlin Heidelberg, Ed. 3, 404 pp.

COMPARATIVE ANALYSIS OF CORRELATION COEFFICIENTS IN MINERALOGICAL AND GEOPHYSICAL DATA FROM THE MINE TAILING SITE “RUDNIK” (SERBIA)

¹Vesna Cvetkov, ²Filip Arnaut, ¹Dragana Životić¹

¹*Faculty of Mining and Geology, University of Belgrade, Djusina 7, Belgrade, Republic of Serbia
vesna.cvetkov@rgf.bg.ac.rs; dragana.zivotic@rgf.bg.ac.rs*

²*Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, Belgrade, Republic of Serbia
filip.arnaut@ipb.ac.rs*

Key Words: Outlier detection, Non-linearity, Data analysis, Regression analysis, Magnetic susceptibility

INTRODUCTION

Quantifying the relationship between data is a crucial undertaking that can result in misinterpretation if the methods are not employed correctly. This communication highlights the disparity between the Pearson's correlation coefficient (PCC) and the Spearman's correlation coefficient (SCC) when used to assess the correlation between mineralogical, chemical, and geophysical data.

Although the PCC is readily accessible in commercially available software, it can yield inaccurate correlation estimates under certain assumptions. However, SCC is less commonly used and familiar to researchers in different fields. Nevertheless, it does not rely on the same assumptions as PCC, which makes it less prone to inaccurate correlation estimates. Therefore, it was considered advantageous to evaluate both correlation coefficients using a wide variety of data obtained from borehole measurements.

The research data was collected from the Rudnik mine tailing site, situated at Rudnik mountain in the Republic of Serbia. The dataset comprises mineralogical, chemical, and geophysical data pertaining to the samples collected from the boreholes.

METHODOLOGY

The purpose of a correlation coefficient, whether it is PCC, SCC, or any other type, is to assess the association between two or more variables. The assessment of the nature and magnitude of the relationship between two variables has been found to be a challenging endeavor, and the interpretation of these correlation coefficients has been primarily based on rules-of-thumb.

The PCC is a parametric measure of correlation that assumes the tested data follows a normal distribution and that the relationship between the data is linear. Microsoft Excel provides the CORREL function, which allows users to calculate PCC. Despite its availability in MS Excel, there are certain drawbacks associated with it, particularly with normally distributed data and the type of correlation between the data as not all data is normally distributed especially in natural sciences.

Conversely, SCC is a non-parametric measure of correlation that does not rely on assumptions about the distribution of the data or the nature of the correlation. SCC can be found in commercially available statistical software like JASP, but it is not included in MS Excel.

RESULTS AND DISCUSSION

The PCC and SCC correlation matrices, as well as the correlation matrix showing the difference between the PCC and SCC correlation coefficients (Figure 1), were calculated using mineralogical, chemical, and geophysical data. The most significant difference can be observed for arsenic and aluminum oxide, with a discrepancy of 0.84. Under those circumstances, the PCC was computed as 0.63, indicating a moderate correlation based on the scale provided by Schober et al. in 2018. However, the SCC was determined to be -0.21, indicating a significant discrepancy and a reversal in the correlation direction. Unlike the positive PCC, which implies that as one variable increases, so does the other, in this case, as one variable increases, the other decreases. The strength of the correlation in the SCC was determined to be weak, as per the scale provided in Schober et al. 2018.

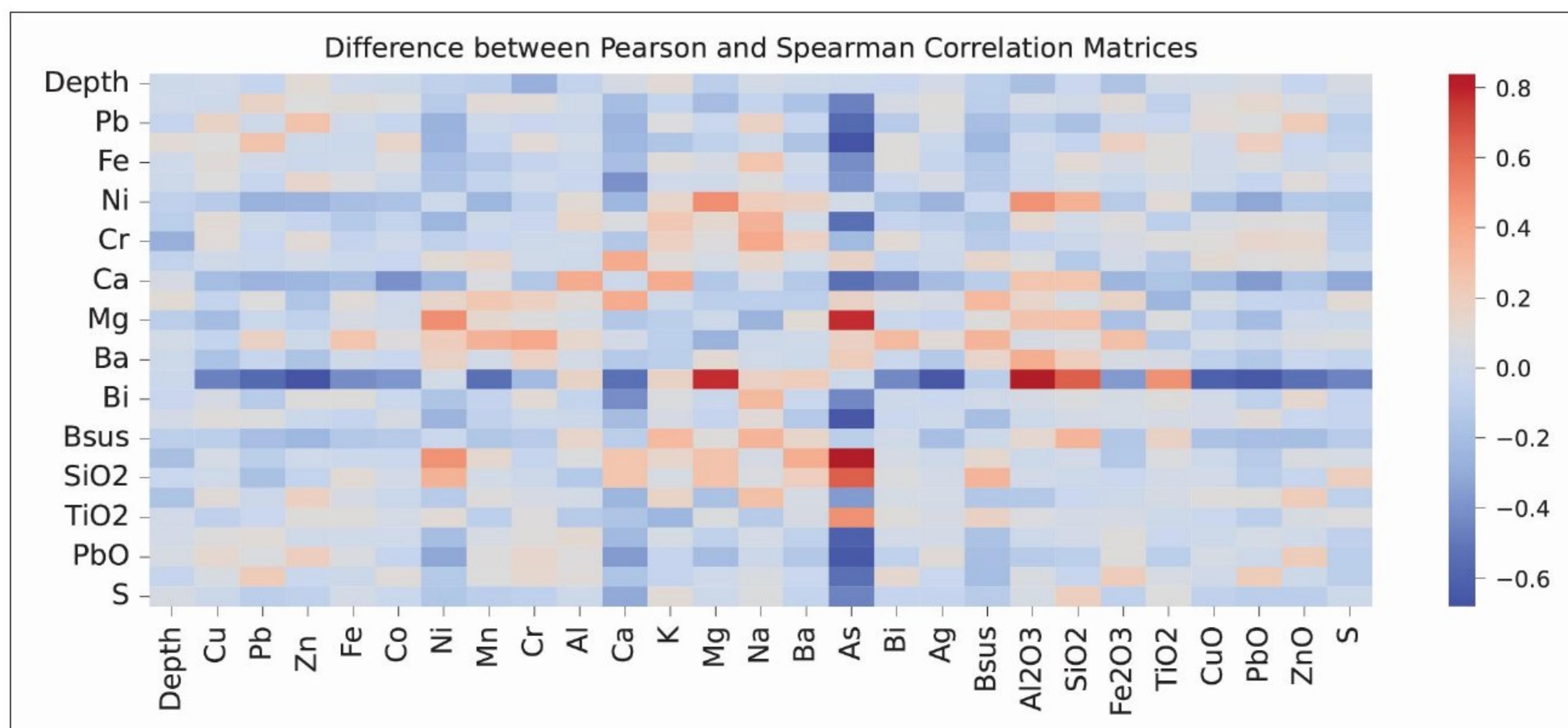


Figure 1. Selected differences between Pearson and Spearman correlation matrices.

An analysis of the cause behind the notable disparity in correlation coefficients for the same dataset revealed intriguing findings. Specifically, during the process of visualizing the data, a data point that deviated significantly from the rest of the data appeared, causing the data to be skewed. The PCC, which assumes a linear relationship between the data, failed to accurately estimate the true correlation coefficient when an outlier data point was present. In contrast, SCC, which does not make any assumptions about the distribution of the data or the type of relationship with the data, provided a more accurate estimation of the correlation coefficient.

Since this study was a preliminary investigation before further research in the field, it can be used as a reliable reference for evaluating the relationship between mineralogical, geophysical, and chemical data. Performing the concurrent computation of PCC, SCC and their discrepancy can provide a more comprehensive examination. The analysis can identify variable combinations that warrant further investigation and those that can be disregarded. If the PCC and SCC closely align, it indicates a stable correlation coefficient estimate. Otherwise, a more thorough examination is required.

CONCLUSION

Correlation coefficients are a powerful tool to assess type and strength of two or more variables relationship when utilized properly. In this brief

communication the discrepancy between PCC and SCC was displayed and a general guideline on how to proceed was also given: If the PCC and SCC are closely aligned, it suggests a reliable estimate of the correlation between the variables. Alternatively, a more comprehensive analysis is necessary.

ACKNOWLEDGEMENT

This research was supported by the Science Fund of the Republic of Serbia, 7522, – REASONING.

REFERENCES

1. Abramović, F., Cvetkov, V., Ilić, A., Životić, D. (2022). Korelacija magnetnog susceptibiliteta i sadržaja metala u flotacijskom jalovištu. XVIII Kongres geologa Srbije, Divčibare, 1-4 jun 2022.
2. Schober, P., Boer, C. and Schwarte, L.A., 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763-1768.

КАТЕГОРИЈА М50

ЗАПИСНИЦИ СРПСКОГ ГЕОЛОШКОГ ДРУШТВА

за 2020. ГОДИНУ

COMPTE RENDUS DES SÉANCES DE LA SOCIÉTÉ SERBE DE GÉOLOGIE

pour les années 2020

REPORTS OF THE SERBIAN GEOLOGICAL SOCIETY

for the year 2020

Beograd, 2020.

Арнаут Филип¹

Arnaut Filip¹

ПРИМЕНА УКРШТЕНОГ КВАДРАТНОГ ДИСПОЗИТИВА ЗА ДЕТЕКЦИЈУ ПРИМАРНЕ ЗОНЕ РАСТА КОРЕНА ДРВЕТА ВРБЕ

APPLICATION OF THE CROSSED SQUARE ARRAY FOR DETECTION OF PRIMARY GROWTH ZONE OF WILLOW TREE ROOTS

ОРИГИНАЛНИ НАУЧНИ РАД – TRAVAIL SCIENTIFIQUE – RESEARCH PAPER

Апстракт. Укрштени квадратни диспозитив користи се од половине двадесетог века за детекцију електричне анизотропије, а касније тај диспозитив улази у стандардну употребу за детекцију пукотинских система у стенама. У овом раду приказано је како се физичким моделовањем проверава применљивост овог система за детекцију примарне зоне раста корена дрвета врбе. Примењен је двојак приступ решавању оваквог проблема, у виду конструкције розети анизотропије (квалитативна анализа) и у виду срачунавања параметара анизотропије (квантитативна анализа). Резултати добијени квалитативном анализом указују на релативни правац пружања коренског система у правцу СЗ-ЈИ. Резултати добијени аналитичким поступком показују правац пружања од СИ-ЈЗ. Разлика квалитативног и квантитативног решења износи око деведесет степени. Истинит правац пружања коренског система дрвета врбе, који је измерен након изведених мерења, износи ССЗ-ЈЈИ. Квантитативна анализа података укрштеног квадратног диспозитива указала је на интензитет привидне анизотропије реда величине пар процената. Физичко моделовање показало је своје предности у виду контролisanja окружења биљке. Добијени резултати указују да се овакав вид истраживања може успешно применити у пољопривреди.

Кључне речи: Електрометријска истраживања, електрична анизотропија, физичко моделовање, мониторинг раста корена биљке.

Abstract. The crossed square array has been used for electrical anisotropy detection since the middle of the twentieth century, and later, the array is mainly used to detect rupture systems in rock formations. The crossed square array application for the primary root growth direction of the willow tree is investigated with the help of physical (analog) modelling. A two-way exploration method was used, one of which is the construction of anisotropy rose diagrams (qualitative analysis) and the second (quantitative analysis) approach. Results obtained with the qualitative analysis show the willow tree root's strike direction to be NW-SE. The findings obtained by the analytical method show the strike to be NE-SW. The difference is about ninety degrees between qualitative and quantitative analysis. The true plant root strike direction is NNW-SSE, which is determined after the end of the acquisition period. The intensity of apparent anisotropy in the order of a few percent was shown by quantitative analysis of the crossed square array data. Physical modelling displayed its advantages in the form of regulating the plants environment. A feasible implementation in agriculture is suggested by the results obtained.

Key words: Electrical resistivity survey, electrical anisotropy, physical (analog) modelling, plant root system monitoring.

¹Универзитет у Београду, Рударско-геолошки факултет, Ђушина 7, 11000 Београд
University of Belgrade, Faculty of Mining and Geology, Djusina 7, 11000 Belgrade
Email: G601-20@rgf.bg.ac.rs

УВОД

Електрометријска истраживања колинеарним диспозитивима не пружају у потпуности адекватну слику испитиваног полупростора у присуству електричне анизотропије која је у одређеној мери увек присутна. Геоелектрична сондирања колинеарним диспозитивима изнад хоризонтално услојеног полупростора са малим коефицијентом анизотропије даваће дијаграме привидне специфичне електричне отпорности (СЕО) са малим одступањима, што неће бити случај приликом појављивања косе слојевитости, шкриљавости, пукотинских система итд. Коришћење укрштеног квадратног диспозитива није ушло у стандардну праксу геоелектричног сондирања иако има многобројне предности у односу на колинеарне диспозитиве (захтева мању површину на терену, испитивања се временски брже обављају, не приказује парадокс анизотропије итд.).

Убрзан развој геофизике омогућио је њену примену у различитим областима (геотехника, археологија, хидрогеологија итд.), као и примену у последњим деценијама за сврхе агрикултуре (агрогеофизика). Примарна инспирација овог рада представља рад аутора WEIGAND & KEMNA (2016), где је приказан мониторинг раста корена биљке коришћењем методе индуковане поларизације. Концепција истраживања би се најближе могла описати као „каротажно“ мерење (истинита томографија)- због позиције електрода (са свих страна биљке). Пошто су оваква мерења на терену деструктивна и веома скупа, циљ истраживања био је да се примени укрштени квадратни диспозитив за детекцију примарне зоне раста корена биљке на минимално деструктивни начин (потпуно недеструктивно испитивање би се постигло коришћењем плочастих електрода).

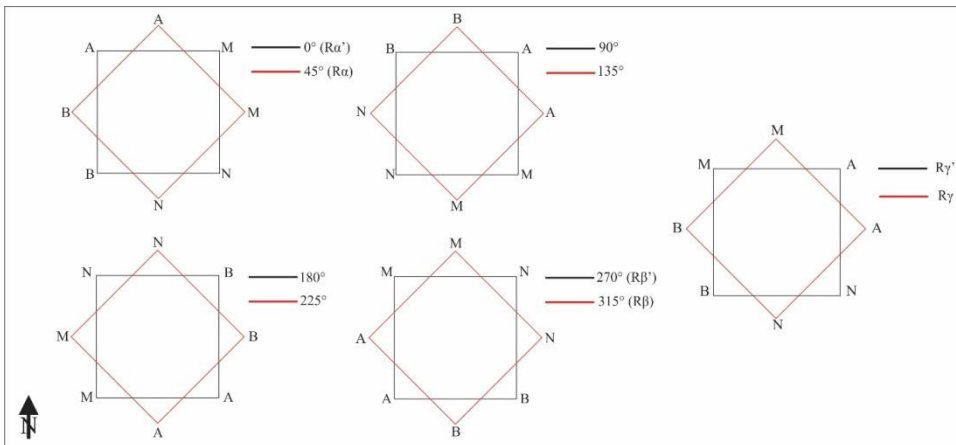
ТЕОРЕТСКЕ ОСНОВЕ

Геофизичке методе истраживања заснивају се на разлици одређених физичких својстава материјала у потповршини. У том смислу електрометрија се заснива на разлици специфичне електричне отпорности одређених природних и вештачких структура које се налазе испод површине. Варијација мерене привидне специфичне електричне отпорности на истом локалитету, али у различитом географском смеру представља електричну анизотропију, стога се за анизотропију може лако закључити да је векторска величина (има свој правац, смер и интензитет). Приликом мерења изотропног система, мерена привидна специфична електрична отпорност не треба да показује варијације у односу на положај диспозитива, односно у реалним условима варијација мерене привидне СЕО треба да буде у опсегу инструменталне грешке мерења (пар процената).

Разлози зашто слојеви у природи испољавају анизотропна својства су разнолики. Код седиментационих средина може се јављати услед транзиционих депозиционих фаза (смена трансгресивних и регресивних система депозиције), шкриљавости стена, као и пукотина које се јављају у хипсометријски вишим деловима Земљине коре, преференцијалних праваца минерала приликом кристализације итд. Претпоставка приликом аквизиције података, а и касније током интерпретације, да је потповршина изотропна, може довести до великих грешака у интерпретацији, попут погрешне процене дубина до границе слојева или погрешне

вредност CEO слоја. Да би се превазишао овај проблем (дирекционалне зависности мерења привидне CEO), HANBERJAM & WATKINS (1967) уведе укрштени квадратни диспозитив, који може да се користи и за геоелектрично картирање, као и за вертикално електрично сондирање (BHATTACHARYA & PATRA, 1968). Предности овог диспозитива су многобројне. Првенствено заузима мању површину од стандардних колинеарних диспозитива за постизање истог дубинског захвата, на њега не утиче парадокс анизотропије и мање је осетљив на латералне хетерогености, а више осетљив на присуство електричне анизотропије (SRETENović & ARNAUT, 2019).

Такав принцип аквизиције коришћен је за време трајања истраживања приказаних у овом раду и омогућио је квалитативну интерпретацију, односно приказ података на розети анизотропије. Подаци су добијени ротирањем диспозитива за 45 степени до пуног круга (Слика 1). Правац под углом, коме је приписан податак мерења, узет је у правцу од електроде В ка електроди А. Правац максималне привидне CEO се конструише, након чега се управно на тај правац поставља правац пружања пукотинског система, односно у овом случају правац пружања коренског система. Аквизиција података која је омогућила квантитативну интерпретацију укрштеним квадратним диспозитивом, врши се тако што се за један диспозитив мере три отпорности- R_α , R_β и R_γ , затим се диспозитив заротира за 45 степени и прикупе се нова мерења R_α' , R_β' , R_γ' . По завршетку овог мерења, диспозитив се у односу на централну тачку повећава за вредност $a^{1/2}$ и мерења се даље настављају по истом принципу као и претходна. Централној тачки диспозитива (на пресеку дијагонала квадрата) се приписује тачка мерења.



Слика 1. Принцип аквизиције квадратним и укрштеним квадратним диспозитивом
Figure 1. Acquisition principle with the square and crossed square array

Потребно је, такође, напоменути да се диспозитив није ротирао сваки пут, већ да је могуће мењати улоге електрода. Двојаким начином узорковања добија се већи број података, а тим и више информација о истражном простору и решавање анизотропне структуре како графички, тако и аналитички.

Добијањем укупно шест података (R_α , R_β , R_γ , R_α' , R_β' , R_γ') могуће је аналитички обрадити следеће параметре:

1. привидна CEO (ρ_α^p , ρ_β^p);
2. осредњена вредност привидне CEO из више правца (ρ_x);
3. средња вредност привидне CEO (ρ_m);
4. изотропно дефинисана привидна CEO ($\bar{\rho}$);

5. трипотенцијални однос (v);
6. одступање од модела хомогеног и анизотропног полупростора (ξ);
7. привидна анизотропија (n);
8. привидан правац пружања (Ω);
9. однос азимуталних нехомогености (A.I.R.);
10. индекс азимуталних нехомогености (A.I.I.); и
11. темпорални однос аномалија ($\Delta\rho_r$).

Начин рачунања ових параметара могуће је видети у HANBERJAM & WATKINS (1967), HANBERJAM (1975) и SAMOUËLIAN ET AL. (2005).

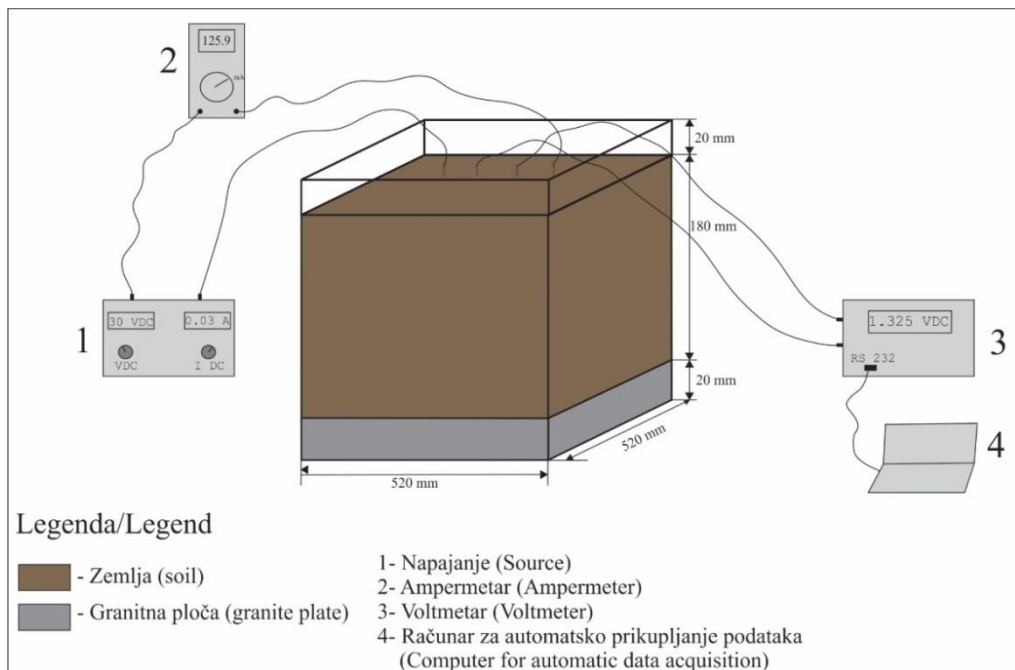
МАТЕРИЈАЛИ И МЕТОДЕ

Коришћени уређаји за мерење разлике потенцијала између пара потенцијалних електрода и предатог интензитета струје из генератора сигнала, спадају у уређаје високе резолуције, реда величине 10^{-3} . Волтметар (*Agilent 34401a*), повезан са рачунаром, врши аутоматско прикупљање података разлике потенцијала у времену.

Напонска функција генератора сигнала (*Voltcraft 2403D*) показала је релативно добру стабилност. Распон вредности (*engl. peak-to-peak*) не надмашује 0,06 V, што се може приписати грешки приликом мерења, односно генератор сигнала пружа у великој мери једносмерни сигнал средње вредности $30,7 \pm 0,03$ V.

По завршетку конструкције модела потребно је тај модел запунити земљом и направити што реалнији физички комплекс. За такву потребу коришћена је релативно хомогена и изотропна гранитна плоча исечена у димензије простора који ограничава истражни простор (плексиглас калупа). Простор изнад гранитне плоче запуњен је земљом киселог до неутралног састава (5-7 pH) укупне дебљине од 18 cm. Финални изглед физичког модела могуће је видети на слици 2.

Избор биљке за мониторинг раста корена квадратним и укрштеним квадратним диспозитивом свео се на врбу (*lat. Salix Integra Hakuro-nishiki*), због релативно мале величине и велике брзине раста корена. Биљка је смештена у центар модела око којег је ротиран квадратни диспозитив, а правац „севера“ односно правац $180-0^\circ$ произвољно је постављен.



Слика 2. Физички модел и мерни систем
 Figure 2. Physical model and measurement system

РЕЗУЛТАТИ И ДИСКУСИЈА

Резултати ће бити приказани на два начина:

1. Квалитативан - розете добијене ротацијом квадратног распореда и
2. Квантитативан (аналитички) - на основу укрштеног квадратног распореда и параметара анизотропије који су одређени аналитичким поступком НАВВЕРЈАМ-а (1975).

Квалитативно испитивање електричне анизотропије

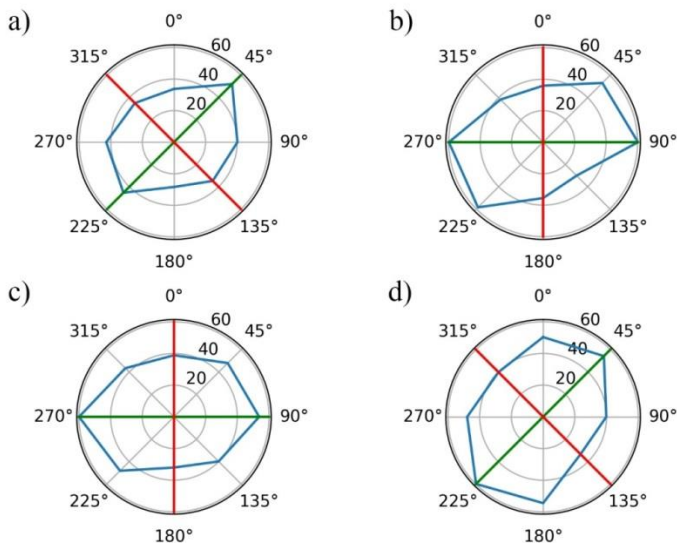
Подаци мерени 20. априла приказују правац пружања коренског система од 135-315 степени (Слика 3а - црвена линија). Срачуната грешка мерења ($\Delta\sigma$) износи 1,135 % док привидна анизотропија (n) узима вредност од 1,06. Средња привидна СЕО (ρ_m) износи 42,65 Ωm .

Мерења изведена 24. априла приказују правац пружања коренског система од 0-180 степени (Слика 3б), срачуната грешка ($\Delta\sigma$) износи 1,2 %, привидна анизотропија (n) приказује исту вредност, као и на претходно изведеним мерењима (1,06). Средња привидна СЕО (ρ_m) има вредност од 45,16 Ωm , што је веће него код претходног мерења.

Подаци прикупљени 28. априла са розете анизотропије показују правац пружања од 0-180 степени (Слика 3с), привидна анизотропија (n) износи 1,05, док је средња привидна СЕО (ρ_m) врло слична средњој привидној СЕО претходног мерења са 45,64 Ωm . Срачуната средња грешка два узастопна мерења на нивоу целе розете ($\Delta\sigma$) приказује веће вредности од претходних мерења за 2,73 %. Високе вредности

грешке мерења вероватно произилазе из промене својстава земљишта (додавање веће количине воде).

Подаци мерени 2. маја приказују исти правац пружања коренског система као мерења изведена 24. априла са правцем пружања од 135-315 степени (слика 3d). Привидна анизотропија (n) приказује највеће вредности у односу на претходна мерења са вредношћу од 1,08. Средња привидна СЕО (ρ_m) износи 46,44 Ωm , док грешка мерења ($\Delta\rho$) износи 1,28%



Слика 3. Розете анизотропије (зелена линија- правац максималне СЕО, црвена линија- правац пружања коренског система, плава линија- вредност мерене СЕО)

Figure 3. Anisotropy rose diagrams (green line- strike direction of maximal apparent resistivity, red line- strike direction of the plant root system, blue line- measured apparent resistivity)

Приказана мерења са розета анизотропије указују на правац пружања, који се налази између 315-360° и 135-180°. Због изабраног угаоног корака укрштеног квадратног диспозитива од 45 степени није било могуће тачније одредити правац пружања. Избором мањег корака био би омогућен детаљнији приказ правца пружања.

Квантитативно испитивање електричне анизотропије

Срачунате привидне СЕО као и срачунати параметри електричне анизотропије по поступку НАВБЕРЈАМ-а (1975), могу се видети у табели 1.

Табела 1. Параметри електричне анизотропије
Table 1. Electrical anisotropy parameters

Datum (Date)	a [m]	ρ_a [Ωm]	ρ_p [Ωm]	ρ_a' [Ωm]	ρ_p' [Ωm]	ρ_m [Ωm]	ρ^- [Ωm]	ρ^+ [Ωm]	ρ_s [Ωm]	Ω [°]	ξ [I]	n [I]	A.I.R. [I]	$\Delta\rho$ [I]
20. April (April 20 th)	0,15	48,87	37,22	42,85	31,29	42,65	43,05	37,07	40,06	33,71	0,0707	1,06	0,27	/
24. April (April 24 th)	0,15	52,97	38,50	50,33	35,76	45,16	45,74	43,04	44,39	32,50	0,0525	1,06	0,32	0,06
28. April (April 28 th)	0,15	46,70	44,61	52,36	39,17	45,64	45,66	45,77	45,71	42,96	0,0343	1,05	0,05	0,01
2. Maj (May 2 nd)	0,15	55,60	38,79	50,45	47,76	46,44	47,20	49,11	48,15	35,09	0,0302	1,08	0,36	0,02

Опсег варирања привидне СЕО у току аквизиције података налази се између 31,29 Ωm и 55,6 Ωm . Средња вредност привидне СЕО (ρ_m), која је оријентационо независна, приказује вредности које се налазе у опсегу од 42 до 46 Ωm . Минимална вредност срачуната је за 20. април, док је максимална вредност срачуната за 2. мај.

Вредност изотропно дефинисане привидне СЕО срачунате су за квадратни (ρ^-), као и за укрштени квадратни распоред (ρ'^-). Могуће је приметити велико поклапање вредности изотропно дефинисане привидне СЕО за квадратни и укрштени квадратни диспозитив на мерењима изведеним 28. априла. Изотропно дефинисана привидна СЕО из правца укрштеног квадратног диспозитива приказује генерално узлазни тренд. Опсег у коме се вредности изотропно дефинисане привидне СЕО налазе су од 37 Ωm до 49 Ωm . Вредности осредњене привидне СЕО из четири правца (ρ_x) налазе се између 40 и 48 Ωm , генерални тренд криве је узлазан.

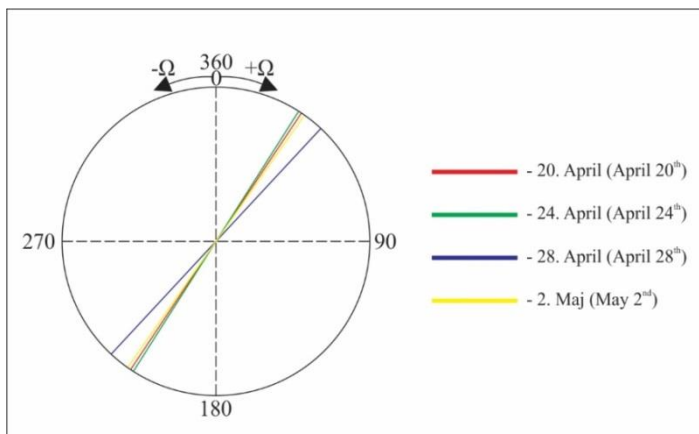
Одступање од хомогеног и анизотропног полупростора максималну вредност приказује код мерења прикупљених 20. априла са вредношћу од 0,07. Минимална вредност може се наћи код мерења изведених 2. маја и износи 0,03. Може се приметити опадајући ток вредности одступања од хомогеног и анизотропног полупростора у току трајања аквизиције података.

Однос азимуталних нехомогености (A.I.R.) за мерења прикупљена 28. априла приказује минималну вредност од 0,05. Ова вредност одступа од претходно и накнадно прикупљених вредности које се налазе у опсегу од 0,27 до 0,36. Потребно је напоменути да је претходно изведеним мерењима тог дана додата већа количина воде земљишту.

Параметар који приказује промену средње привидне СЕО у односу на референтно мерење (20. април) - темпорални однос аномалија средње привидне СЕО у току трајања аквизиције података не прелази вредности од 6 % (24. април), а минимална вредност износи 1 % (28. април).

Поређење квалитативног и квантитативног приступа аквизицији података могуће је урадити на основу претходно приказаних розета анизотропије и срачунатог привидног правца пружања (Ω). Са розета анизотропије може се видети да се примарни правац пружања коренског система налази између 135-180 степени и 315-360 степени, односно средњи правац пружања би имао правац северозапад-југоисток.

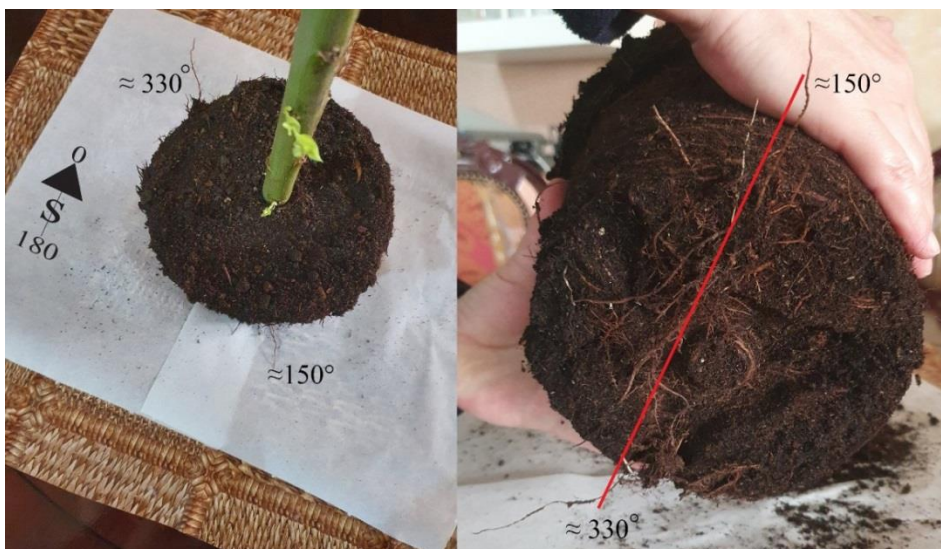
На графичком приказу срачунатог привидног правца пружања (Слика 4) може се видети разлика у односу на податке добијене квалитативном анализом података. Привидни правци пружања показују генерални правац пружања североисток-југозапад што се разликује за око деведесет степени у односу на розете анизотропије. Могу се такође приметити врло мала варирања вредности привидног правца пружања током аквизиције података. Вредности варирају у опсегу од 32 до 42 степена.



Слика 4. Срачунат привидан правац пружања коренског система
Figure 4. Calculated apparent strike direction of the plant root system

Непоклапање квалитативне и квантитативне анализе података добијеним квадратним и укрштеним квадратним диспозитивом показала су и мерења од стране геолошке службе Сједињених Америчких Држава (USGS, 2005).

Предност физичког моделовања представља могућност провере резултата након завршене аквизиције података. Приликом завршетка мерења и добрим делом оформљеног мишљења који је могући примарни правац кретања коренског система, биљка је откопана и резултати су упоређени са истинитим пружањем коренског система биљке (Слика 5).



Слика 5. Истинити правац пружања коренског система
Figure 5. True strike direction of the plants root system

Са слике 5 види се да је истинит правац пружања коренског система дрвета врбе, оквирно у правцу 150-330 степени (ССЗ-ЈИ).

ЗАКЉУЧАК

Испитивање примарне зоне раста корена дрвета врбе изведено је двојако. Квалитативан начин испитивања електричне анизотропије (розете анизотропије) показао је правац пружања коренског система СЗ-ЈИ. Велики угаони корак није омогућио детаљнију анализу пружања примарне зоне раста корена. Квантитативан (аналитички) приступ решавању правца пружања коренског система на основу примене укрштеног квадратног диспозитива показао је интензитет привидне анизотропије реда величине пар процената. Правац пружања који је аналитички срачунат разликује се од правца пружања који је добијен квалитативном анализом. Правац пружања добијен аналитичким поступком има генералан правац СИ-ЈЗ, што се у односу на квалитативне податке разликује за око деведесет степени. Истинит правац пружања коренског система, који је добијен након завршене аквизиције података износи око 150-330, односно ССЗ-ЈЈИ.

Укрштени квадратни диспозитив показао је своје предности у односу на конвенционално сондирање и томографију у присуству одређене анизотропије, али се може видети да је потребно имати и дводимензионалне секције привидне СЕО, као и моделе добијене након инверзије ради лакше визуелизације полупростора. Такви подаци би комплементирали једни друге и дали квалитетнију слику потповршинске архитектуре.

Поред секција привидне СЕО и модела добијених након инверзије, пожељно је решења тражити и у методи индуковане поларизације, као и у методи сопственог потенцијала. Овакав вид аквизиције података пружио би далеко већу спознају интеракције комплексног физичко-хемијског система какав је земља и корена дрвета биогеног порекла.

Физичко моделовање пружа бројне предности у односу на мерења изведена на терену. Главна од тих предности представља могућност контролисања окружења биљке (светлост, температура, количина воде итд.). Иако физичко моделовање представља скупљи вид истраживања, предности које пружа ту цену оправдавају.

Резултати добијени овим истраживањем се не могу узети као да су високо прецизни, већ они представљају наговештај да постоји могућност за њихову евентуалну имплементацију у пољопривреди. Поновљена мерења, за које постоји потреба да се изведу, имаће за циљ да утврде најоптималнију процедуру примене методе СЕО у одређивању примарне зоне раста корена биљака као и потенцијалне грешке које се могу јављати приликом аквизиције података. Поновљена мерења такође, са квалитетнијим подацима, показаће да ли су подаци довољно прецизни за потребе пољопривреде. Након додатних мерења, изведених у лабораторијским условима, могуће је прећи у реални физички комплекс, где се степен примене ове методе може у потпуности проучити на већим површинама.

APPLICATION OF THE CROSSED SQUARE ARRAY FOR DETECTION OF PRIMARY GROWTH ZONE OF WILLOW TREE ROOTS

INTRODUCTION

Colinear array resistivity method surveys do not provide an entirely adequate image of the subsurface in the presence of electrical anisotropy, which is often present in some form. Geoelectric soundings with colinear arrays over horizontally layered mediums with small coefficients of anisotropy will show apparent resistivity diagrams with small deviations, what will not be the case with surveys performed on oblique lamination mediums, schists, and rupture systems, etc. While the cross square array demonstrated several advantages over the traditional colinear arrays (requires less surface area, surveys are conducted in less time, is not sensitive to the paradox of anisotropy etc.), the crossed square array did not enter standard geoelectric sounding convention.

Over the last few decades, the rapid development of geophysics has allowed its application in various fields (geotechnics, archeology, hydrogeology, etc.) and agriculture (agrogeophysics). The primary inspiration for this study was WEIGAND & KEMNA'S (2016) work in which growth monitoring of plant root systems with the induced polarization method was seen. Due to the electrode location position, the principle of that research can be best defined as well-logging or "true" tomography (around every side of the plant). Because those field measurements are very expensive and destructive, the main objective of this research paper is to test the application of the crossed square array to detect the primary growth path of the willow tree root in a minimally destructive way (to obtain a truly nondestructive data acquisition system with the resistivity method, flat based electrodes must be used).

THEORETICAL FOUNDATIONS

Geophysical exploration methods are based on the difference in the subsurface medium of certain physical properties. Electrical resistivity survey methods are based on the difference in the subsurface resistivity of natural or artificial structures. On the same sounding station, but in different geographical directions, the variation of the measured apparent resistivity is electrical anisotropy. Anisotropy can be concluded to be a vector (it has direction and intensity). The calculated apparent resistivity cannot show variations on the same measurement station (with fixed depth) when measuring isotropic subsurface structures, in fact the apparent resistivity shows certain variations corresponding to the instruments measuring error (a few percent).

The reasons why anisotropic properties are manifested in certain subsurface structures are different. The explanation behind electrical anisotropy is due to transitional deposition phases (shift between transgressional and regressional deposition systems) when it comes to sedimentation environments, cleavage systems that occur in rock formations, as well as rupture systems in hypsometric higher terrains of the Earth's crust and preferential mineral orientations throughout crystallization.

The presumption that the subsurface is isotropic during data acquisition and later during data interpretation may lead to erroneous misinterpretation that can manifest itself either as an (over/under) estimation of depth to layer boundaries or an (over/under) estimation of resistivity values. HABBERJAM & WATKINS (1967) introduced the crossed square array to solve the problem of directional dependence during resistivity data

acquisition, which can be used both in geoelectric mapping and vertical electrical sounding (BHATTACHARYA & PATRA, 1968). The advantages of this array are various, first, to achieve the same investigation depth, it needs less surface area than colinear arrays, also the crossed square array is not susceptible to the paradox of anisotropy and it is less sensitive to lateral heterogeneities and much more sensitive to the presence of electrical anisotropy (SRETENOVIC & ARNAUT, 2019).

Data acquisition was carried out at 45-degree increments by rotating the crossed square array until a complete circle rotation was reached (Figure 1). This method of data acquisition enabled qualitative data analysis, i.e. the creation of rose diagrams of apparent resistivity (anisotropy). Before the beginning of the measurements, the angle direction to which the data point was attributed was set and it was the direction from current electrode B to current electrode A. The direction of maximum apparent resistivity was constructed after constructing the rose diagrams, after which the direction of the cleavage system (or in this case the root growth system direction) can be constructed perpendicular to that direction. A different data acquisition process, consisting of measuring three resistance values R_α , R_β , R_γ , then rotating the square array by 45 degrees and measuring the same three values R_α' , R_β' , R_γ' again, facilitated quantitative interpretation. The electrode spacing is increased by $a^{1/2}$ and the measurements are repeated after completing the six resistance measurements for a constant square array length. The data point is attributed to the square array's intersection of the diagonals.

It should be noted that for every measurement the array was not rotated, it is possible to adjust the "roles" of electrodes (current or potential). A larger number of data points is obtained with a two-way data sampling process, so more information is available about the subsurface, as well as the possibility to approach the anisotropic subsurface both graphically and analytically.

With a total of six data points for one depth level (R_α , R_β , R_γ , R_α' , R_β' , R_γ') it is possible to calculate the following parameters:

1. Apparent resistivity (ρ_α^p , ρ_β^p);
2. Mean resistivity from two directions (ρ_x);
3. Mean geometric resistivity (ρ_m);
4. Isotropic apparent resistivity ($\overline{\rho}$);
5. Tripotential value (v);
6. Anisotropy non compatibility ratio (ξ);
7. Apparent coefficient of anisotropy (n);
8. Apparent electrical strike (Ω);
9. Anisotropy inhomogeneity ratio (A.I.R.);
10. Anisotropy inhomogeneity index (A.I.I.); and,
11. Temporal anomaly ratio ($\Delta\rho_T$).

In HABBERJAM & WATKINS (1967), HABBERJAM (1975) AND SAMOUËLIAN ET AL. (2005), the theory behind the presented parameters is given.

MATERIALS AND METHODS

The instrument used to measure the potential difference between two potential electrodes belongs to instruments of a high resolution class which measure potential and current differences in the order of 10^{-3} . The voltmeter (Agilent 34401a) was attached to a computer, which automatically samples possible potential differences in time.

The signal generator voltage function (Votcraft 2403D) showed good time stability, the peak-to-peak value does not exceed 0.06 V, which can be due to error of measurement. In the majority of cases, the signal generator provides a direct current signal with a value of 30.7 ± 0.03 V.

The model has to be filled with soil after building the physical model, in order to get it closer to a real physical complex. A relatively homogeneous and isotropic granite plate, which was cut into the dimensions of the plexiglass mold, was used for that need. The volume above the granite plate was packed with soil that was neutral to acidic (5-7 pH) and had a total thickness of 18 cm. The physical model's final appearance can be seen in figure 2.

Due to its small size and high root growth velocity, the willow tree (*lat. Salix Integra Hakuro-nishiki*) was the preferred plant for monitoring with the crossed square array. In the physical model around which the square and the crossed square array were rotated, the plant was positioned in the centre. The north direction was established arbitrarily.

RESULTS AND DISCUSSION

The results will be given as:

1. Qualitative analysis- rose diagrams obtained with the rotation of the square array and,
2. Quantitative analysis-parameters that were determined using the HABBERJAM (1975) method and the crossed square array.

Qualitative analysis of electrical anisotropy

Data measured on April 20th shows the direction of crop root growth at around 135-315 degrees (Figure 3a- red line). The measurement error estimated ($\Delta\sigma$) is 1.135 percent, while apparent anisotropy (n) takes values of about 1.06. The apparent mean resistivity is 42.65 Ωm .

Measurements taken on April 24th show the direction of growth of the plant root strike at 0-180 degrees (Figure 3b), measurement error ($\Delta\sigma$) is estimated at 1.2 % and apparent anisotropy (n) has not changed since the first measurements four days before (1.06). Mean apparent resistivity (ρ_m) of about 45.16 Ωm displays an increase with regard to previous measurements.

Data obtained on April 28th and graphically presented on a rose diagram suggest that the direction of growth has not changed from the measurements carried out on April 24th (0-180 degrees, Figure 3c). The apparent anisotropy (n) is 1.05, while the mean resistivity has decreased moderately and values are about 45.64 Ωm . The measurement error ($\Delta\sigma$) estimated from two consecutive measurements is 2.73 %. The higher measurement error values can possibly be due to the change in soil properties (more water was given to the plant that day).

Measurements obtained on 2nd of May display the same values for the direction of root growth as those taken on 20th of May, 135-315 degrees (Figure 3d). Apparent anisotropy increased to 1.08, while apparent mean resistivity increased to 46.44 Ωm . The calculated measurement error is 1.28 %.

A relative root growth direction that lies between 315-360 degrees and 135-180 degrees is shown by data estimated over a two-week period and graphically illustrated on anisotropy rose diagrams. A more detailed direction of growth was not plausible to evaluate because of the selected angle step of the crossed square array of 45 degrees. When performing these types of research, it would be suggested to choose a smaller angle step, such as 30 or 15 degrees, to obtain a more detailed subsurface image.

Quantitative analysis of electrical anisotropy

Calculated apparent resistivities and parameters of electrical anisotropy obtained by the HABBERJAM (1975) method can be seen in Table 1.

During the acquisition period, the range of apparent resistivity variation lies between 31.29 Ωm and 55.6 Ωm . Mean apparent resistivity, which is directionally independent, indicates values that can be found between 42 and 46 Ωm . The minimum value was calculated on April 20th, while the maximum value was measured on May 2nd.

Isotropically defined apparent resistivity values calculated for the square and crossed square array can be observed on April 28th to show broad overlaps. An upward trend is demonstrated by isotropically defined apparent resistivity measured for the square array. The range of variation in apparent resistivity is between 37 Ωm and 49 Ωm . The mean apparent resistivity values obtained from four directions have a resistivity range between 40 and 48 Ωm .

The anisotropy non-compatibility ratio indicates a maximum value of 0.07 during the measurements performed on 20th of April. A minimum of 0.03 can be found during the measurement acquired on May 2nd. The general trend during the measurement time of the anisotropy non-compatibility ratio is downward.

For measurements performed on April 28th, the anisotropy inhomogeneity ratio showed a minimum value of 0.05. This value deviates from those between 0.27 and 0.36 that were previously and subsequently obtained. It should be noted that these values were obtained by measurements carried out on 28th of April after more watering the soil than had been previously done.

Parameter showing the shift in apparent resistivity relative to the reference measurement (April 20th)-Temporal apparent resistivity anomaly varies from 1 % (April 28th) to 6 % (April 24th).

A distinction of both approaches can be made using the anisotropic rose diagrams (qualitative) and the measured apparent strike (quantitative). Rose diagrams of anisotropy display a relative strike direction between 135-180 and 315-360 degrees, i.e. the mean strike direction will be in the direction of NW-SE.

The apparent strike data can be shown graphically (Figure 4) where a divergence from the data used to make up the qualitative analysis is seen. The apparent strike direction indicates a general SE- SW strike direction, which varies by approximately ninety degrees from the qualitative interaction. During the acquisition period, minor variations of the apparent strike can be seen only looking at the apparent strike direction graph. The values range from 32 to 42 degrees.

The United States Geological Survey research team also showed the disparity between qualitative and quantitative analysis obtained by the crossed square collection (USGS, 2005).

The main advantage of physical (analog) modelling is the verification of the result after the measurement period is over. The plant was excavated after completing the measurements and having a reasonably clear understanding of what would be the main direction of growth of the willow tree root, and the findings were compared with the true strike direction of the tree root (Figure 5).

The true root growth direction of the willow tree of approximately 150-330 degrees is apparent in figure 5 (NNW-SSE).

CONCLUSION

Two techniques were used to evaluate the primary growth direction of the willow tree. The strike direction of the NW-SE was shown by qualitative study of electrical anisotropy (rose diagrams). A more thorough study could not be made possible by a large angular step. The intensity of apparent anisotropy in the range of a few percent was shown by a quantitative (analytical) method based on crossed square array measurements. The measured strike direction varies from the direction of the qualitative strike. The strike direction obtained by the crossed square array has a NE-SW direction that deviates by ninety degrees from the qualitative strike direction. The true direction of growth of the willow tree root, which was evaluated after measurement completion, is approximately 150-330 degrees, i.e., NNW-SSE.

In the presence of electrical anisotropy, the crossed square array showed its advantages from traditional sounding and tomography, but both 2D apparent resistivity profiles and models obtained after inversion could improve the subsurface visualization. The crossed square resistivity data would be complemented by that form of data and give a better insight of the subsurface architecture.

In relation to field measurements, physical (analog) modelling offers various benefits. The main benefit is regulating the plant environment (light, temperature, water content etc.). Even though physical modelling is a more expensive form of research, the advantages that it gives outweigh the cost.

The findings provided in this research paper cannot be perceived as being highly accurate and indicate that it is possible to apply them in agriculture. The objective of repeated measurements, which are needed, will be to improve the most efficient technique of applying the resistivity method in the detection of the tree root's primary growth direction. Repeated measurements of better-quality data most likely would demonstrate that the data is reasonably reliable for the application in agriculture. After repeated tests under laboratory conditions, it would be possible to translate it into a real physical complex, where the degree of application will be checked in substantial areas.

ЛИТЕРАТУРА- REFERENCES:

- BHATTACHARYYA, P. K., PATRA, H. P. (1968): *Direct current geoelectric sounding: principles and interpretation*. Elsevier
- GEOLOGICAL SURVEY (USGS). (2005): *Square-Array Direct-Current Resistivity Measurements Conducted at Nye County near Borehole NC-EWDP-29P*. Nye County Nuclear Waste Repository Project
- HABBERJAM, G. M., WATKINS, G. E. (1967): *The use of a square configuration in resistivity prospecting*. Geophysical prospecting, 15(3): 445-467
- HABBERJAM, G. M. (1975): *Apparent resistivity, anisotropy and strike measurements*. Geophysical prospecting, 23(2): 211-247
- SAMOUELIAN, A., COUSIN, I., TABBAGH, A., BRUAND, A., RICHARD, G. (2005): *Electrical resistivity survey in soil science: a review*. Soil and Tillage research, 83(2): 173-193
- SRETENOVIC, B., ARNAUT, F. (2019): *Inadequacy of 1D, 2D and 3D Resistivity Inverse Modelling in the Presence of Electrical Anisotropy*. Earth, 8(2): 102-116
- WEIGAND, M., KEMNA, A. (2017): *Multi-frequency electrical impedance tomography as a non-invasive tool to characterize and monitor crop root systems*. Biogeosciences, 14(4)

Bulletin of the Natural History Museum, 2021, 14: 45-61.

Received 1 Nov 2021; Accepted 11 Nov 2021.

doi:10.5937/bnhmb2114045A

UDC: 902.2:550.837(497.11)

903.4"634"(497.11)

Original scientific paper

2D ELECTRICAL IMAGING INVESTIGATIONS AT THE NEOLITHIC SETTLEMENT "PLOČNIK"

FILIP ARNAUT, BRANISLAV SRETENović

University of Belgrade, Faculty of Mining and Geology, Department of
Geophysics, Đušina 7, Belgrade, Serbia

G601-20@rgf.bg.ac.rs, branislav.sretenovic@rgf.bg.ac.rs

At the Neolithic settlement "Pločnik" near Prokuplje in Serbia, the electrical imaging method (ERT) was deployed. At the sounding site S-3, 1D electrical survey revealed a low resistivity layer with a thickness of 0.8 meters and a resistivity of about 30 Ωm . A sharp and steep boundary divides the high resistivity medium and the low resistivity medium, according to the 2D electrical survey. The resolution of the ERT method was tested using 3D forward and inverse modeling to see whether it could distinguish the high resistivity pit from the surrounding high resistivity gravel.

Key words: Electrical imaging (ERT), Neolithic settlement "Pločnik", Forward modelling, Inverse modelling.

INTRODUCTION

The archaeological Neolithic settlement "Pločnik" is located near the town of Prokuplje in central Serbia, covering an area of about 80 hectares

and belonging to the well-known "Vinča" culture. Earlier archeological investigations, which were carried out in the traditional way by digging archeological trenches at several sites, revealed the presence of a large settlement and, more importantly, the possibility of metallurgy operations. The discovery of various copper tools and jewelry, as well as copper ores and minerals (chalcopyrite, malachite), is linked to the open-profile conical and cylindrical pits with diameters up to 9 meters and depths of 3.5-4 meters at the vertical cliff. These pits have been filled with ash deposits, most likely as a result of copper processing. Smaller and shallower pits (with diameters up to 3 meters and depths of 1.5-3 meters) are also present, often containing animal bones, pottery fragments, and anthropomorphic terracotta (Vukadinovich *et al.*, 1999).

Given the size of the potential targets, their burial depths, and the potential contrast of their physical properties with the surrounding material, it was agreed to use electrical mapping to detect relatively large pits filled with ash. The geophysical methods used successfully identified the most important anomalous zones and significantly reduced the region available for possible archeological, 2D, and 3D geoelectrical scanning investigations.

LOCATION AND GEOLOGICAL SETTING OF THE SURVEY AREA

Figure 1a depicts the survey area, which includes two sublocalities (Pločnik I and Pločnik II) where the survey was conducted. The findings from the Pločnik I locality will be presented in this research paper. Figure 1b shows an archeological profile discovered as a result of coastal erosion along the river Toplica. The existence of a pit with a radius of a few meters and a depth of about 4 meters can be clearly seen.

The vertical cliff formed by the river flow reveals a basic geological structure consisting of an almost homogeneous loess-like, four- to five-meter-thick clayey layer that covers the alluvial gravel of the Toplica river. This fact greatly improves the possibility of successful application of the geophysical methods by reducing the geological interference in delineating anomalies caused by archeological remainders.

1D GEOELECTRICAL SURVEY RESULTS

The inquiry was approached in a variety of ways at the survey location (Figure 2). Geoelectrical mapping, as well as 2D electrical imaging and vertical electrical sounding (VES) were deployed. There were twelve parallel geoelectrical mapping profiles (I-XII), four electrical imaging

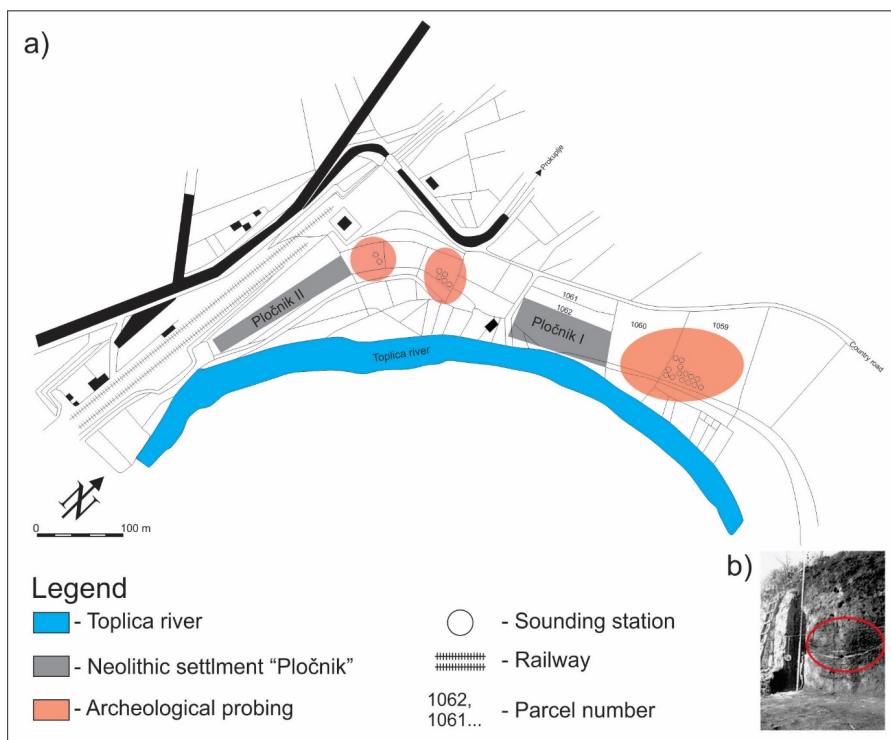


Fig. 1. – (a) The location of the survey area; (b) the archeological profile with the marked pit location.

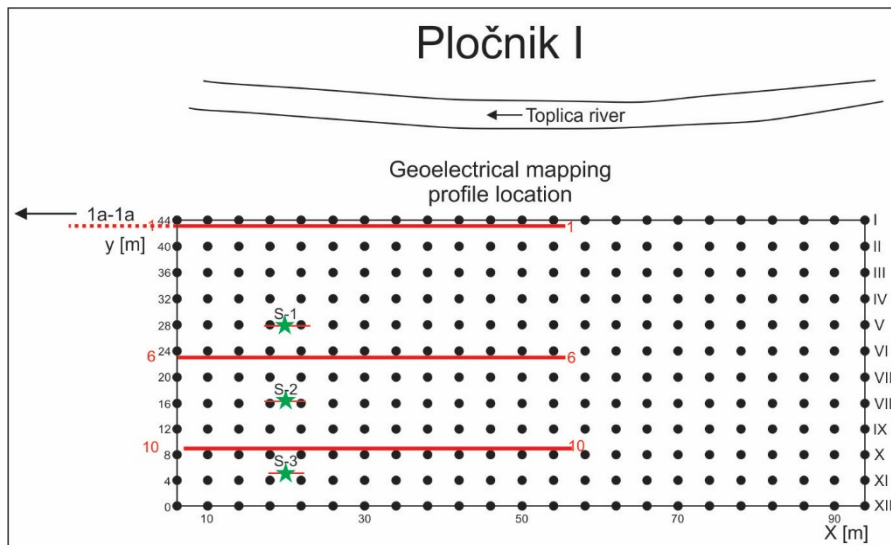


Fig. 2. – Locations of geoelectrical mapping profiles (I-XII), 2D electrical mapping profiles (1-1, 1a-1a, 6-6, 10-10) and geoelectrical sounding locations (S-1, S-2, S-3).

profiles (1-1, 1a-1a, 6-6, 10-10), and three VES locations used in total (S-1, S-2 and S-3).

The presence of a thin, low resistivity layer that overlays a high resistivity substratum is indicated by 1D interpretation of measured geoelectrical sounding curves with a relatively shallow depth of investigation ($AB/2_{max} = 20$ m) (Figure 3). The current electrode crossing over local near surface inhomogeneities causes distinct lateral effects in the geoelectrical sounding curve S-1, particularly in the near surface part of the diagram ($AB/2 = 1.5$ m to $AB/2 = 3$ m). As a result, 1D inversion is not used to interpret the S-1 sounding curve.

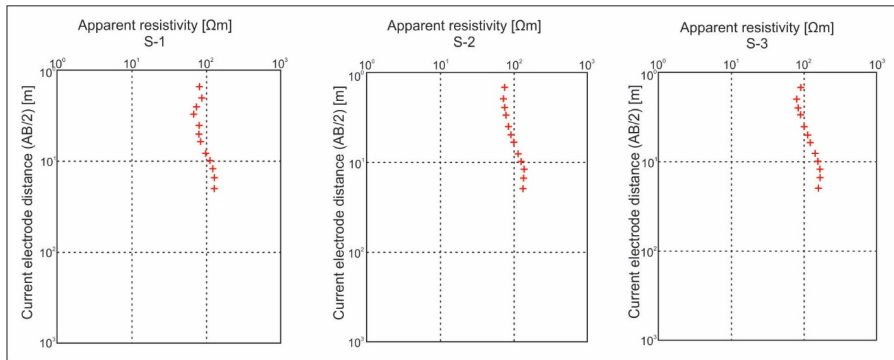


Fig. 3. – Measured sounding curves S-1, S-2 and S-3.

The sounding curve S-3 was interpreted in two different ways, yielding two equivalent models that are shown in Figure 4. In both cases, a low resistivity layer with a thickness of 0.8 meters and a resistivity of about 35 Ωm is present.

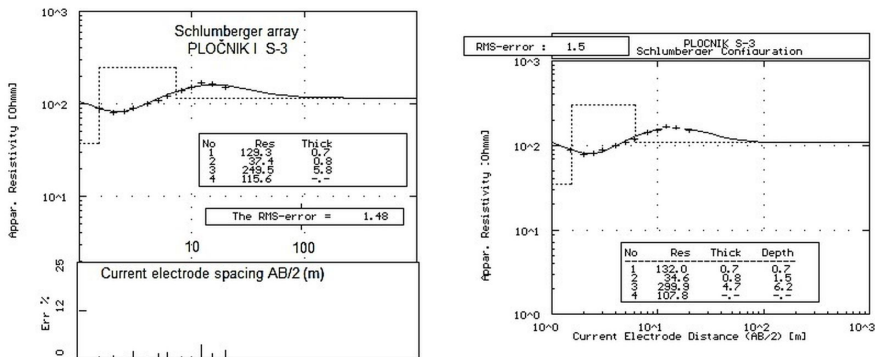


Fig. 4. – Equivalent 1D models obtained after inversion for the sounding curve S-3.

The measurement arrays were oriented parallel to the strike of the vertical cliff formed by Toplica River erosion (with a height of 4 to 5 meters). At sounding station S-3, the cliff's effect is the smallest. The

topography effect would first appear as an increase in measured apparent resistivity values as the Schlumberger array dimensions were increased, followed by a descending trend (Figure 5). In practice, the topographic effect will produce effects similar to the change in resistivity with depth, or a K-type sounding curve. The magnitude of this effect is determined by the cliff height-to-sounding-station-distance ratio (H/X), as well as the cliff angle (the highest being for $\alpha = 90^\circ$). The topographic effect is practically absent as the distance between the sounding station S-3 and the cliff increases (around 40 meters) (Figure 5). Holcombe & Jiracek (1984)

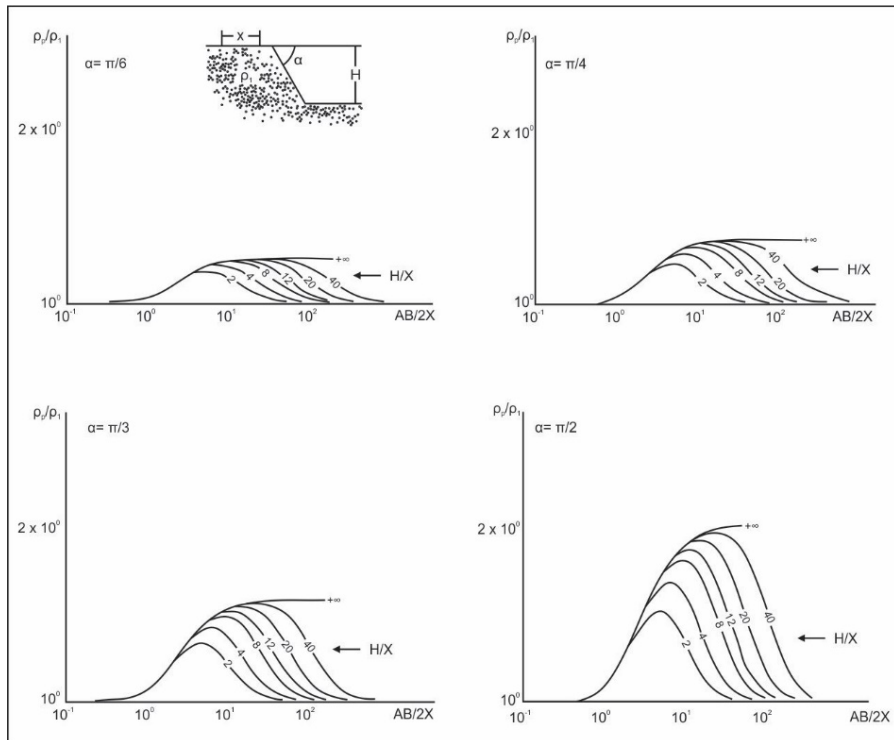


Fig. 5. – Topographic effect of the cliff (array orientation is parallel to the cliff strike).

developed a three-dimensional numerical model that demonstrated the possibility of removing the topographic effect from resistivity data.

2D GEOELECTRICAL SURVEY RESULTS

The apparent resistivity map (Figure 6) showed two distinct resistivity regions, which were apparently separated by a steeply dipping boundary (in the zone of 25-35 meters). The contact between the lower resistivity zone

(yellow zone with resistivity values of 55 to 70 Ωm) and the higher resistivity zone (red zone with apparent resistivity values up to 140 Ωm) is essentially two-dimensional. Three anomalies are indicated within the high resistivity region, indicating the existence of 3D structures. On the apparent resistivity map, the topographic effect of the vertical cliff can be seen as an increase in apparent resistivity values in the vicinity of the cliff (upper part of the map, blue rectangle). The apparent resistivity map was also used to position the 2D electrical imaging profiles (1-1, 1a-1a, 6-6 and 10-10 in Figure 2) for thorough inspection of the most interesting zones (3 closed anomalies) and 2D interpretation of the apparent resistivity pseudosections.

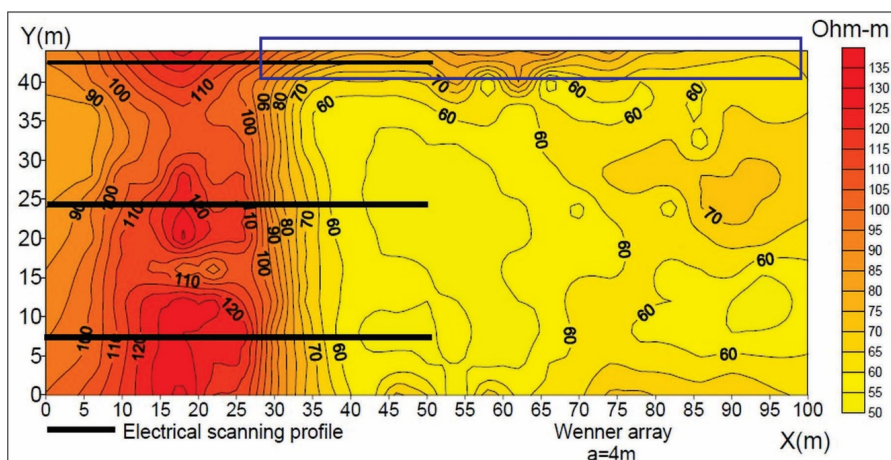


Fig. 6. – Apparent resistivity map constructed with the grid given at Figure 2 (measurements are attached to the midpoint of every Wenner array during the geoelectrical mapping, the length of the Wenner array is $a=4\text{m}$, with the mapping step $\Delta X=4\text{m}$).

2D electrical imaging method with 26 electrodes (unit electrode spacing $a = 2\text{m}$) was used to further investigate the contact between the two regions, as well as to obtain information about the depth extent and shape of the structures causing 3D anomalies. Surveying was first carried out at five depth levels, along three parallel profiles that were set perpendicular to the 2D contact. Despite the fact that the structures within the high resistivity region are 3D, the results of 2D interpretation were satisfactory (Figure 7). The resistivity of the sources of anomalies is several hundred Ωm , according to two-dimensional models. The highest true resistivity values, as defined by profile 1-1, are influenced in part by the presence of a vertical cliff, as shown on the apparent resistivity map (Figure 6). The same value scale was used for all three profiles during the construction of the 2D

models, which influenced the appearance of the 2D models with lower resistivity ranges (profiles 6-6 and 10-10).

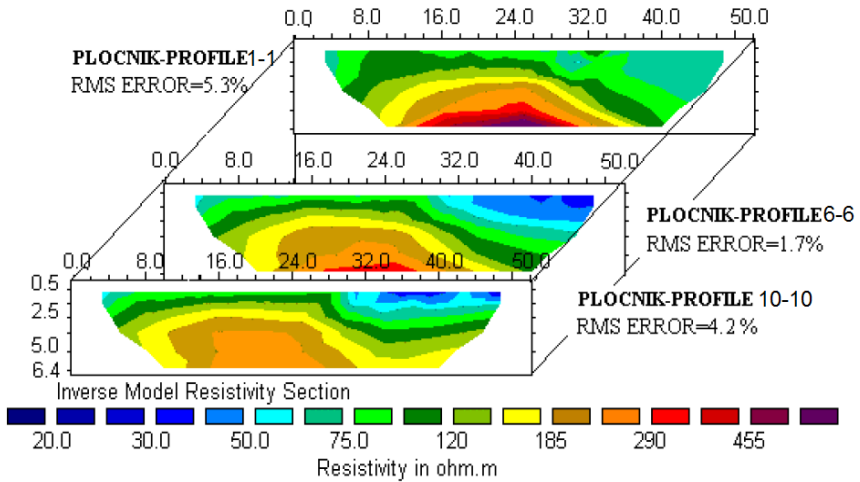


Fig. 7. – Results of 2D inversion along three parallel profiles (1-1, 6-6 and 10-10).

Figure 8 displays the result of 2D inversion along the electrical scanning profile 6-6 with a value scale that better illustrates the subsurface architecture.

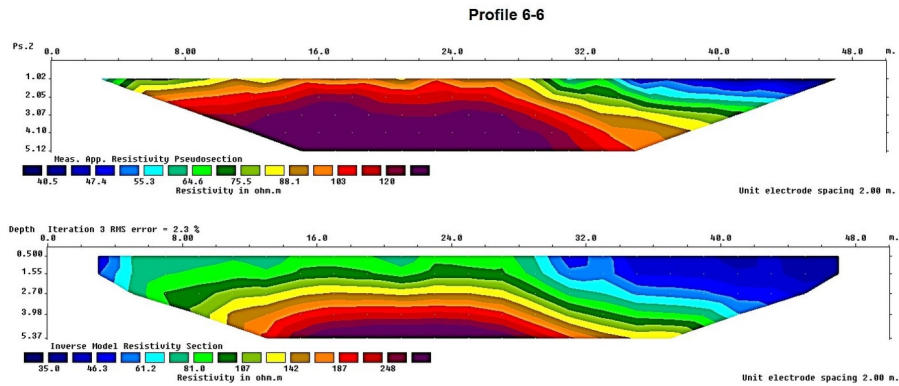


Fig. 8. – Result of 2D inversion along profile 6-6 (bottom) and the measured apparent resistivity pseudosection (top).

Since the results show that the high resistivity zone sinks before the start of the 2D profile, additional measurements, such as the extension of profile 1-1, were taken (Figure 9). In comparison to the subsurface

architecture shown along profile 1-1, the extended results indicate a much simpler subsurface model (close to 1D conditions).

Pločnik I- 2D Inversion of profile 1-1 and 1a-1a

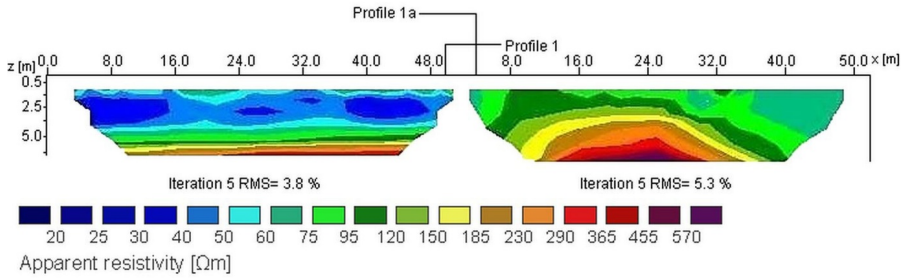


Fig. 9. – Results of inversion along profile 1-1 and its extension 1a-1a.

2D FORWARD MODELLING AND MODEL REFINEMENT

Figure 10 shows two possible models for the entire profile 1-1 and 1a-1a, both of which include the high resistivity pit. An apparent resistivity section was calculated using the forward modeling method (in RES2DMOD, Loke 2016b). This section was later used to invert calculated data and to obtain a 2D model using these calculated data as "measured" apparent resistivity data (in RES2DINV, Loke 2016a).

Based on the analysis that contains forward and inverse modelling it is clear that pits, regardless of their geometry and dimensions cannot be differentiated from the higher resistivity substratum even when they are located at relatively shallower depths and have a high resistivity contrast relative to the surrounding medium. The presence of the pit unambiguously leads to distinct anomalies relative to the simple, 1D geological profile that has been detected along profile line 1a-1a.

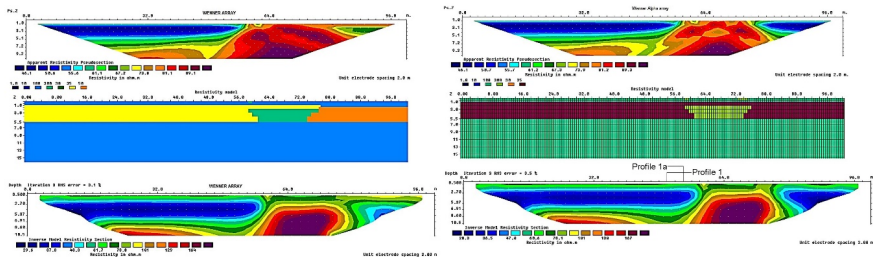


Fig. 10. – Two possible models for the combined profiles 1-1 and 1a-1a (middle), apparent resistivity sections (top) and 2D models obtained after inversion (bottom).

The option "Refine model" was used to perform a 2D inversion along profile 6-6, which allows for a more detailed investigation of near-surface inhomogeneities. The cell width is half the unit electrode spacing when this option is selected. Figure 11 shows the effects of that inversion process when data from all five depth levels are used (Figure 11, top) and when data from just four depth levels are used (Figure 11, bottom). The use of four depth levels is intended to help distinguish the pit from the high resistivity substratum (gravel from the Toplica River).

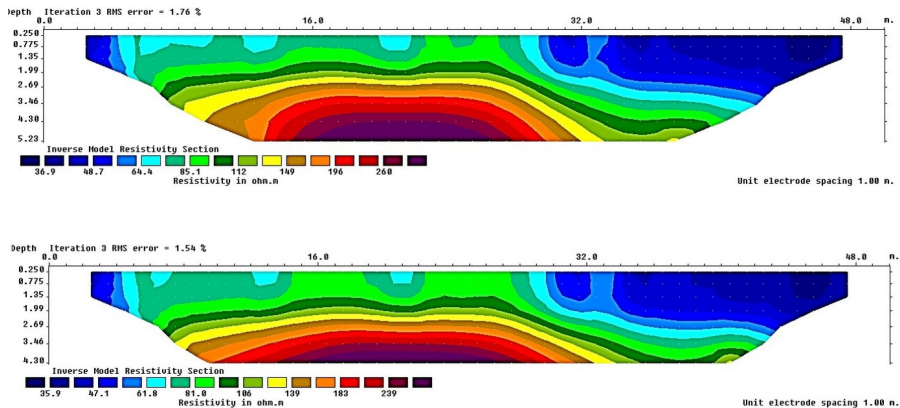


Fig. 11. – 2D inversion of profile 6-6 with the “Refine model” option enabled.

In order to focus on the zone containing the high resistivity anomaly, a 2D inversion of a symmetrically shortened profile 6-6 (first electrode at the 6th meter, last electrode at the 44th meter) was performed. In this case, the “Refine model” option was used, but it did not clearly distinguish the pit from the high resistivity substratum (Figure 12). The usage of the "Refine model" is responsible for the high resistivity surfacing. This option can also be used when looking for other archeological artifacts (other shallower pits etc.).

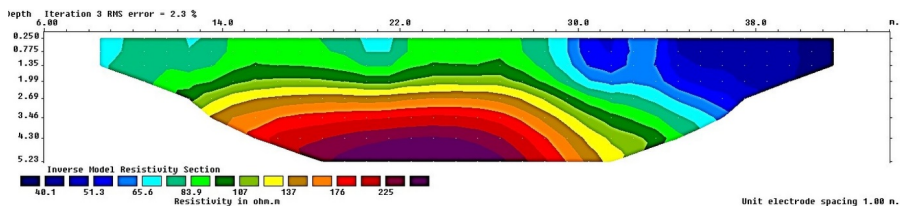


Fig. 12. – 2D inversion along the symmetrically shortened profile 6-6 with the “Refine model” option enabled.

The anomalous zone caused by the presence of the high resistivity pit filled with metallurgy byproducts (ash, slag etc.) was successfully singled out by 2D inversion of an unsymmetrically shortened profile 6-6 (first

electrode at the 0 meter, and the last at the 40th meter) (Figure 13). In this case, the "Refine model" option was also used, which highlighted near-surface inhomogeneities (ripple near surface effect). As a result, the 2D model of true resistivity values is shown in linear scale with two equidistance values (20 and 22.5 Ωm) to reduce the near-surface inhomogeneity effect and highlight the high resistivity anomaly induced by the survey objective (pit).

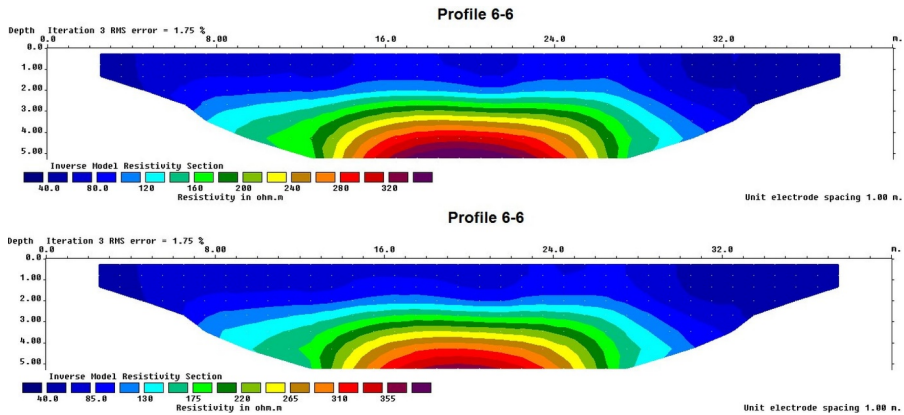


Fig. 13. – 2D inversion of the unsymmetrically shortened profile 6-6 with the “Refine model” option enabled.

Shortening the profile 6-6 (to a total length of 36 meters) resulted in a 2D profile containing only the high resistivity region and not the normal, undisturbed surrounding mediums (loess like clays that lie on top of gravel from the Toplica River). Figure 14a shows the measured apparent resistivity section, which only includes the high resistivity zone from figure 6 (the only exception being the near surface zones at the start and the end of the profile). The measured data clearly do not include data from the deeper lying gravel in the undisturbed zone; instead, they only include data from the anomalous pit zone. The results of 2D inversion were seen in both linear and logarithmic scales (Figure 14b, equidistance of 20 Ωm) (figure 14c, equidistance is increased by 14.5 percent). In both cases, equidistance is chosen to provide a visual representation of the model that does not obliterate any interesting zones. The logarithmic value scale of apparent resistivity best represents information about the near surface region. The high resistivity anomaly (ash pit) has lateral extent of about 10 meters, which corresponds to knowledge about the pit's geometry. The obvious resistivity method's vertical resolution makes it impossible to distinguish between the pit and gravel anomalies. Since the pit is in the lower part of the archeological profile, very close to the alluvium of the Toplica River (Figure 1b), no geophysical method enables the resolution to distinguish the

two anomalies. The other method used was ground penetrating radar (GPR), which was unable to distinguish between the two anomalies.

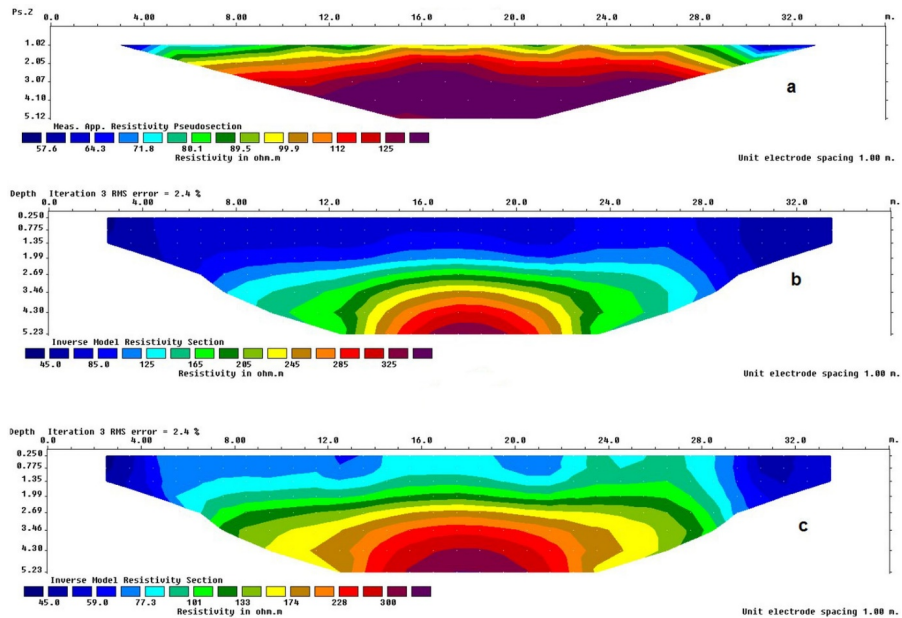


Fig. 14. – (a) Measured apparent resistivity pseudosection; (b) 2D model obtained after inversion with a linear value scale; (c) 2D model obtained after inversion with a logarithmic value scale.

GROUND PENETRATING RADAR SURVEY

Figure 15 shows maps of reflection amplitudes of electromagnetic waves at various two-way travel times (TWT= 15 ns, TWT= 18 ns, and TWT= 22 ns). Only when the TWT value is 18 ns can a contact between high and low resistivity zones be seen on the GPR map of reflection amplitudes as brighter shades. The poor quality of GPR amplitude map results can be interpreted in two ways. First, the presence of a thin low resistivity layer (second layer, thickness 0.8 meters, and resistivity of 35 Ωm) detected with 1D inversion of sounding curve S-3 could cause significant attenuation of the electromagnetic signal. In addition to this, the fact that the subsurface is made up of loess-like clay indicates the presence of fairly homogeneous medium with no permittivity discontinuities, which is a key physical property that determines the magnitude of amplitudes of a GPR section. The GPR method did not yield satisfactory results despite the existence of a relatively high resistivity pit.

3D FORWARD AND INVERSE MODELLING OF THE ANOMALOUS ZONE

The high resistivity zone (depicted as red zone) on Figure 15 (left) could be caused by three distinct anomalous bodies with a 3D geometry but creating a 2D-like contact on the apparent resistivity map. Figure 16 shows

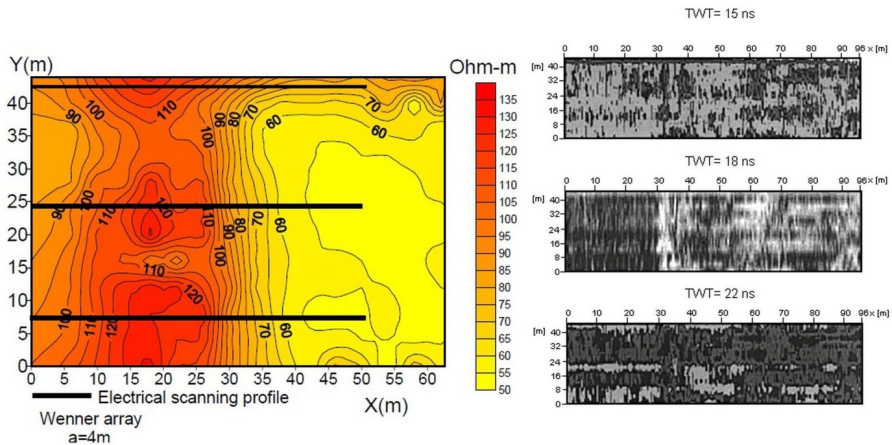


Fig. 15. – Apparent resistivity map (left) and GPR amplitude map obtained for different TWT (right).

a 3D model with a 1.1-meter-thick surface layer and a resistivity of 100 Ω m. The second layer is a low resistivity layer with a thickness of 0.5 meters and a resistivity of 30 Ω m, close to determined resistivity (about 35 Ω m) by the 1D inversion of sounding curve S-3. Three independent, high resistivity (300 Ω m) pits can be seen inside the low resistivity layer. The three pits are conical in shape and reach a depth of 3.2 meters. Finally, a high resistivity substratum (300 Ω m) was modelled at a depth of 4.4 meters, which corresponds to the gravel of the Toplica River alluvium.

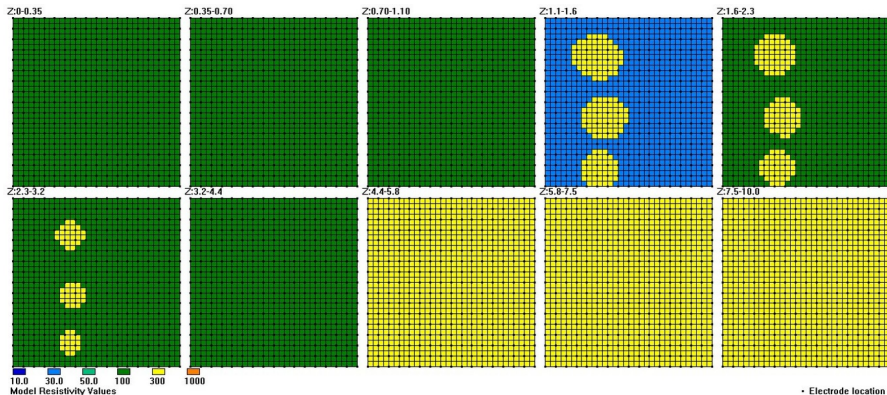


Fig. 16. – 3D forward model of high resistivity pits.

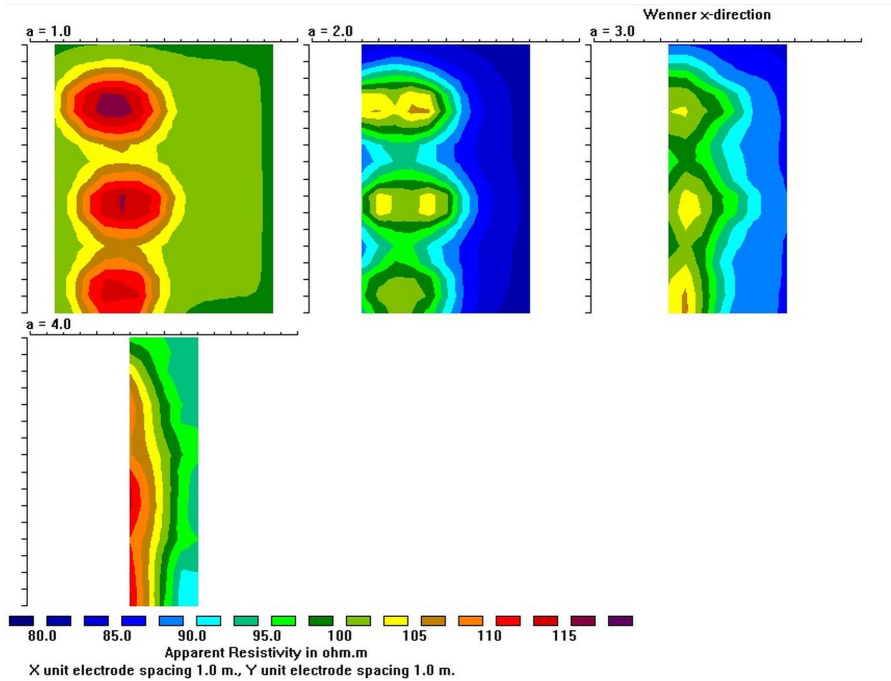


Fig. 17. – Theoretically calculated apparent resistivity maps for the Wenner array.

Figure 17 shows apparent resistivity maps for three array lengths ($a = 1\text{ m}$ to $a = 4\text{ m}$), or three depth levels (from approximately 0.5 meters to 2 meters), when the Wenner array is oriented along the X-axis. As the depth of investigation is increased, the 3D character of the individual anomalies is lost, and the three distinct anomalous zones become one 2D anomalous zone due to a lack of horizontal resolution. When the Wenner array's

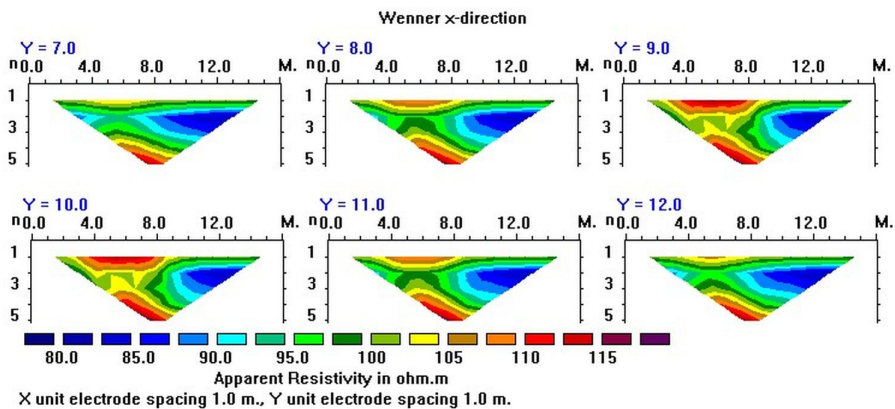


Fig. 18. – Apparent resistivity sections for the Wenner array oriented along the X-axis (Y-axis values range from 7 to 12 meters).

dimensions is $a = 4\text{m}$, the apparent resistivity map from Figure 15 corresponds best with the theoretically calculated apparent resistivity map.

Figure 18 shows apparent resistivity sections that are very similar to 2D survey sections, confirming that the vertical resolution is insufficient to distinguish individual pits from the high resistivity substratum.

The results of a 3D inversion of synthetic data confirm that resolution decreases with depth, and as a result, three individual anomalies are gradually lost as depth increases (Figure 19).

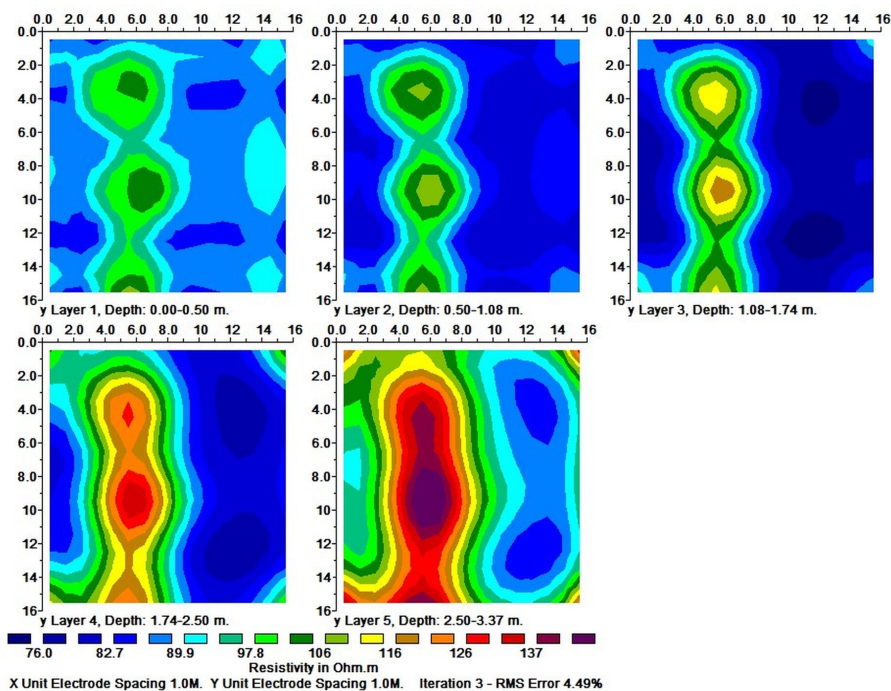


Fig. 19. – 3D model obtained after inversion of synthetic data.

Figures 20 and 21 show maps of 3D inversion using the Pole-pole array for the model shown in figure 16. A Pole-pole array is a two-electrode array in which one potential and one current electrode are separated from the other two electrodes by a large gap (usually ten times the separation of the two measuring electrodes). In comparison to the Wenner array, which is widely used for 2D data acquisition, Pole-pole and Dipole-dipole arrays are commonly used when conducting 3D geoelectrical surveys. The advantages of the Pole-pole array over the Wenner array can be found in better surface coverage which in turn gives a better subsurface image of the 3D anomaly in the high resistivity zone as can be seen from Figures 20 and 21.

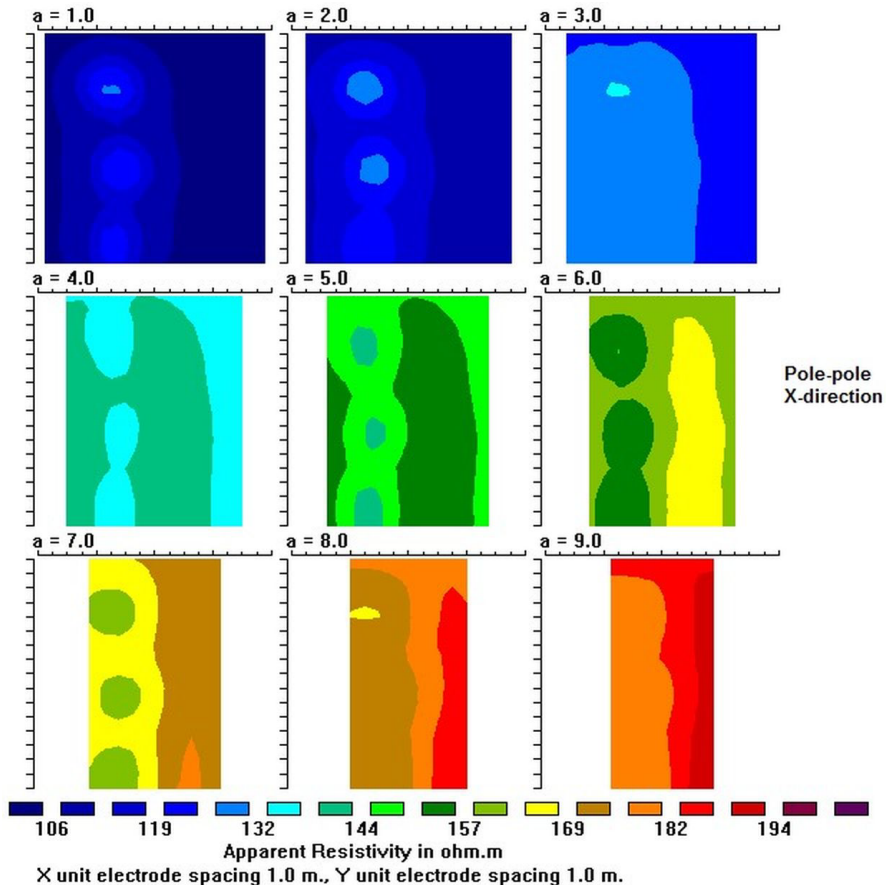


Fig. 20.—Theoretically calculated apparent resistivity maps for the Pole-pole array.

CONCLUSION

The thin, low resistivity layer (0.8 meters, about 35 Ωm) was discovered during 1D geoelectrical survey at sounding station S-3. Low resistivity layer could attenuate electromagnetic wave propagation, which was later determined to be one of the contributors to the low quality GPR results.

The contact between the high resistivity medium and the low resistivity medium was determined using geoelectrical mapping. Following that, 2D electrical imaging profile 1-1 was extended to obtain additional details about the subsurface architecture. The extension of profile 1-1 (1a-1a) was found to have simple, almost 1D subsurface geometry.

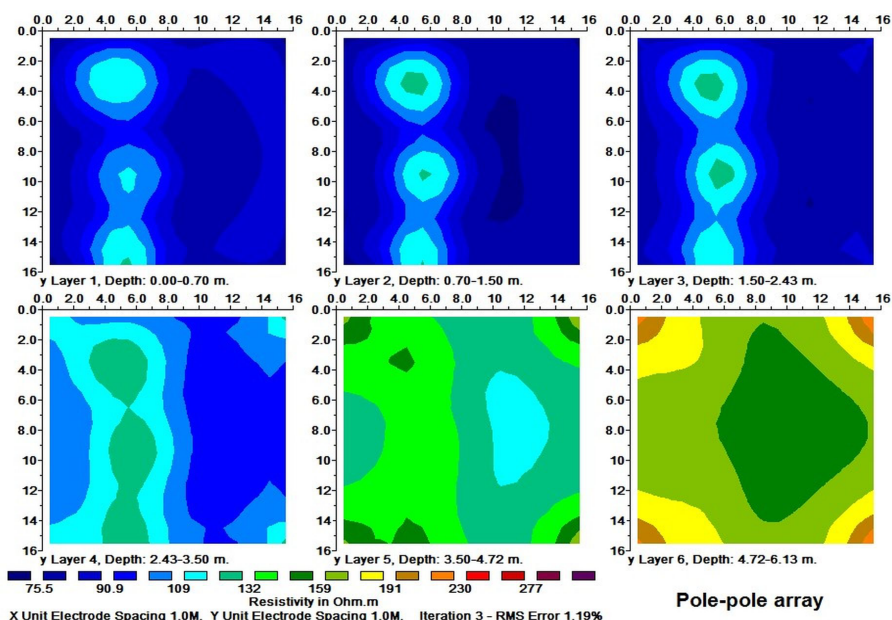


Fig. 21. – 3D model obtained after inversion of synthetic data for the Pole-pole array.

The benefits of model refinement and profile shortening (profile 6-6) were demonstrated with the goal of isolating the high resistivity anomaly. The use of this approach has the disadvantage of illustrating local near-surface inhomogeneities (ripple near-surface effect), which can be less visible by using linear scales and adjusting the equidistance scale. It was also determined that the ERT method's resolution, as well as that of other geophysical methods, is inadequate to distinguish the metallurgy-filled, high-resistance pit from the underlying high-resistance gravel (river Toplica alluvium).

The 3D forward model verified that as depth of investigation increases, three individual anomalies become 2D-like. 3D geoelectrical survey would provide the most accurate subsurface image, but would probably not yield a 3D model where the pits could be distinguished from the underlying high resistivity Toplica River gravel.

REFERENCES

- Holcombe, H. T., Jiracek, G. R. (1984): Three-dimensional terrain corrections in resistivity surveys. – Society of Exploration Geophysicist (SEG). – **Geophysics** 49 (4): 439–452.
- Loke, M. H. (2016a): Tutorial: 2D and 3D electrical imaging surveys. – Unpublished revised course notes. <http://www.geotomosoft.com/downloads.php>

- Loke, M. H. (2016b): Rapid 2D resistivity and I.P. forward modeling using the finite-difference and finite-element methods. – Unpublished software manual. <http://www.geotomosoft.com/downloads.php>
- Vukadinovich, S., Sretenovic, B., Slijivar, D. (1999): Combined Geoelectrical and GPR Investigations of the neolithic archeological settlement “Plocnik” - Serbia. In: 5th EEGS-ES Meeting, September 1999/ 1999 EEGS-ES Abstracts: 35–36. – European Association of Geoscientists & Engineers.

2Д ЕЛЕКТРОМЕТРИЈСКО СКЕНИРАЊЕ НА НЕОЛИТСКОМ НАЛАЗИШТУ “ПЛОЧНИК”

ФИЛИП АРНАУТ, БРАНИСЛАВ СРЕТЕНОВИЋ

Р Е З И М Е

Електрометријска метода скенирања (ЕРТ) коришћена је на локалитету неолитског налазишта „Плочник“ у околини Прокупља. Локација сонде вертикалног електричног сондирања (ВЕС) С-3 приказала је нискоотпорну зону дебљине 0,8 метара са вредности специфичне електричне отпорности од око 30 Ω . Оштра и стрма гранична површ која дели високоотпорну и нискоотпорну средину откривена је методом 2Д геоелектричног картирања. Резолуција електрометријског скенирања проверена је 3Д директним и инверзним моделовањем ради провере могућности диференцирања три индивидуалне високоотпорне јаме од околног високоотпорног шљунка.

Original scientific paper

2D GEOELECTRICAL RESISTIVITY TOMOGRAPHY APPLICATION AT THE FORMER CITY WASTE DUMP “ADA HUJA” – ECO - GEOLOGICAL PROBLEM

Branislav Sretenović¹, Filip Arnaut¹, Ivana Vasiljević¹, Vesna Cvetkov¹

Received: April 18, 2019

Accepted: June 3, 2019

Abstract: Rapid yearly technology development has enabled the possibility of renewing already acquisitioned data using new and improved software capabilities. The potential to gather 3D inversion results from 2D acquisitioned data is one of many new software options that were not available twenty years ago. Data obtained by ground penetrating radar (GPR) using 500 MHz and 300 MHz frequencies were used to confirm information attained by electrical resistivity tomography, hence the GPR survey was only conducted on a small portion of one profile line. Suitable location of the former communal waste site “Ada Huja” and the degree of possible environmental pollution are one of many reasons why it was crucial to reinterpret the 2D inversion data. Two decades after the initial scanning the data was used to solve the eco-geological problem which is a side effect of the decade long disposal and non-aiding of the communal waste dump.

Keywords: *2D inversion; geoelectrical method; Wenner array; ecology; Ada Huja*

1 INTRODUCTION

The location of the former city waste dump “Ada Huja” was labeled as *IBA (Important Bird Area)*, where refuge was found by some endangered species of birds like the black-headed seagull, big and small cormorant and the white-tailed eagle. Ada Huja, which is located only 4 km from the city center of Belgrade, was under the supervision from the former Federal executive council of SFRY and declared a communal waste dump in the year 1960. From then on it is representing an unsolved ecological problem. Such a solution for the daily production of waste causes several threats for the ecosystem such as the production of methane, hydrogen-sulfide, carbon-dioxide and carbon-monoxide.

Based on the research of the disposed waste origin and the knowledge of building material production during the twentieth century, there exists a founded assumption of environmental pollution with fiber (asbestos) silicates which increase the chance of

¹ University of Belgrade – Faculty of Mining and Geology

Emails: branislav.sretenovic@rgf.bg.ac.rs; G94-16@rgf.bg.ac.rs; ivana.vasiljevic@rgf.bg.ac.rs; vesna.cvetkov@rgf.bg.ac.rs

getting a variety of cancer types. Housing development “Deponija”, which is proclaimed to be the most unattractive housing settlement in Belgrade, is located near the former waste dump site “Ada Huja”. The lack of statistical reports on the health condition of the people who live in the near vicinity of the waste site indicates a neglecting attitude for the health of people, but also for the condition of the environment in general. This is best noticed in the completely wrong classification of forest classes in that area.

Lateral resistivity changes in the disposed waste were detected using 2D geoelectrical scanning method along three profile lines. The chosen Wenner array proved to be the most adequate because of the present urban noises, economical aspects of the survey and required staff. Geoelectrical methods for surveying subsurface structures serve as economically efficient and time-saving process (with regards to other subsurface surveying methods). The main limitation of exploration drilling is a high cost (drills, required staff etc.), and in this case (communal waste dump) the lack of utilization of geological logic in connecting the acquired data.

The classification of geophysical (geoelectrical) methods is based on the number of dimensions shown, thus, we have one-dimensional, two-dimensional and three-dimensional geophysical methods. One-dimensional methods (1D) are used only when we want to get the information of the thickness and depth of layers, without knowing what kind of lateral changes are present. This kind of survey method requires the least amount of data that has to be acquired and therefore is the cheapest. Two-dimensional (2D) geoelectrical surveys are giving us the information about the lateral changes of resistivity, because of that they demand more acquired field data (hundreds or a couple of hundreds) for the subsurface to be graphically presented. Generally speaking, 2D surveys give the best ratio of needed funds to be executed, required field acquisitioned data and gained information of the subsurface. That is why they are a favorable option of geoelectrical surveying methods (Dahlin, 1996). Three-dimensional (3D) surveys provide information about the whole geometry of the surveyed half-space. For this kind of study, the largest amount of acquired data is needed (couple of thousands) (Loke, 2004).

2D geoelectrical surveying was applied at the former communal waste dump “Ada Huja” (Figure 1). This type of method was used given that immense lateral resistivity changes were expected in the hoarded material and accounting for the time and funding aspect, this kind of method was a logical choice.

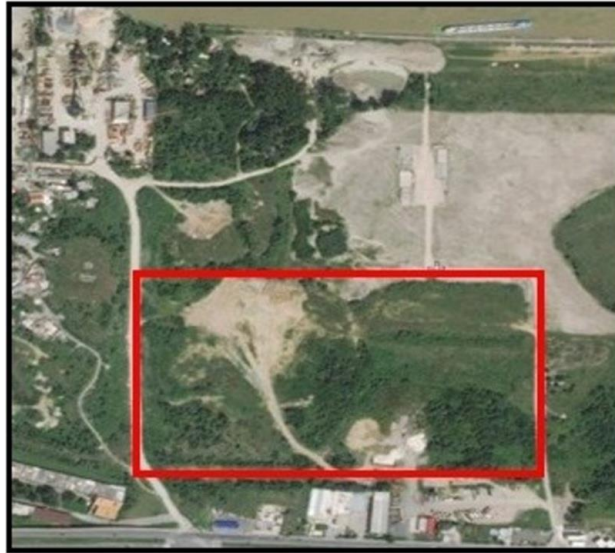


Figure 1 Satellite image of the surveyed field (Google Earth)

Accounting for the occurrence of the subsurface material (anthropogenic) the stratigraphic column is of little use, except to provide the information about the impermeable marl base. Geotechnical profile (Figure 2), constructed using data from 22 exploration drill holes is an improved source of information about the subsurface. It is noticeable that there exists a surface embankment layer (sand, gravel, construction material etc.), which generally enhances the already bad load capacity and subsidence characteristics of the terrain. The embankment mostly lies on the formerly disposed waste, with a smaller part laying on a layer of sand. The base which meanders from 7.5 m to 12 m is composed of marl type rocks. That type of rock is porous but impermeable, in a hydrogeological sense (Šperl and Tročkova, 2008). Water level varies in the range of about 3 m, passing through the deposited waste of the surveyed subsurface forming an ecologic problem that will be discussed in the latter chapters.

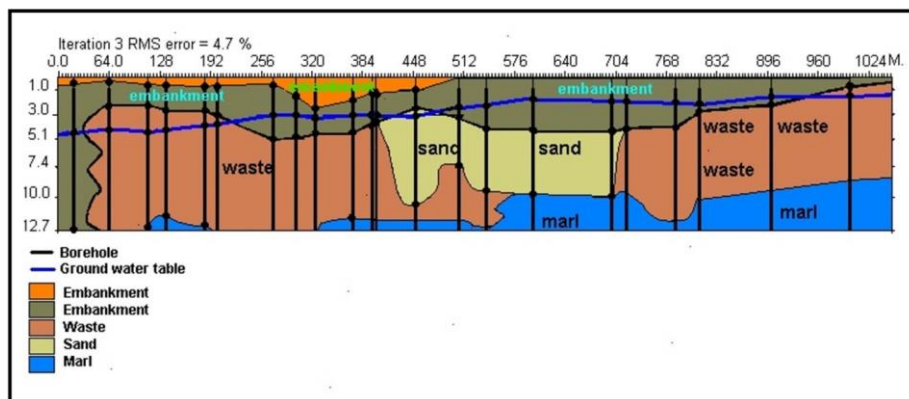


Figure 2 Geotechnical profile of the surveyed terrain (Sretenović, 1998)

2 RESULTS OF THE 2D GEOELECTRICAL SCANNING METHOD

The survey was carried out with three parallel profile lines, east-west heading (Figure 3a), “profile line I” is the longest and consists out of six parts, the other two profile lines (II and III) consist out of five parts. The construction of resistivity models and sections was enabled using the software package Res2DINV from the developer Geotomosoft (Loke, 1998). Further surveying was executed with another three parallel profile lines heading north-south (Figure 3b) with a total length of 279 m.

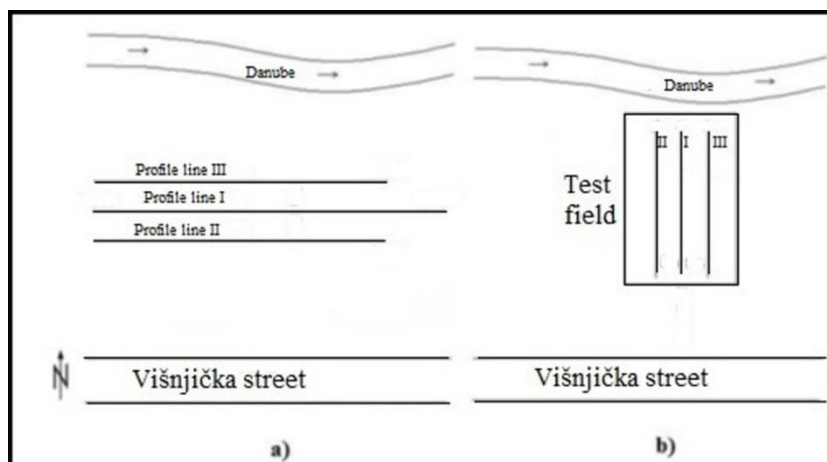


Figure 3 Location of profile lines (Sretenović, 1998)

Primary research was conducted using 32, 51 and 52 electrode arrays according to the accessible length of the scanned unit. Electrode spacing was 4 m, which in the high contrast resistivity subsurface gave out a horizontal resolution of 2 m for the first depth level. The horizontal resolution further decreases with the increasing of the depth level. Mutual profile line distance was 5 m.

Farther scanning on the “test field” was carried out with a goal of getting a better understanding of the geotechnical characteristics. The north-south bearing additional profile lines were shorter in length, with a distance between the electrodes set at 3 m. This was done for the improvement of the horizontal resolution. Like the first scanning that was carried out, the additional one used the Wenner Alpha array with 32 electrodes.

All profile lines were scanned on six depth levels, which amounted to a depth reach of around 12 m. The value of the depth of investigation did not enable the detection of the marl base all along the profile line. This was caused by the low resistivity value zones (1-2 Ωm) which made a “screen” type effect and made it difficult for the penetration of electric current to deeper levels of the half-space, i.e. regardless of Wenner array dimensions.

Due to the large resistivity value variations (1-200 Ωm), the establishment of one unique resistivity scale and observing the profile line as a whole would produce the loss of model features. Because of the overwhelming heterogeneity of the subsurface it was strived to maintain the maximal amount of details of 2D models, and because of that units of the same profile line can have different resistivity scales.

Part three of the profile line I has a length of 124 m (Figure 4). A high value resistivity zone with a length of 56 m (252 m to 308 m of the profile line) can be noticed and it represents most likely sand, gravel and other construction materials. Decomposed waste has a lens-like geometry and retains a low resistivity value. It obtains that low value from the reaction with electrolytic leachate. Marl type rocks that make up the basis of the half-space also have a low resistivity value, but one order of magnitude greater than the deposited waste.

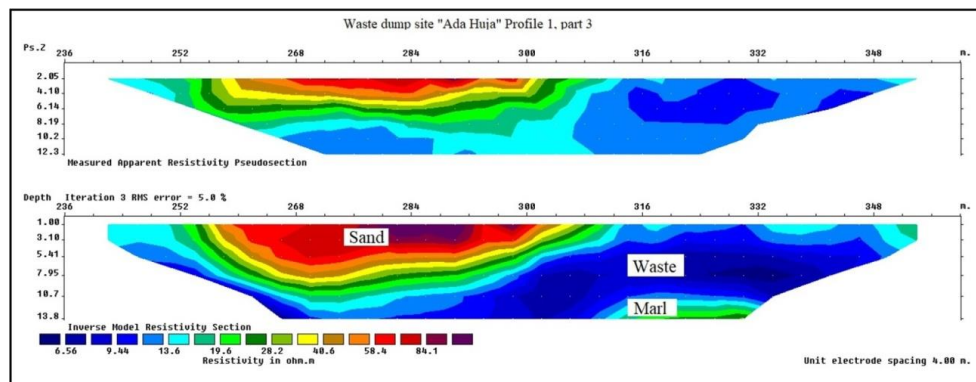


Figure 4 Sections of apparent (top) and true (bottom) resistivity values of profile line I, part 3 (modified after Sretenović, 1998)

Information acquired from the fourth part of profile I (Figure 5) can be compared with the geotechnical profile (borehole data). Four boreholes confirmed the presence of decomposed waste material as well as a sand lens and surface embankment layer consisting of previous construction material. Like the previous section (Figure 4), the decomposing waste is represented as a distinctive laterally changing low resistivity layer.

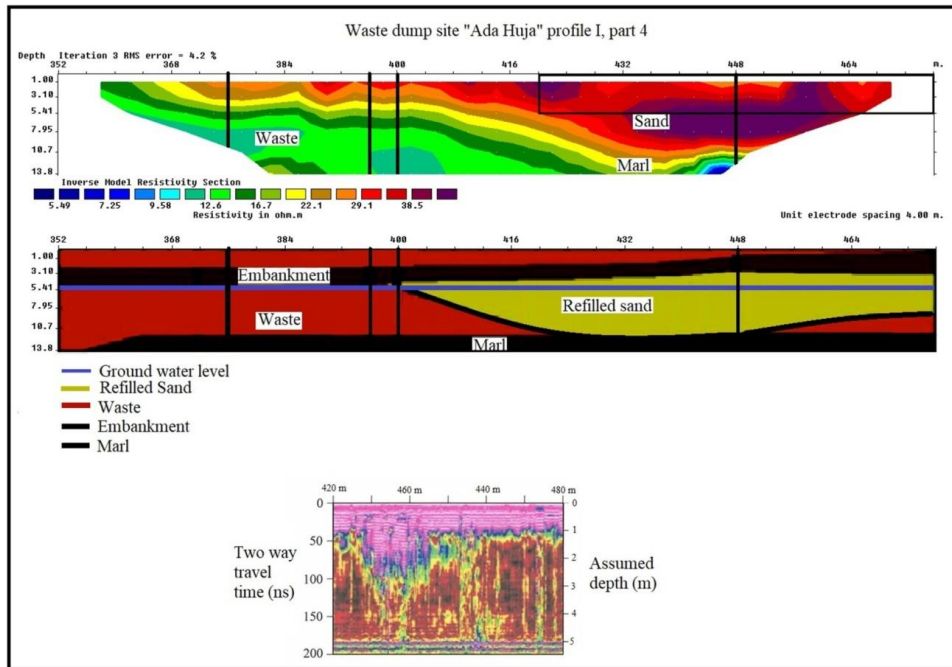


Figure 5 Profile I part 4, model-section of true resistivity values (top), geotechnical profile with borehole data (middle), ground penetrating radar (GPR) cross section (bottom) (modified after Sretenović, 1998)

Furthermore, the survey was conducted using ground penetrating radar (GPR). The depth of investigation achieved with GPR is generally very limited along the whole examined half-space because of the electromagnetic signal attenuation due to low value resistivity layers. The GPR section (Figure 5, bottom) corresponds to the marked rectangle on the 2D model-section of true resistivity values (Figure 5, top) where the presence of relatively high resistivity sands had to enable the least amount of attenuation to electromagnetic signals. At the depth that corresponds to the two-way travel time of 200 ns, the horizontal reflection is noticed that could indicate a horizontal interface, that was not detected with the geoelectrical scanning method. Acquired borehole data indicated a presence of subsurface water at a depth of around 5 m (Figure 5, middle) which makes a good reflective surface that divides two diametrically opposite resistivity values and a dielectric constant value which are in contact. Taking into account that the GPR survey was conducted in the urban environment where the presence of noise is at a higher level, the horizontal interface could be coherent noise that was picked up by the receiving antenna which looks like a horizontally continuing reflection that is hard to distinguish from the primary interface reflections. Within the first 1 to 2 m of the GPR cross section thin layers (laminae) can be seen that indicate a high vertical resolution of the GPR equipment.

Display of the sand lens distribution (Figure 6) as well as the marl basis can greatly help with getting a better understanding of the subsurface geometry. The sand lens is active around the mark of 360 m and it gains maximal thickness along the middle survey line. The last profile line (Figure 6, bottom) shows again a decrease of thickness of the sand lens. The profile lines are at a mutual distance of 5 m.

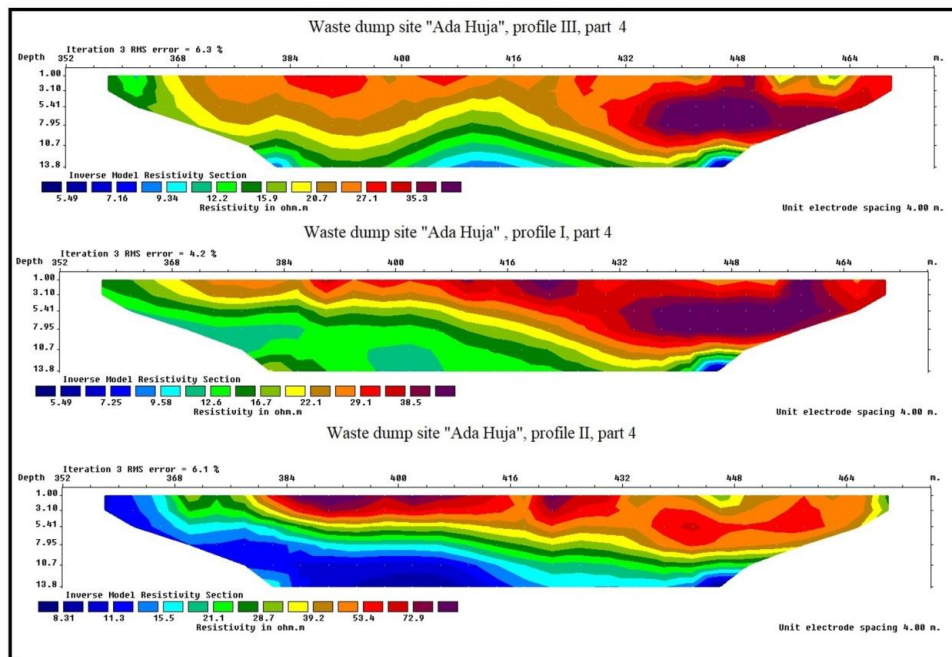


Figure 6 Display of the sand lens distribution through three parallel profile lines (I, II, III) part 4 (modified after Sretenović, 1998)

3D inversion results (Figure 7) obtained on six depth levels can be used to further understand the overall geometry of the examined half-space. Layer one, two and three are high resistivity layers due to the disposal of sand, gravel and construction material. This type of resistivity trend can be followed to the depth of 6.95 m. Beneath the first three layers, a lower value of resistivity can be observed, due to the leachate waters reacting with decomposing waste. The deepest layer (layer six) is located on the contact with marl that has its upper layers alternated by means of leachate waters, hence, it has a low resistivity value as well.

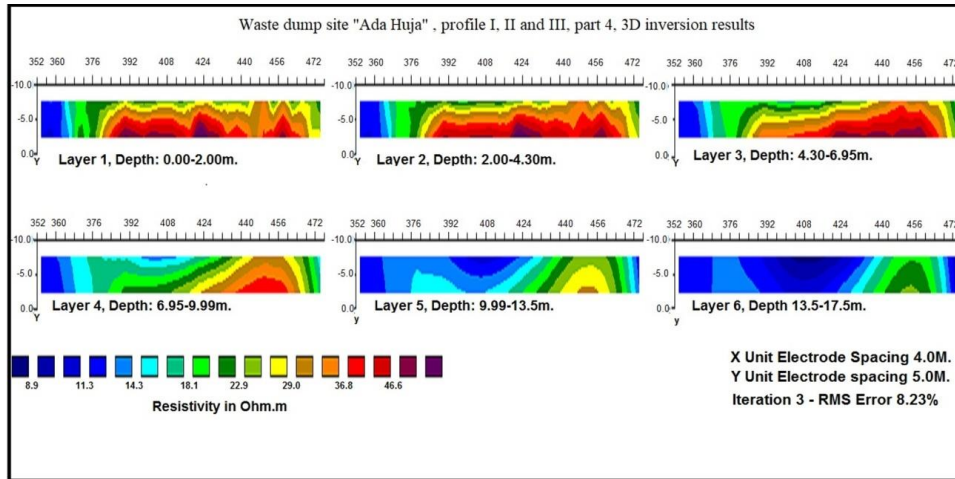


Figure 7 3D inversion results of profile I, II and III, part 4

Profile I is “compressed” and compared with the geotechnical profile (Figure 8). This procedure was done with partial overlap of the previously scanned parts. Detail loss is greatly expressed in comparison with individually scanned parts, but with this procedure the whole spectrum of lateral resistivity changes can be observed. Correlation with the geotechnical profile is to a great extent good even though the geotechnical profile is a simplified version of the profile obtained by 2D inversion process. Near the end of the cross section some very low resistivity values can be noticed ($1.2 \Omega\text{m}$) which further disrupt the resistivity scale and the level of details.

In addition to the good correlation between three mutually parallel profile lines (similar geometry and resistivity distribution), for the sake of confirmation and justification of the 2D geoelectrical scanning method, the root mean square value (RMS) should be viewed. Considering the instrumental error, ambient (urban) noise and the pronounced lateral resistivity changes of the subsurface, the RMS% values could be considered satisfactory. The enhancement of the RMS% value in some parts of the profile could be attributed to the generally 3D character of the examined half-space.

It is necessary to consider the possibility of contamination of the lower layers (marl rocks) by means of electrolytic fluids. These marl type rocks are located near the maximum depth of investigation under the low resistivity layered waste and for that reason the contact between marl and waste can produce a “screen” type effect. To check the existence of the “screen” effect profile I was examined from 652-856 m. Near the mark of 784 m (Figure 8), the black rectangle shows a very low resistivity value zone ($1.2 \Omega\text{m}$). Comparing it to the geotechnical profile the waste is laying on top of the marl basis which starts to be detected at the depth of around 8-8.5 m. Using the 2D geoelectrical scanning method, that zone was not detected even though marl has an order of magnitude higher resistivity value.

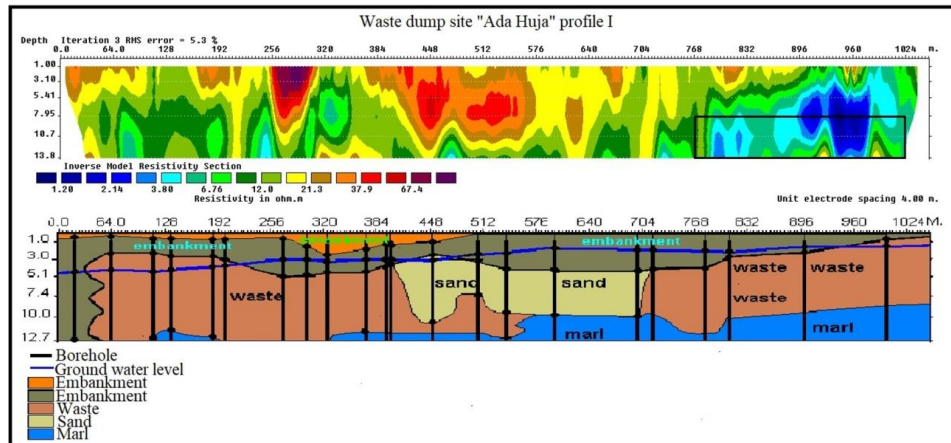


Figure 8 Entire length of profile line I compared with the geotechnical profile (modified after Sretenović, 1998.)

Forward modeling using the program Res2DMOD (Geotomosoft) was applied and a 2D model with twenty depth levels was obtained. Using that model and putting it through the process of inversion resulted in theoretical models calculated to the depth of forty meters (Figure 9).

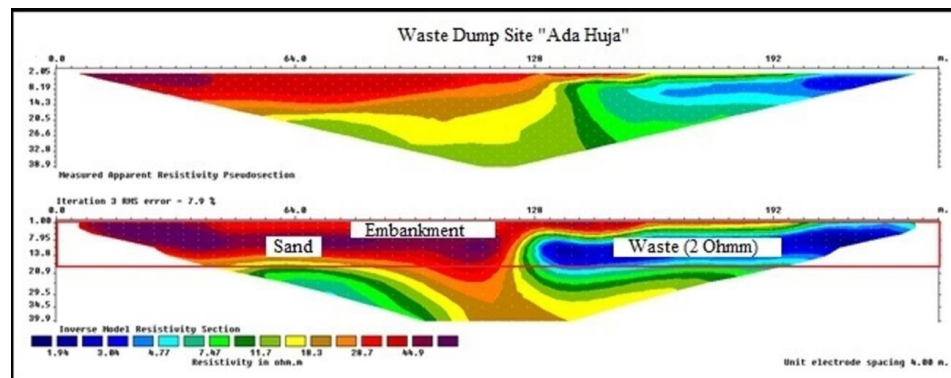


Figure 9 Theoretically calculated model for obtaining the information about the lower marl base

With a resistivity value of 2 Ωm (waste) the marl base can be detected i.e. the basic principle was held (resistivity values increase with depth). The red rectangle indicates a depth of investigation larger than the one originally used (six depth levels), but even when using a ten-depth layer survey, the marl basis can not be detected all along the profile line. It is necessary to perform a survey with more than ten depth levels, because the low resistivity electrolytes do have a “screen” like effect and deflect the electrical current from its natural flow. Furthermore, it is vital to conduct an exploratory borehole at the contact of the waste and marl to get a better understanding of the contaminated uppermost marl layers.

3 ENVIRONMENTAL PROBLEM OF THE FORMER WASTE DUMP SITE “ADA HUJA”

3.1 Basic Information

Convenient location of the former waste site “Ada Huja” (4 km from the Belgrade city center, first city zone) resulted in the housing development for 900 to 2000 people built on the most unsanitary locality in Belgrade (housing development “Deponija”). In the beginning of the year 1960, under the jurisdiction of the former Federal executive council of SFRY, “Ada Huja” became a waste dump site. A variety of birds was endangered as consequence of this decision, one of which was the white-tailed-eagle, the largest eagle in Europe. Through the decades that followed, rapid agglomeration caused the habitation of the surroundings, but the site between the Pančevo bridge and the Danube riverside remained unregulated.

Palilula municipality is currently the largest in area and the second largest by the number of inhabitants (Republic statistical institute, 2011). A couple of thousands of people pass through Višnjički Boulevard on a daily basis not knowing that there is a serious possibility of environmental contamination on the edge of the boulevard.

Plot areas should be taken with a reserve considering that they go out of the examined terrain, hence nobody can know for sure if the contaminated trend is maintained. Figure 10b shows the real area of the surveyed terrain and Figure 10a shows the cadastre plots of the terrain.



Figure 10 Maps: a) Location of cadastre plots (National Spatial Data Infrastructure, 2019) and b) the real area of the surveyed terrain (Google Earth)

Table contents (Table 1) should give adequate information about the land class. Parcels 7/36, 7/38 and 7/40 are labeled as “Class I Forest” (ash and oak tree) and they occupy around 19280 m² (192.8 are). Class I forest represents the highest quality of wood (Official Gazette of the Republic of Serbia, 1994).

Table 1 Basic information about the cadaster parcels on the investigated terrain
(Republic Geodetic Authority, 2019)

Parcel Number	Owner/User	Way of Land Use	Type of Land	Status of the Construction (if it exists)	Area [m ²]
7/1	Belgrade Pier AS	Land under Construction Buildings/Class I Forest	Cities Urban Construction Land	Unknown Purpose, Constructed Without a Permit	77897
7/2	City of Belgrade	Canal	Cities Urban Construction Land	/	2716
7/3	City of Belgrade	Land under Construction Buildings and Other Objects	Cities Urban Construction Land	Unknown Purpose, Constructed Without a Permit	460
7/7	City of Belgrade	Deep Pond	Cities Urban Construction Land	/	11
7/32	Belgrade Pier AS	Land under Construction Buildings and Other Objects	Cities Urban Construction Land	Unknown Purpose, Constructed Without a Permit	1995
7/35	City of Belgrade	Land under Construction Buildings and Other Objects	Cities Urban Construction Land	Unknown Purpose, Constructed Without a Permit	1531
7/36	City of Belgrade	Class I Forest	Cities Urban Construction Land	/	9377
7/37	City of Belgrade	Canal	Cities Urban Construction Land	/	35
7/38	City of Belgrade	Class I Forest	Cities Urban Construction Land	/	236
7/39	City of Belgrade	Deep Pond	Cities Urban Construction Land	/	57864
7/40	City of Belgrade	Class I Forest	Cities Urban Construction Land	/	9667
7/42	City of Belgrade	Artificially Made Infertile Land	Cities Urban Construction Land	/	359

Visiting the examined terrain (Figure 11), it was quickly obvious that there is no indication of a Class I Forest.



Figure 11 Examined terrain photographed on 29th March 2019.

3.2 Possible Environmental Threats

The list of the possible ecosystem threats regarding the locality of the former communal waste dump site “Ada Huja” can be lengthy, but without thorough examination it remains speculative i.e. theoretical. Investigation based on theoretical assumptions can be quite appropriate, if conducted correctly and relying on already accepted previous research.

Broadly speaking, the main waste dump site contaminants are asbestos (if construction material that was assembled during the twentieth century was deposited), hydrogen-sulfide (H_2S), carbon-monoxide (CO), carbon-dioxide (CO_2) and the side effects of decomposition of polyethylene-terephthalate (PET).

3.2.1 Asbestos

Because of its mechanical properties, asbestos was used in civil engineering from the beginning of the twentieth century. During the 1980’s health consequences were established (a vast variety of carcinogens and respiratory diseases) that were directly

linked to air contamination by asbestos fibers. From that point on, any handling is heavily regulated (from 2008 banned for civil use by the Republic of Serbia). Guided by the rule book “Handling of asbestos contaminated waste” published in the journal “Official Gazette of the Republic of Serbia” number 75/2010 main concerns will be expressed.

The term *asbestos* does not relate to one substance, but to a whole group of fiber silicates that will be cited:

- Crocidolite,
- Actinolite,
- Anthophyllite,
- Chrysotile
- Amosite,
- Tremolite.

It is important to note that the Republic of Serbia does not have a waste dump site for the disposal of environmental threatening and industrial waste where asbestos, mineral wool, waste with adsorption properties and ionic masses belong (Environmental protection secretariat, 2008). Because of this kind of assertion, it is not unlikely that during the lifespan of the waste dump site “Ada Huja” construction material containing asbestos was deposited, and was not handled according to Article 4 (Official Gazette of the Republic of Serbia, 1994) i.e. did not succumb to regular air emission measurements and fluid leakage.

Results of geoelectrical measurements and borehole data did show a groundwater level existed, that water reacts with the decomposing waste and creates an electrolyte with a very low resistivity value. The fibrous minerals do not react with water nor with acid solutions except chrysotile which in a reaction with acidic solutions can release toxic matter to the ecosystem.

3.2.2 Side effects of continuous disposal of polyethylene-terephthalate (PET containers)

Polyethylene-terephthalate (PET) is half-crystalline, thermoplastic polyester which is one of the most used varieties of polyester due to its superior mechanical and chemical properties (Supawee and Wanlop, 2014). Primary use of PET containers in everyday life is for water and soft drink packaging, as well as electronics, home appliances, textiles etc.

According to data of the “*General characterization of waste*” (Singh et al., 2015) which was not conducted for the Belgrade locality but can be used with some reserves, organic waste boasts a maximum of household waste with 43%, but that waste is biodegradable so it is not accounted for in the “Ada Huja” case. This is also true for paper and paper-

like products. Plastic contributes 21% with the highest decay rate and is the primary ecosystem contaminator (Figure 12).

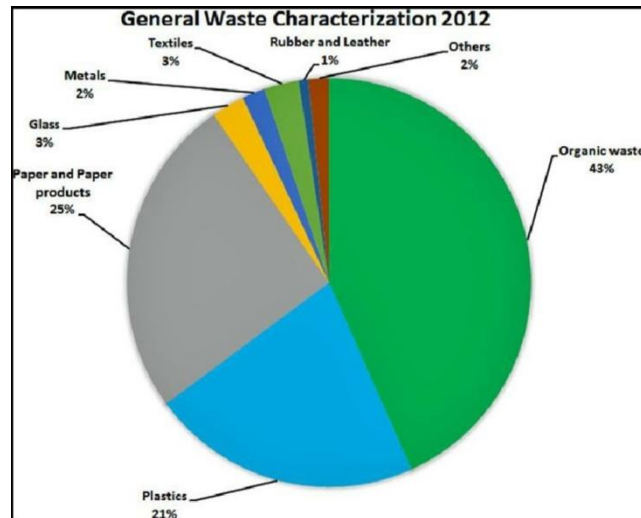


Figure 12 General waste characterization (Singh et al., 2015)

There are five ways in which plastic degrades in the environment (Andrady, 2011):

- Biodegradation,
- Photo-degradation,
- Thermo oxidative degradation,
- Thermal degradation (not and environmental degradation process),
- Hydrolysis.

Natural plastic degradation starts with the process of photodegradation. UV rays emit energy which is used to incorporate oxygen atoms into the polymer, and because of this, plastic becomes susceptible to strain (easier to break apart). This process is repeated until plastic gets a low enough molecular mass to be consumed by microorganisms who further whiter away the plastic and convert it to carbon-dioxide. The whole process can last about fifty years (Webb et al., 2013).

Even though the whole process lasts for about half a century, in the case of “Ada Huja” that number is not an issue because the waste is deposited for 59 years (in the year 2019). The mechanism is accelerated with the presence of underground water that was detected.

There are secondary problems that arise with the disposal of PET containers in landfills, such as production of benzene, toluene and ethyl-benzene in gas or leachate form. There is evidence that PET disposal can lead to the upsurge of hydrogen-sulfide concentrations (potentially deadly gas in higher concentration values) by sulfide-reducing bacteria (Webb et al., 2013).

3.2.3 Methane

Methane is a carbonic compound formed when one hydrogen atom is linked to four oxygen atoms with single bounds. Methane is an odorless, colorless gas which is not toxic but is highly flammable and it is a side effect of every landfill especially the ones that are not regulated by the state (safely secured).

Not sanitizing the methane released from a landfill will inevitably lead to its release into the atmosphere. Methane has a 23 times higher global warming index than the same volume of carbon-dioxide (Themelis and Ulloa, 2006)

Not long after the waste is disposed in a landfill, organic compounds initiate biochemical reactions in an air stimulated environment. The product of this reaction is carbon-dioxide and water vapor. First phase is bacteria fermentation, which in a process of hydrolysis dissolve complex organic compounds into smaller molecules. Second phase is the transformation of those molecules into organic acids, carbon-dioxide and hydrogen. The last phase is methane production by methanogenic bacteria which either dissolve those acids into methane and carbon-dioxide or reduce carbon-dioxide with hydrogen (Themelis and Ulloa, 2006).

Looking at the research done by the Environmental and Energy Study Institute (Pierson, 2013), and neglecting the cost of installation of such a plant, the calculated values were that one million tons on a day to day basis can produce 12233 m³ of landfill methane, that is equivalent to 0.78 MW of electrical energy.

Taking into account the previously exhibited data, and looking at the 2017 methane stimulated fire at the landfill "Vinča", which posed a threat to a million and a half inhabitants of Belgrade, such a mini-plant would not be an enormous investment. Daily production of waste for the city of Belgrade reaches 1700 tones (Mitrović, 2014).

3.2.4 Other hazardous compounds

Presence of other compounds which take less than 1% in the produced landfill substances cannot be ignored. Those compounds are mostly organic-sulfides, oxidized and other organic compounds. The biggest accompanying problem of these gases is the smell threshold, a value that can be quantified. The best representative compound is hydrogen-sulfide (H₂S) with its well-known smell (like rotten eggs) which has less than 1% intake in the total discharged gas. Measuring the smell threshold in a landfill hydrogen-sulfide ranks in the top three causes of repulsive odor just behind methyl-sulfide and ethyl-acetate (Yue et al., 2012).

Gases that contribute to global warming cannot be forgotten, they predominantly consist of carbon-dioxide and carbon-monoxide. Alongside methane which makes up 45-60% of the total emitted gas from a landfill, carbon-dioxide takes the rest 40-60% leaving just around 1% to other gases (Marković and Stevović, 2016). The presence of water vapor

can speed up the process of gas production. This is of great value for the examined terrain “Ada Huja” because underground water level was detected.

4 CORRECT ECOLOGICAL AND GEOLOGICAL CONDITIONS FOR A LANDFILL

From a geological point of view, the most important aspect of a controlled and safe landfill is the isolation of the disposed waste. This implies scouting the terrain for favorable conditions (not in contact with a water spring or any underground water, appropriate rock type basis, further enough from housing estates but still a good road connection etc.) and after this the engineering part starts (correct isolation of the hole, pipelines for oxygen pumping and methane recovery, mini-station for leachate treatment etc.)

Geoelectrical and geotechnical investigations that were carried out on the terrain established a marl foundation. Pure marl is a carbonate-clay rock that consists out of 75-40% calcium-carbonate (CaCO_3) and around 43-22.4% calcium-oxide (CaO) (Đorđević et al., 1991). Clay type rocks are most suitable for a good foundation for a controlled landfill. Additionally, boreholes have to be present to monitor the subsurface water level even though the waste is “incubated” from the surrounding rocks (Figure 13).

This kind of a landfill not only uses the waste to produce electric energy, but it provides several job openings, and reduces the risk of a methane stimulated fire. Monetary investments for this kind of a project can be steep, but it is of most importance to make a good economical calculation for taking a part in this kind of project. United States of America is a country that has the most initiative for methane exploitation with 519 plants (Alexander, 2010).

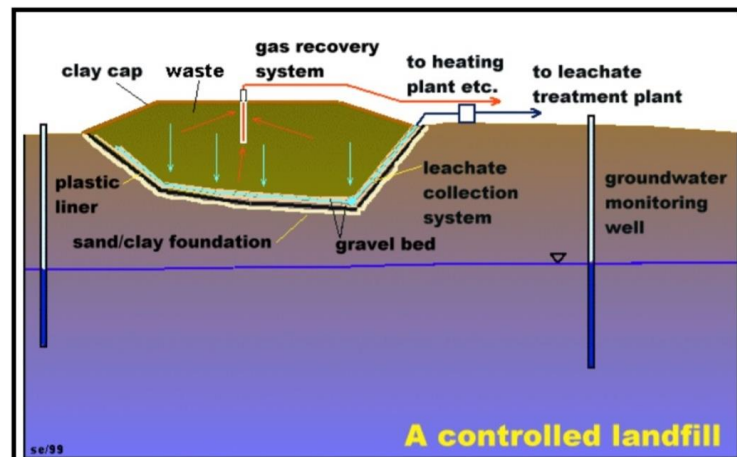


Figure 13 Controlled landfill with a methane induced mini-plant (Earle, 2005)

5 CONCLUSION

Geoelectrical surveys for obtaining a better understanding about the subsurface geometry have shown its full potential in the form of needed funds, reliability and acquisition speed. Further surveys showed a good correlation between the borehole data and 2D inversion process.

The location of the surveyed area is a very interesting terrain in a form of civil engineering opportunities, because of its proximity to the Belgrade city center. It is of great importance that any sort of civil engineering project that takes place on the examined terrain should be heavily regulated. The release of subsurface gasses into the atmosphere will yield side effects on the environment.

To get a completely detailed picture of the ecological situation of “Ada Huja” further research should be undertaken. This research should include additional geoelectrical measurements with more depth levels, pH value measurements (to reveal whether the electrolyte is a strong acid or a strong base), measurements to determine how much subsurface gas is present (methane, hydrogen-sulfide, carbon-monoxide, carbon-dioxide etc.), and hydrogeological measurements for the quality of the underground water.

REFERENCES

- ALEXANDER, A. (2010) An Overview of Landfill Gas Energy, Landfill Methane Outreach Program (LMOP).
- ANDRADY, A. (2011) Microplastics in the marine environment. *Marine Pollution Bulletin*, 62, pp. 1596-1605.
- DAHLIN, T. (1996) 2D Resistivity surveying for environmental and engineering applications. *First Break*, 14, pp. 275-284.
- EARLE, S. (2005) *Geology 312*. Malaspina University College.
- ĐORĐEVIĆ, V. and ĐORĐEVIĆ P. (1991) *Petrology*. Belgrade: Nauka.
- ENVIRONMENTAL PROTECTION SECRETARIAT (2008) Report on life quality in the Belgrade area in 2008. Belgrade, Serbia.
- GOOGLE EARTH (2019) Google Earth. [Online] Available from: <https://earth.google.com/web> [Accessed 13/04/19]
- LOKE, M.H. (2004) Tutorial: 2-D and 3-D electrical imaging surveys. Geotomo Software, Res2dinv 3.5 Software.
- MARKOVIĆ, J. and STEVOVIĆ, S. (2016) The process of creation and analysis of the landfill gas from the landfill in the region of Pčinja. *Journal on Processing and Energy in Agriculture*, 20, pp. 63-68.

- MITROVIĆ, S. (2014) Applied Ecology- guidebook, chapter 3 (Green potential of the city of Belgrade). Ministry of agriculture and environmental protection, pp. 208-235.
- NATIONAL SPATIAL DATA INFRASTRUCTURE (2019) Geosrbija. [Online] Available from: <https://geosrbija.rs/en/> [Accessed 13/04/19]
- OFFICIAL GAZETTE OF THE REPUBLIC OF SERBIA (1994) Rulebook on the cadastre land classing, 37/94 and 49/94. Belgrade, Serbia.
- OFFICIAL GAZETTE OF THE REPUBLIC OF SERBIA (2010) Rulebook on handling asbestos contaminated waste, 75/2010. Belgrade, Serbia.
- PIERSON, R. (2013) *Landfill Methane, Fact Sheet*. Washington: Environmental and Energy Study Institute.
- REPUBLICAN STATISTICAL INSTITUTE (2011) Republic of Serbia Census report 2011. Belgrade, Serbia.
- REPUBLIC GEODETIC AUTHORITY (2019) Real-estate cadaster (eCadaster). [Online] RGZ, Available from: <http://www.rgz.gov.rs/> [Accessed 13/04/19]
- SINGH, R.K. et al. (2015) Analyzing waste problems in developing countries: Lessons for Kathmandu, Nepal through analysis of the waste system in Tsukuba City Japan. *Journal of Scientific Research and Reports*, 8 (6), pp. 1-13.
- SRETENOVIĆ, B. (1998) Report on the geoelectrical survey conducted at the former waste dump site ar Karaburma. Faculty of Mining and Geology, Geotechnical Institute, Belgrade.
- SUPAWEE, M. and WANLOP, H. (2014) Rheological and mechanical properties of recycled PET modified by reactive extrusion. *Energy Procedia*, 56, pp. 547-554.
- ŠPERL, J. and TROČKOVA, J. (2008) Permeability and porosity of rocks and their relationship based on laboratory testing. *Acta Geodynamica et Geomaterialia*, 5 (1), pp. 41-47.
- THEMELIS, J. N. and PRISCILLA, A. U. (2006) Methane generation in landfills. *Renewable Energy*, 32, pp. 1243-1257.
- WEBB, H.K. et al. (2013) Plastic degradation and its environmental implications with special reference to Poly(ethylene terephthalate). *Polymers*, 5, pp. 1-18.
- YUE, D. et al. (2013) Composition and Distribution of Non-Methane Organic Compounds at Municipal Solid Waste Landfill Surfaces. Semantic Scholar. [Online] Available from: pdfs.semanticscholar.org/82a1/fb621d96c57c62874fc693b9b78386268d7e.pdf [Accessed 13/04/19]

ЗАПИСНИЦИ СРПСКОГ ГЕОЛОШКОГ ДРУШТВА (ЗА 2023. ГОДИНУ)

COMPTEs RENDUS DES SÉANCES DE LA SOCIÉTÉ SERBE DE GÉOLOGIE
pour les années 2023

REPORTS OF THE SERBIAN GEOLOGICAL SOCIETY
for the year 2023

Beograd, 2024

Владимир Симић¹, Стефан Петровић¹, Филип Арнаут², Весна Цветков¹,
Милена Костовић¹, Драган Радуловић³, Јовица Стојановић³, Владимир
Јовановић³, Дејан Тодоровић³, Нина Николић⁴, Јелена Сенћански⁵,
Грозданка Богдановић⁶, Драгана Мариловић⁶

ПРОЈЕКАТ ПРИЗМА: КАРАКТЕРИЗАЦИЈА И ТЕХНОЛОШКИ ПОСТУПЦИ ЗА РЕЦИКЛАЖУ И ПОНОВНУ УПОТРЕБУ ФЛОТАЦИЈСКЕ ЈАЛОВИНЕ РУДНИКА „РУДНИК“

PRIZMA PROJECT: CHARACTERISATION AND TECHNOLOGICAL PROCEDURES FOR RECYCLING AND REUSING OF THE RUDNIK MINE FLOTATION TAILINGS (REASONING)

ПРИКАЗ ПРОЈЕКАТА – PROJECT PRESENTATION

Апстракт: Пројекат Карактеризација и технолошки поступци за рециклажу и поновну употребу флотацијске јаловине из рудника „Рудник“ (REASONING) је мултидисциплинарни истраживачки пројекат испитивања јаловине кроз геолошка, минералозна, геохемијска, геофизичка научна испитивања и развој одговарајућих технолошких метода за употребу јаловине у индустрији. REASONING се реализује кроз програм „Призма“ Фонда за науку Републике Србије и бави се испитивањем јаловине настале током вишедценијске експлоатације и флотацијске прераде руде из рудника „Рудник“. Примарни циљеви пројекта су карактеризација, развој технолошких метода за ефикасно издвајање корисних метала и других елемената присутних у јаловини и оцена

¹ УБ, Рударско-геолошки факултет, Бушина 7, 11000 Београд, Србија;
University of Belgrade, Faculty of Mining and Geology, Djusina 7, 11000 Belgrade, Serbia

² УБ, Институт за Физику, Београд, Србија
UB, Institute of Physics, Belgrade, Serbia

³ УБ, Институт за технологију нуклеарних и других минералних сировина, Београд, Србија
UB, Institute for Technology of Nuclear and Other Mineral Raw Materials, Belgrade, Serbia

⁴ УБ Институт за мултидисциплинарна истраживања, Београд, Србија
UB, Institute for Multidisciplinary Research, Belgrade, Serbia

⁵ УБ, Институт за општу и физичку хемију, Београд, Србија
UB, Institute of General and Physical Chemistry, Belgrade, Serbia

⁶ УБ, Технички факултет у Бору, Бор, Србија
UB, Technical Faculty in Bor, Bor, Serbia

могућности коришћења јаловине као сировине. Истраживања ће се реализовати кроз радне пакете који ће имати за циљ: мерење и праћење промена физичко-хемијских услова дуж вертикалног профила тела јаловине, корелацију магнетне суспензибилности са садржајем тешких метала са дубином и дефинисање зона обогаћивања, утврђивање дистрибуције минералних асоцијација, варијација и трендова у саставу порних вода, тешких метала и пратећих елемената, истраживање дистрибуције стабилне минералне асоцијације и дистрибуције елемената кроз теоријско геохемијско моделовање, проналазак оптималне шеме за екстракцију елемената теоријским приступом, електрохемијско одвајање елемената из синтетичких и лужених раствора и технолошки поступак валоризације корисних компоненти. Пројектни тим чине истраживачи из геологије, минералне индустрије, технологије, хемије из шест домаћих и једне научноистраживачке институције из иностранства. Пројекат REASONING ће допринети проучавању дистрибуције елемената, поступцима екстракције и начину коришћења јаловине из јаловишта рудника „Рудник“ а резултати овог пројекта требало би да буду и стандард за будућа истраживања и комплексну валоризацију јаловине из бројних јаловишта у Србији.

Кључне речи: Призма, REASONING, Рудник, јаловина, Фонд за науку Републике Србије

Abstract: The project Characterization and technological procedures for recycling and reuse of flotation tailings from the "Rudnik" mine (REASONING) is a multidisciplinary research project of tailings examination through geological, mineralogical, geochemical, geophysical scientific investigations and the development of appropriate technological methods for the use of tailings in industry. REASONING, funded by the "PRIZM" program of the Science Fund of the Republic of Serbia, examines tailings formed during decades of exploitation and flotation processing of ore from the "Rudnik" mine. The primary objectives of the project are characterization, development of technological methods for efficient extraction of useful metals and other elements present in tailings and assessment of the possibility of using tailings as raw material. Research will be carried out through work packages that will aim to: measure and track changes in physical-chemical conditions in the tailing with depth, correlate magnetic susceptibility with heavy metals contents with depth and to define the enrichment zones, determine distribution of mineral associations, variations and trends in the composition of pore waters, heavy metals and accompanied elements in tailing with depth, investigate distribution of stable mineral association and elements distribution through theoretical geochemical modeling, investigate optimal pathway for elements extraction through theoretical approach, investigate the electrochemical separation of elements from synthetic and leached solutions, and investigate technological procedure for the valorization of useful components from the tailing material. The project team consists of researchers from geology, mineral industry, technology, chemistry from six domestic and one scientific research institution from abroad. The REASONING project will contribute to the study of the distribution of elements, extraction procedures and the way of using tailings from the "Rudnik" mine tailings, and the results of this project should be a standard for future research and complex valorization of tailings in Serbia.

Key words: Prizma, REASONING, Rudnik, tailings, Science Fund of the Republic of Serbia

УВОД

Пројекат „Карактеризација и технолошки поступци за рециклажу и поновну употребу флотацијске јаловине из рудника Рудник“ (REASONING) реализује се кроз програм „Призма“ Фонда за науку Републике Србије. Пројекат REASONING представља мултидисциплинарни приступ проучавању флотацијске јаловине кроз геолошка, минералошка, геохемијска, геофизичка

научна испитивања и развој технолошких метода у сврху оцене могуће употребе јаловине у индустрији на конкретном примеру јаловине рудника „Рудник“.

Флотацијска јаловина представља материјал различитог састава који настаје у процесу вађења корисних елемената из руде. Овај нуспродукт рударства се акумулира у јаловиштима која се због природе материјала и начина одлагања означавају као потенцијална опасност по животну средину. Прерадом полиметаличне руде, пре свега Pb и Zn, настала је бројна јаловишта, која су друга по значају врста отпада у рударској индустрији у Србији.

Флотацијско јаловиште настало прерадом полиметаличне руде Pb, Zn, Cu и Ag из рудника „Рудник“ једно је од најпознатијих јаловишта у Србији. Одлагање јаловине је почело након изградње постројења за припрему и прераду руде 1953. године, када је изграђена и прва брана јаловишта. Од почетка прераде полиметаличне руде до данас у јаловишту рудника „Рудник“ депоновано је >11 милиона тона, односно ~7 милиона м³ нуспроизвода флотацијске прераде. Флотацијска јаловина рудника „Рудник“ представља финозрнаст материјал (-0,40+0,00 мм) састављен углавном од алуминосиликатних минерала кварца и фелдспата (>60%), мањих количина минерала глине (5-10%), док остатак представљају минерали Pb, Cu, Zn, Fe са констатованим присуством тешких метала до 1%.

Примарни циљ пројекта REASONING је минералшко-геохемијска карактеризација јаловине, развој технолошких метода за ефикасно издвајање корисних метала и других елемената пронађених у флотацијској јаловини рудника „Рудник“ и оцена могућности коришћења преосталог дела јаловине као сировине за грађевинску индустрију. У сврху реализације овог циља спровешће се комплексна мултидисциплинарна истраживања кроз учешће истраживача из домаћих и страних научноистраживачких организација. Стечена знања и предложене методе биће основа за развој алтернативних еколошки прихватљивих метода екстракције корисних компоненти из јаловине и коришћења јаловине као минералне сировине.

ЦИЉЕВИ ПРОЈЕКТА И ИСТРАЖИВАЧКИ ТИМ

Циљеви истраживачког пројекта REASONING су:

Мерење и праћење промена физичко-хемијских услова дуж вертикалног профила тела јаловине. Мерења физичко-хемијских параметара порних флуида у јаловишту (рН, редокс потенцијал (ORP), температура, проводљивости, укупни растворени кисеоник, отпорност раствора).

Корелација магнетне суспензибилности са садржајем тешких метала са дубином и дефинисање зона обогаћивања. Одређивање магнетне суспензибилности ниског поља и планирање додатних минералшких и геохемијских истраживања у зонама магнетних аномалија које указује на високо обогаћење тешким металима.

Утврђивање дистрибуције минералних асоцијација, варијација и трендова у саставу порних вода, тешких метала и пратећих елемената у јаловини са дубином. Минералшка и геохемијска истраживања: комбинацијом

инструменталних техника, оптичке микроскопије, електронске микроскопија (SEM/EDS), рендгенске дифракције (XRD), FTIR, RAMAN и ICP-OES спектроскопије. Одређивање дистрибуције секундарних минералних асоцијација по дубини биће упоређено са резултатима теоријског моделовања у циљу дефинисања зоне секундарног обогаћивања елемената.

Истраживање дистрибуције стабилне минералне асоцијације и дистрибуције елемената кроз теоријско геохемијско моделовање. Примена теоријског моделовања у сврху дефинисања врсте новонасталих минералних фаза и присуства стабилних минералних асоцијација. Планирана је употреба софтверских пакета PHREEQC и PhreePlot за теоријско геохемијско моделовање водених раствора, прорачун индекса засићености и конструисање Pourbaix дијаграма, одређивање поља стабилности јонских врста у воденим растворима који транспортују тешке метале и других елемената као и поља стабилности минералних и аморфних фаза

Установљавање оптималне шеме за екстракцију елемената теоријским приступом. Теоријско моделовање оптималног начина екстракције елемената из луженог раствора. Моделовање лужених раствора вршиће се у сврху екстракције фаза жељеног састава. Успостављени начини екстракције биће тестирани на синтетичким растворима путем електрохемијских експеримената да би се потврдила валидност резултата теоријског моделовања.

Електрохемијско одвајање елемената из синтетичких и лужених раствора. Електрохемијски експерименти на синтетичким растворима. Експеримент ће укључивати тестирање различитих комбинација материјала и облика електрода, одређивања услова за проналажење одговарајућих процедура за ефикасно одвајање и екстракцију елемената. Биће извршено и тестирање на стварним растворима излуженим из јаловине како би се показало да је утврђена процедура валидна.

Технолошки поступак валоризације корисних компоненти из јаловине. Поред електрохемијске екстракције, биће изведени алтернативни приступи за рекулерацију ПГЕ, црних и обојених метала како би се добио „чист” јаловински материјал пречишћен од глине, присутних токсичних црних и обојених метала у флотацијској јаловини. Циљ је да се произведу концентрати вредних метала, као и да се произведе чврсти остатак без тешких метала који се састоји од силикатних, алумо-силикатних и оксидних минерала који неће представљати опасност за животну средину и који се као такав може даље користити у грађевинарству и путној индустрији. Овај циљ ће се остварити у низу од три подактивности: 1. успостављање процедура валоризације ПГЕ, 2. успостављање процедура валоризације обојених метала које садрже одговарајући минерали и 3. верификација потенцијалне примене алумосиликатног јаловинског материјала за путну индустрију.

За реализацију предложених циљева пројекта формиран је тим истраживача из области геологије, минералне индустрије, технологије, хемије, који поседују искуства и резултате претходних сличних истраживача. Истраживачки тим чине истраживачи из шест домаћих научно-истраживачких институција и једна из иностранства (Немачка) (Табела 1). Шеф истраживачког

тима пројекат REASONING је Проф. Др Владимир Симић са Рударско-геолошког факултета Универзитета у Београду.

Табела 1. Истраживачки тим пројекта REASONING

Институције	Истраживачи
Универзитет у Београду Рударско-геолошки факултет	Проф. Др Владимир Симић
	Проф. Др Весна Цветков
	Проф. Др Милена Костовић
	Стефан Петровић
Институт за технологију нуклеарних и других минералних сировина	Др Драган Радуловић
	Др Јовица Стојановић
	Др Владимир Јовановић
	Др Дејан Тодоровић
Универзитет у Београду - Институт за мултидисциплинарна истраживања	Др Нина Николић
Универзитет у Београду - Институт за општу и физичку хемију	Др Јелена Сенћански
Универзитет у Београду - Технички факултет Бор	Проф. Др Грозданка Богдановић
	Драгана Мариловић
Универзитет у Београду - Институт за физику	Филип Арнаут
Универзитет RWTH Ахен, Немачка Факултет за георесурсе и инжењерство материјала	Др Срећко Стопић

РАДНИ ПАКЕТИ ПРОЈЕКТА

Минералозна и геохемијска карактеризација јаловине. Почетак реализације пројекта подразумева прикупљање узорака за геохемијска, геофизичка и минералозна испитивања из четири истражне бушотине дубине ~30 метара. Посебна пажња биће посвећена спровођењу мерења и анализа дефинисаних кроз циљеве 1-4, јер ће они дати основу за даља истраживања. Ово се посебно огледа у одређивању садржаја елемената и процени физичко-хемијских услова у јаловини, јер ће ти подаци бити коришћени за планирање корака током електрохемијске екстракције и избор одговарајућих мембрана. Физичко-хемијски параметри ће се мерити помоћу вишепараметарског инструмента и применом различитих анализа које ће пружити свеобухватно разумевање промена физичко-хемијских својстава на различитим дубинама јаловишта. Добијене информације су од значаја за одређивање стабилности минералних асоцијација, дистрибуције метала и других елемената. Истраживање ће допринети корелацији између магнетне суспектибилности и концентрације тешких метала, а мерења ће се фокусирати на интервале од 10 цм да би се идентификовала подручја обогаћивања у високој резолуцији. Подаци високе резолуције добијени овим истраживањем пружаће информације за накнадне анализе (спроведене у нижој резолуцији), које се односе на обогаћене зоне. Претходно приказани подаци о физичко-хемијским параметрима, као и

анализирани хемијски састав порозних течности, послужиће као улазни параметри за теоријско моделовање стабилних минералних асоцијација у нормалним условима.

Електрохемијско одвајање и екстракција елемената. Други пакет обухвата теоријско моделовање екстракције елемената из луженог раствора, као и тестирање електрохемијске сепарације и екстракције. Циљ теоријског моделовања елемената из излужених раствора је одређивање најефикаснијих метода за екстракцију ових елемената. Почетна фаза тестирања укључиваће преглед лабораторијски припремљених синтетичких раствора користећи методологију покушаја и грешака (*engl. Trial-and-test*). Процес електрохемијске сепарације и екстракције ће се спроводити на припремљеним растворима који садрже жељене елементе у облику сулфата (нпр. Pb, Zn, Cu, Fe, Mn, Al, Cd, Ni) у односима одређеним из хемијских анализа порозних течности. Примарни циљ радног пакета је одредити најефикасније методе за сепарацију и екстракцију елемената.

Технолошка истраживања и поступци за валоризацију корисних компоненти из јаловине. Овај пакет се састоји од три одвојене подгрупе, свака фокусирана на специфичне процедуре везане за валоризацију ПГЕ и обојених метала, као и рециклирање и примену флотацијских јаловина у друмској индустрији. Узорци од 150 кг ће се прикупљати за процену ПГЕ и накнадне процесе магнетске сепарације, млевења, и лабораторијских тресућих столова. Поступци за процену валоризације обојених метала, врши ће се на узорцима тежине од 350 кг, укључују процес испирања који се спроводи под природним атмосферским климатским условима. Током експеримената, вршиће се праћење рН, сланости, температуре и концентрације елемената у раствору током времена. Процес одвајања минерала глине из ће се спроводити користећи сита. Десилтовани узорак ће проћи кроз магнетску сепарацију користећи различите параметре магнетних сепаратора. На крају, процедуре рециклирања и примене флотацијских јаловина у друмској индустрији ће проћи темељно физичко и механичко тестирање како би се утврдила њихова погодност за изградњу путева.

ЗАКЉУЧАК

Предложена методологија пројекта REASONING никада није примењена у изучавању флотацијске јаловине Србији. Нови концепт и приступ изучавању огледа се у примени најсавременијих аналитичких метода и дефинисању технолошких поступака рециклаже јаловине. Пројекат REASONING ће обезбедити сазнања о вертикалној и бочној дистрибуцији главних и микроелемената и пратећој геохемијској асоцијацији у флотацијској јаловини рудника „Рудник“ а добијени резултати биће основа за примену одговарајућих технолошких поступака прераде. Реализација пројеката ће имати позитиван утицај на научну заједницу, друштво, привреду, индустрију, образовање и животну средину. Резултат овог пројекта треба да оствари мултипликативни ефекат коришћења јаловине на начин да се минимизирају сви негативни еколошки и социјални утицаји, а максимизирају позитивни утицаји у смислу односа индустрије, локалне заједнице и регионалног економског развоја.

Предложена методологија и очекивани резултати пројекта REASONING требало би да буду стандард за будућа истраживања и комплексну валоризацију јаловине у Србији.

Захвалност

Истраживање спроведено уз подршку Фонда за науку Републике Србије, Пројекат 7522, Карактеризација и технолошки поступци за рециклажу и поновну употребу флотацијске јаловине рудника „Рудник“ - REASONING / This research was supported by the Science Fund of the Republic of Serbia, #Grant No 7522, Characterisation and technological procedures for recycling and reusing of the Rudnik mine flotation tailings – REASONING.

Landslide Investigations Using the geoelectrical Scanning Method and Ground Penetrating Radar - case Studies

FILIP D. ARNAUT, University of Belgrade,

Faculty of Mining and Geology, Belgrade

BRANISLAV B. SRETENović, University of Belgrade,

Faculty of Mining and Geology, Belgrade

Professional paper

UDC: 624.131.537(497.11)

550.8.08(497.11)

DOI: 10.5937/tehnika2104419A

Geotechnical investigations of landslides for future mitigation are a complex task. To obtain relevant information, various geophysical methods are used, with varying degrees of success. The geoelectrical scanning method, also known as Electrical Resistivity Tomography (ERT) has been successfully applied in several locations in Serbia in the last three decades. The geoelectrical scanning method was used during investigations of landslides: Umka, Lukovska banja, Tara, and Trandžament, while both geoelectrical scanning and Ground Penetrating Radar (GPR) were utilized at the Trandžament landslide. The achieved results from both methods are mutually correlated at the Trandžament landslide. Correlation of GPR data with geoelectrical scanning data was only possible since there were no near-surface low resistivity zones in the Trandžament landslide body. Otherwise, electromagnetic signal attenuation would be high in the presence of near-surface low resistivity zones, and a quality signal would be impossible to detect at the receiving antenna.

Key words: *Electrical Resistivity Tomography (ERT), Ground Penetrating Radar (GPR), Landslide geophysical surveys*

1. INTRODUCTION

Landslides are a phenomenon that can be categorized in a variety of ways, the most common being according to landslide activity, rate of movement, water content, material and type of movement [8]. From a geophysical standpoint, two important aspects can be singled out - sliding surface and the contrast between the sliding mass and bedrock. The first aspect can be subdivided into a unique slide surface or a zone with several sliding surfaces, while the second one is based on the physical and mechanical properties of the sliding mass, sliding surface, bedrock and groundwater distribution.

In this research paper, two main types of landslides are presented - Landslides in Neogene sediments (mainly clays, marls and sands) at the brim of the Panonian basin, and landslides in deluvial material that lie over older Mesozoic and Paleozoic bedrock.

Further surveys in the Neogene sediments enabled the discovery of two forms of landslides - Landslides on the right riverbank of the Sava and Danube River and, landslides in Central Serbia (inland Neogene) [1].

Landslides on the right riverbank of the Sava and Danube River are formed in clay and marl weathering zones and slide over undisturbed sediments. There is a distinct sliding zone. In terms of hardness and water content, there are major differences between the sliding mass and the bedrock. The Umka landslide is one of the most representative landslides that form on the right riverbank of the Sava and Danube River [1]. The Umka landslide was investigated by various geotechnical, geophysical and geological methods and extensive data was collected [1]. Investigation by permanent GNSS monitoring categorized the Umka landslide as slow to very slow [5].

The second type of landslide is distinguished by significant variations in composition, physical, and mechanical properties between the sliding mass and the bedrock. The consolidated bedrock, which is mainly made up of crystallized slates, limestones, and flysch (landslides „Tara“ and „Lukovska banja“), slides over the clay mass, which includes more or less fragments of the bedrock [7].

Author's address: Filip Arnaut, University of Belgrade, Faculty of Mining and Geology, Belgrade, Đušina 7

e-mail: G601-20@rgf.bg.ac.rs

Paper received: 21.04.2021.

Paper accepted: 11.08.2021.

2. METHODOLOGY

In this study are presented results of two well-known geophysical methods already applied on aforementioned landslides: Electrical Resistivity Tomography (ERT) and Ground Penetrating Radar (GPR).

The direct current resistivity method can be performed as 1D resistivity sounding, 2D investigations also known as electrical resistivity tomography or 3D technique. It is a conventional geophysical method that uses two current and two potential electrodes. Current electrodes inject direct current into the subsurface, and potential electrodes measure the resulting potential difference. Measurements can be taken inline (1D and 2D measurements) or over an area (3D measurements), and the results can be graphically shown as sounding curves or pseudosections of apparent resistivity [9] or apparent resistivity maps.

The electromagnetic reflective method (ERM), often known as ground penetrating radar, is used to study structures at relatively small depths. The GPR method relies on the reflection of high-frequency electromagnetic waves (8 MHz to 4 GHz) at various subsurface medium boundaries, where the dielectric constant (permittivity) and resistivity change. While conducting GPR investigations, depth of penetration and resolution have to be taken into account. Higher transmitter antenna frequency yields better resolution, on the other hand, lower frequency increases the depth of penetration [3].

3. CASE STUDIES

3.1. Landslide „Umka“

Umka is a Belgrade suburb that stretches for a few kilometers along the Sava River. The entire estate is built on top of a landslide that has varying depths to slip surface (block structure) [1], which causes pro-

blems such as house and road collapse. The landslide has also impacted the local road that connects Belgrade and Obrenovac. Due to the presence of a landslide that could not be mitigated, the newly planned road had to be rerouted to the other side of the Sava River. ERT measurements were conducted during the geotechnical investigation of the landslide zone Umka and Duboko for the preliminary design of the E-76 Belgrade-Obrenovac highway (1991/1992). The ERT profiles had a total length of 2 kilometers [1, 2].

The Umka landslide is one of the most investigated landslides in Serbia. Various methods (permanent GNSS monitoring, photogrammetry, geotechnical and geophysical investigations, boreholes and explorations shafts etc), have been employed by [1, 2, 5], in order to characterize the landslides causative and triggering factors, overall geometry etc. Research by [1] displayed that the Umka landslide is a slow to very slow landslide with a reactivation period of several years to several decades.

Even though the sliding zone is distinct, and the differences in geotechnical properties (hardness and water content) between the sliding mass and the bedrock are large, an essential physical property (resistivity contrast) is relatively small (2.5: 1). In those circumstances, the geoelectrical scanning method's performance was debatable.

The scanning profile and the exploration shaft (located on the „A Block“ of the Umka landslide [1]) is depicted in Figure 1. The apparent resistivities calculated in the nearby exploration shaft show a strong correlation with the 2D interpretation. This form of study shows the presence of zones with even lower resistivity values (around 2 Ω m at the 50th and 60th meter) than the zone associated with water leakage in the exploration shaft (around 4.5 Ω m).

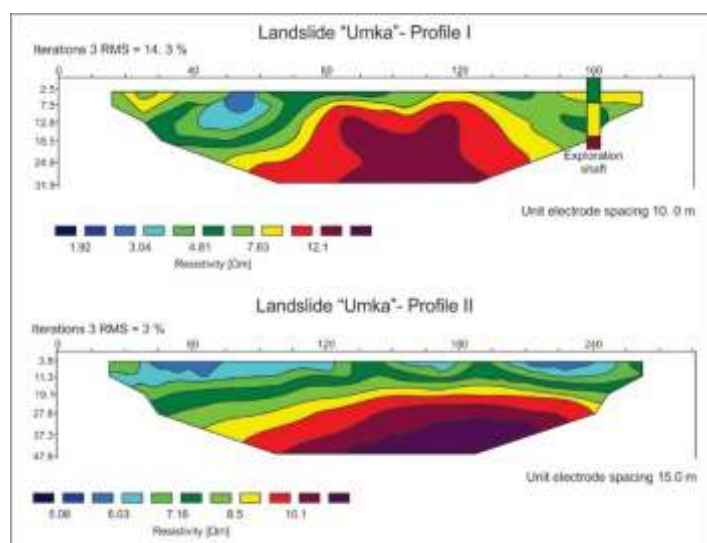


Figure 1 - Results of 2D inversion of ERT data at „Umka“ landslide [7]

In comparison to conventional geotechnical methods such as boreholes and exploration shafts, the 2D interpretation of the geoelectrical scanning method provided valuable experience that successfully demonstrated the application of the geoelectrical scanning method, which showed improved survey area coverage (2.2 kilometres were scanned) while also significantly reducing survey expense.

3.2. Landslide „Lukovska Banja“ and „Tara“

The second type of landslide is described by significant differences in composition, physical, and mechanical properties between the sliding mass and bedrock. The clay mass, which includes bedrock fragments, slides along the solid bedrock in this case. The

geoelectrical scanning method was used to survey these two types of landslides along parallel profiles, which helped with the assessment of 2D model divergence during inversion.

The geological structure of the landslide „Lukovska banja“ (figure 2) is complex. The landslide's body is made up of fragments of marly, shaly limestones, marls, metasandstone, clayey, and chlorite-sericite shales.

The presence of a 3D block of rock with a volume of up to 10 m³ was discovered using geological mapping and borehole data. The landslide's bedrock is serpentinite, as well as metamorphosed peridotites, shaly limestones, clays, and marls.

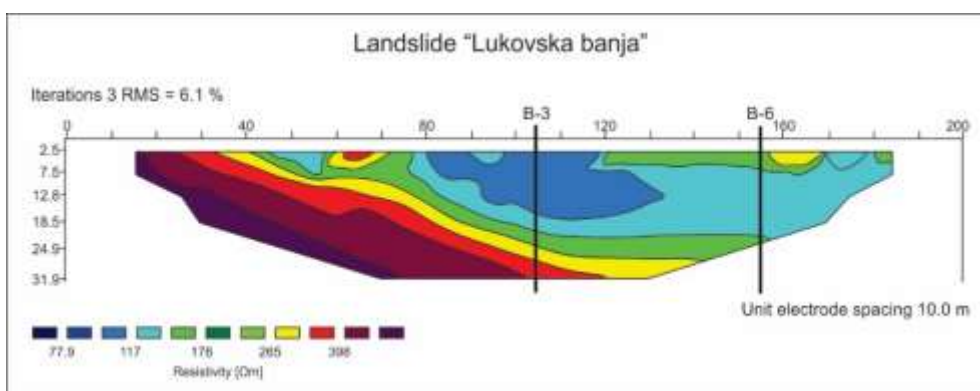


Figure 2 - Result of 2D inversion of ERT data at „Lukovska banja“ landslide [7]

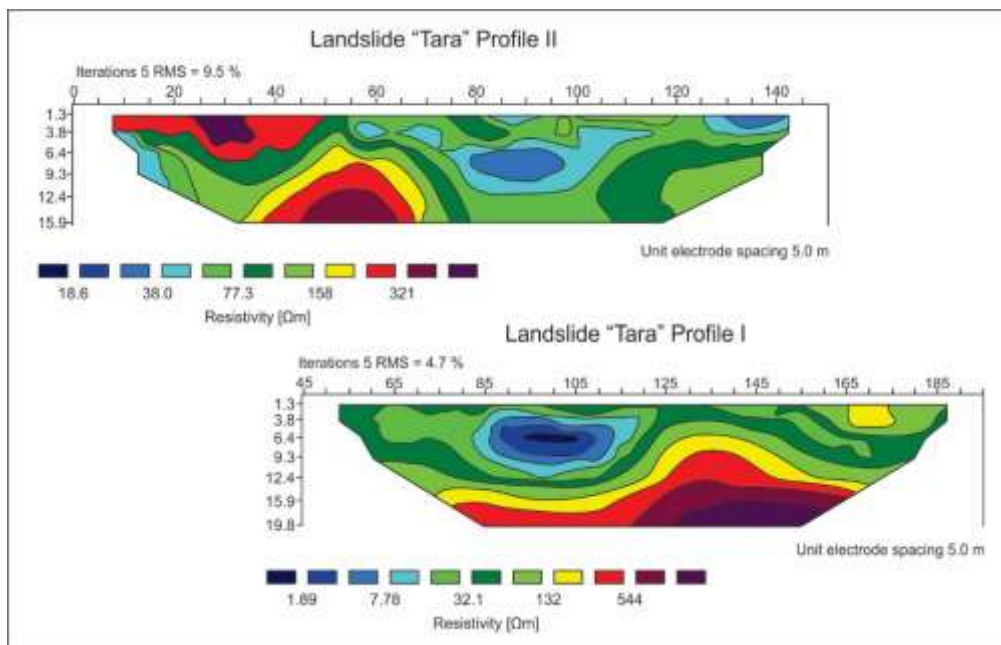


Figure 3 - Result of 2D inversion of ERT data at „Tara“ landslide [7]

The landslide „Tara“ was surveyed during its remediation (figure 3), that is, when a concrete retaining wall was installed. The survey's aim was to check the depth to the bedrock, which had already been estimated prior. The fresh concrete injected during the

measurement period is most likely to blame for the exceptionally low resistivity values on profile I. Under the road, profile I was measured at the same level as the concrete injection site, but five meters below profile II. The retaining wall's founding quality was

determined by measurements. Zones of low apparent resistivity can be seen (from the 125th meter) outside the landslide zone. In order to reduce the impact of topography, Profile II was located on the road over the landslide on the side away from the landslide. When taking into account that profile I began at the 45th meter of profile II, a good correlation with profile I can be made. Since profile II is closer to the landslide scar, the depth to the bedrock is lower. The achieved root-mean-square error is 9.5 percent, which is consistent with topography's effect on measurements.

A preview of the „Umka“, „Lukovska banja“ and „Tara“ landslides is also given in [6] and [7]

3.3. Landslide „Trandžament“

Since the landslide occurred along a road underneath the Petrovaradin fortress, various materials (gravel, slag, etc.) were used as a filling to keep the road working, this landslide could also be viewed as a wild waste dump. Low resistivity marls provide a strong geological basis for GPR surveys when a high resistivity surface layer is present. A 500 MHz antenna was used to perform the GPR surveys. The landslide „Trandžament“ is situated on the road underneath Petrovaradin fortress. The formation is thought to have been aided by a drainage system that drains excess water from the fortress to the Danube. This drainage system, which is 1 meter in diameter and runs beneath the road, dates from the Austro-Hungarian Empire. Since the drainage system is made of bricks, it leaks and creates an unstable zone around it when a greater amount of water passes in. Filling the road with slag and gravel has been undertaken for years in order to keep the road in good shape.

A GPR and geoelectrical survey were conducted to prove the hypothesis about the landslide's causative and triggering factors. Given the relatively high resistivity values of the surface layer and the low resistivity values of the drainage system with water and the underlying clay and marl layers, it was believed that the drainage system full of water would be a good target for GPR investigations. In comparison to the GPR survey, geoelectrical scanning has good survey resolution, so identification of the wet zone and a deeper depth of investigation were expected.

Geoelectrical scanning measurements were taken along two parallel profiles, each 125 meters long, on the left and right sides of the road. There were 26 electrodes, with a unit electrode spacing of 5 meters, and six depth levels, allowing for a depth of investigation of around 15 meters. The drainage system was located in the center of the profile.

A 500 MHz antenna was used to conduct a GPR survey from a vehicle along a profile line that was a few kilometers long. The vehicle was moving along the road, closer to the Danube River and profile II of geoelectrical scanning. Figure 4 shows a GPR section of the landslide, as well as the position of the hyperbolic reflection caused by diffraction from the drainage system. There is significant electromagnetic attenuation in the drainage region, in contrast to the areas of high resistivity filled material. Similarly, from the 38th to the 43rd meter of the GPR section, multiple reflections can be seen that expand along the entire section with a decrease in amplitude. These multiple reflections reveal good reflectors in the area around the antenna, i.e., near the asphalt lane.

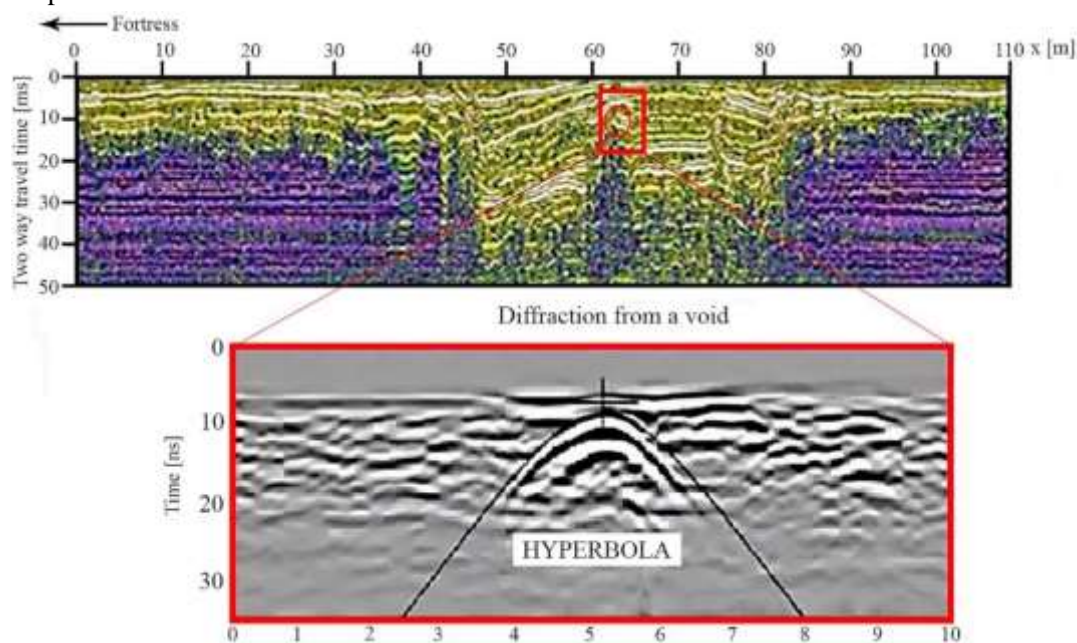


Figure 4 - GPR survey in the Trandžament landslide zone

The 2D model obtained with geoelectrical scanning along profile I, which is closer to the fortress, is shown in the upper part of figure 5.

A low resistivity zone was detected in the middle of the profile, which may indicate the position of the damp zone around the drainage system, while a high resistivity zone was detected above the drainage

system, which reflects the filled material. The attenuation anomaly zone (62nd meter of the section at 1.5-meter depth) at the position of high resistivity materials that allow quality reflections is also visible in the GPR section. The overall geometry of the filled material as well as the landslide is determined using geoelectrical scanning method.

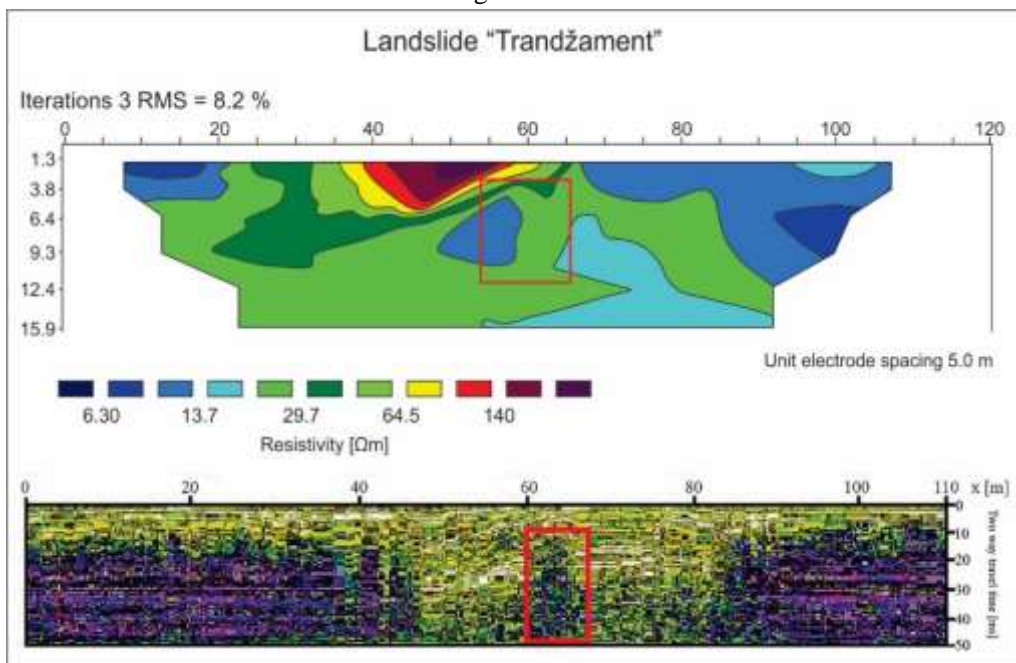


Figure 5 - Results of GPR and ERT surveys along profile I

Figure 6 shows a comparison of geoelectrical scanning profile II and the GPR section. Since the GPR section and the geoelectrical scanning profile are closer together, a better correlation in the form of the length and geometry of the filled material can be established.

Geoelectrical scanning, like the previous example, provides a better depth of investigation and overall geometry, especially in low resistivity clayey marls. The electromagnetic signal was completely attenuated in those zones, in comparison.

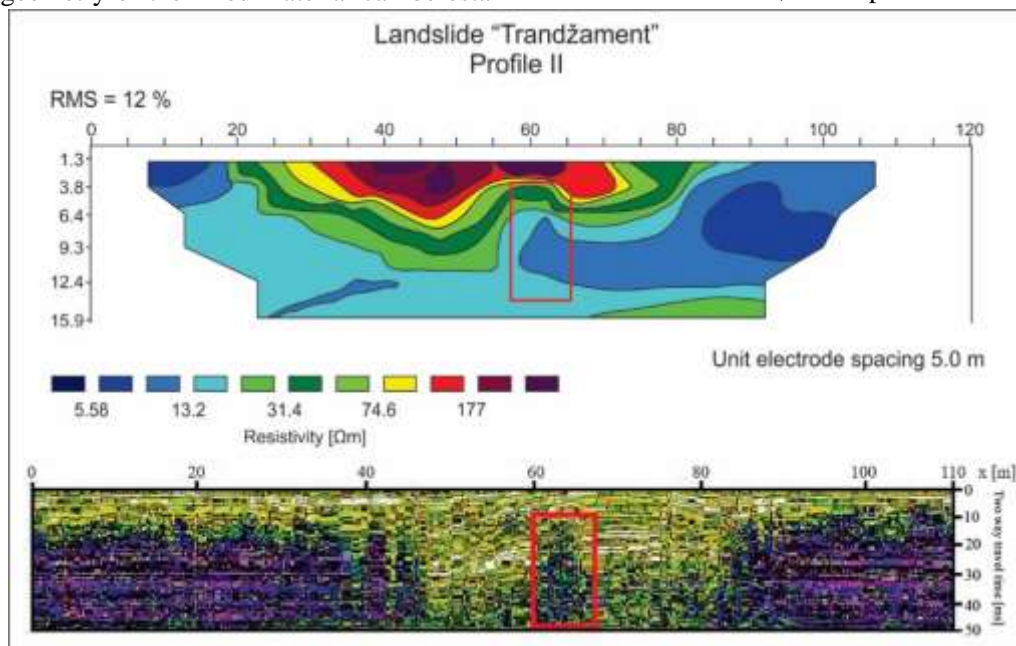


Figure 6 - Results of GPR and ERT surveys along profile II

The landslide extends along profile II (figure 6) from the 20th to the 80th meter, with a maximum depth of about 10 meters. At the end of the profile, the resistivity values decrease with depth as well as laterally. In other words, the lower resistivity values in the right part of the profile decrease to 6 Ωm . It's worth noting that the geoelectrical scanning results are skewed by lateral and topography effects. This was reflected in the RMS value, which was 8% for profile I and 12% for profile II. The existence of 3D inhomogeneities did play a role in the deviation from 2D conditions. The topography effects are caused by notches on both sides of the road, which strike in a relatively gradual pattern along the longest part of the geoelectrical scanning profiles.

4. DISCUSSION AND CONCLUSION

A multi-geophysical approach should be used when dealing with a complex subsurface structure such as landslides. GPR and ERT measurements provide additional data that can substantially aid in the interpretation process and provide valuable insight into the subsurface architecture, which can further complement geotechnical data. Even when dealing with extremely attenuated GPR sections, by correlating them with ERT profiles, some useful near-surface information can be retrieved, as illustrated in [4]. When dealing with landslides, high density data acquired quickly is preferable [3]. Another advantage of integrating GPR and ERT data is the lower investigation cost compared to other investigation methods (boreholes, exploration shafts etc.).

The Umka landslide demonstrated that without supplementary data such as exploration shafts, boreholes, or other geotechnical or geophysical data, the ERT method could be debatable even when the sliding zone is distinct, but the resistivity contrast is low. Furthermore, the Umka landslide demonstrated that in order to completely identify such a complex geological setting, a multidisciplinary approach and a variety of methods should be used. The extensive data collected from permanent monitoring stations is quite useful. This would also apply to continuous monitoring using the ERT method at various times of the year.

The Lukovska banja landslide revealed useful information on the landslide's composition when ERT was cross-referenced with geological mapping and drilling data. The approximation of the bedrock volume was made possible due to additional data.

Landslide Tara revealed a high correlation between the two parallel ERT profiles, but also demonstrated that topography has an effect on measurements (increased RMS value of 9.5 percent) and should be considered.

2D geoelectrical surveys are successfully correlated with ground penetrating radar surveys in the absence of near-surface low resistivity zones („Trandžament“ landslide). Ground penetrating radar surveys have a limited depth of investigation due to signal attenuation in low resistivity zones. Topography had a non-negligible effect in the RMS error in the Trandžament landslide, as well in the Tara landslide case. Apart from the topography effect, the presence of 3D inhomogeneities also contributed to the deviations from 2D conditions, hence increasing the RMS error value.

GPR will never be the first choice in landslide studies when other geophysical methods are available, as [4] noted, but it has proven to be a valuable auxiliary (supplementary) method. The 2D electrical imaging technique offers satisfactory results when dealing with various engineering geology problems, and, because of its relatively low cost and fast acquisition time, should accompany geotechnical data.

REFERENCES

- [1] Abolmasov B, Milenković S, Marjanović M, Đurić U. and Jelisavac B, A geotechnical model of the Umka landslide with reference to landslides in weathered Neogene marls in Serbia. *Landslides*, 12(4), pp.689-702, 2015.
- [2] Đurić U, *Kvantitativna procena rizika od klizišta Umka kod Beograda*. Doktorska disertacija, Univerzitet u Beogradu, Rudarsko-geološki fakultet. p 254, 2020.
- [3] Ristić A, Abolmasov B, Govedarica M, Petrovački D. and Ristić A, Shallow-landslide spatial structure interpretation using a multi-geophysical approach. *Acta Geotechnica Slovenica*, 9(1), pp.47-59, 2012..
- [4] Sass O, Bell R. and Glade T, Comparison of GPR, 2D-resistivity and traditional techniques for the subsurface exploration of the Öschingen landslide, Swabian Alb (Germany). *Geomorphology*, 93(1-2), pp.89-103, 2008.
- [5] Samardžić-Petrović M, Popović J, Đurić U, Abolmasov B, Pejić M. and Marjanović M, Permanent GNSS monitoring of landslide Umka. In *XIV International Conference On Contemporary Theory And Practice In Construction XIV Stepgrad XIV Proceedings, 2020*. University of Banja Luka Faculty of Architecture, Civil Engineering and Geode, 2020..
- [6] Sretenović B, Marčetić D, Determination of the internal geometry of a landslide using electrical tomography. *54th Meeting of the EAEG*: Paris, France pp. 354-355, 1992.

- [7] Sretenović B, Lokin P. i drugi, Some possibilities of electrical scanning method in landslide exploration, *6th Meeting of Environmental and Engineering Geophysics*, September 3-7,2000, Bochum-Germany, 2000.
- [8] Varnes D. J. and Cruden, D. M, *Landslide types and processes. Landslides: investigation and mitigation*, Transportation Research Board special report, 247, 1996.
- [9] Zhou B. and Kanl I, *Electrical resistivity tomography: a subsurface-imaging technique. In Applied geophysics with case studies on environmental, exploration and engineering geophysics*. London, UK: IntechOpen.

REZIME

ISTRAŽIVANJE KLIZIŠTA METODAMA GEOELEKTRIČNOG SKENIRANJA I GEORADARA – NEKI PRIMERI

Istraživanje klizišta u cilju projektovanja sanacionih mera predstavlja kompleksan geotehnički zadatak. Pri tome, različite geofizičke metode se koriste radi dobijanja relevantnih podataka sa različitim stepenom uspeha. U prethodne tri decenije, metoda geoelektričnog skeniranja je uspešno primenjena tokom istraživanja većeg broja klizišta. U radu su prikazane studije slučaja klizišta: Umka, Lukovska banja, Tara i Trandžament, metodom geoelektričnog skeniranja i elektromagnetometrijskom metodom (georadar). Dobijeni rezultati obe metode su međusobno korelisani na primeru klizišta Trandžament. Iako georadarska metoda ima relativno male vrednosti dubinskog zahvata, korelacija je bila omogućena u slučaju nedostatka bliskopovršinskih niskootpornih zona (klizište Trandžament). U slučaju prisustva bliskopovršinskih niskootpornih zona, atenuacija elektromagnetnih signala je bila prevelika za dobijanje kvalitetnog korisnog signala na prijemnoj anteni.

Ključne reči: *Metoda geoelektričnog skeniranja (ERT), georadar (GPR), istraživanje klizišta*

РЕТРОСПЕКТИВНА СТУДИЈА О ПОВРЕДАМА У САОБРАЋАЈНИМ НЕЗГОДАМА У РЕПУБЛИЦИ СРБИЈИ ТОКОМ ЗАТВАРАЊА УСЛЕД ПАНДЕМИЈЕ COVID-19

Александар Медаревић,¹ Катарина Боричић,¹ Филип Арнаут²

¹ Институт за јавно здравље Србије „Др Милан Јовановић Батут”, Београд, Србија

² Универзитет у Београду, Институт за физику, Београд, Србија

A RETROSPECTIVE STUDY ON ROAD TRAFFIC INJURIES IN THE REPUBLIC OF SERBIA DURING THE COVID-19 LOCKDOWN

Aleksandar Medarević,¹ Katarina Boričić,¹ Filip Arnaut²

¹ Institute of Public Health of Serbia “Dr Milan Jovanović Batut”, Belgrade, Serbia

² University of Belgrade Institute of Physics, Belgrade, Serbia

Сажетак

У овој студији испитивали смо појаву и карактеристике опасних ситуација у саобраћају (енгл. RTI) у Србији током периода затварања у априлу 2020. године. При статистичкој анализи коришћени су јавни подаци Агенције за безбедност саобраћаја Републике Србије, Завода за статистику и Универзитета Џонс Хопкинс. У 2020. години дошло је до општег опадања укупног броја повређених и погинулих у поређењу са претходном годином, и то за 15,5% односно 8%. Највећи пад забележен је у априлу када су на снази биле најстроже мере затварања. Број повређених и погинулих лица био је за 55% и 49% нижи у односу на исти месец 2019. године. Највеће смањење броја повреда у обе категорије током априла забележено је током периода затварања од 20.00 часова увече до 4.00 часа ујутро, као и викендом.

Кључне речи: COVID-19, саобраћајне незгоде, саобраћајни судари, затварање

Abstract

This study examined the occurrence and characteristics of road traffic incidents (RTIs) in Serbia during the lockdown period in April 2020. Statistical analysis utilized public data from the National Traffic Safety Agency of Serbia, the Bureau of Statistics, and Johns Hopkins University. In 2020, there was a general decline in the total number of non-fatally and fatally injured when compared to the preceding year by 15.5% and 8%, respectively. The greatest decrease was noted in April, which was the period of the most stringent lockdown measures. The number of non-fatally and fatally injured persons was 55% and 49% lower compared to the same month in 2019. The largest decrease in the number of injuries in both categories during April was recorded during the closing period from 8 pm to 4 am and on weekends.

Keywords: COVID-19, Traffic accidents, Road traffic collisions, Lockdown

Увод

Извештај Светске здравствене организације о стању безбедности на путевима показује да су опасне ситуације у саобраћају (енгл. RTI) проузроковале 1,35 милиона смртних случајева, као и 50 милиона повреда и инвалидитета [1, 2]. На појаву опасних ситуација у саобраћају пре свега утичу природа и количина саобраћаја, заједно са факторима животне средине и „људским фактором” [1, 3]. Иако је током последњих година дошло до опадања, опасне ситуације у саобраћају и даље остају осми водећи узрок смрти у свету [1]. Међутим, оне су примарни узрок смртности код деце и младих [1]. Већина случајева са смртним исходом догодила се у земљама са ниским и средњим нивоом прихода [1]. Поред тога, процењује се да је између 20 и 50 милиона људи доживело повреде без смртог исхода у друмском саобраћају [1]. Ове повреде обухватају трауматске повреде мозга и кичмене мождине, фрактуре, ампуације, посттрауматски стресни поремећај, анк-

Introduction

The World Health Organization's report on road safety reveals that road traffic injuries (RTIs) caused 1.35 million fatalities with 50 million injuries and disabilities [1, 2]. The occurrence of RTIs is primarily influenced by the quantity and nature of traffic, along with environmental factors and the “human factor” [1, 3]. Although there has been a decline in recent years, RTIs remain the eighth cause of death globally [1]. However, they are the primary cause of mortality for children and young adults [1]. The majority of fatalities occurred in nations with lower and middle levels of income [1]. Furthermore, an estimated 20 to 50 million individuals experienced non-lethal road traffic injuries [1]. The injuries encompass traumatic brain and spinal cord injuries, fractures, amputations, post-traumatic stress disorder, anxiety, and depression. Accidents of this nature can cause injuries that result in enduring disabilities, which can impact one's ability to move, physical strength, and overall physical functioning. Severe non-fatal injuries can impede

сиозност и депресију. Незгоде ове врсте могу бити узрок повреда које доводе до трајног инвалидитета, што може утицати на способност кретања, физичку снагу и опште физичко функционисање појединца. Тешке повреде без смртног исхода могу ограничити радну способност особе и могућност запослења, што доводи до финансијске нестабилности [4]. Трошкови повезани са лечењем, рехабилитацијом и континуираном негом могу бити значајни, што ствара финансијски притисак на пацијенте и њихове породице. Чен и др. су утврдили да су укупни трошкови повреда са и без смртног исхода широм света износили око 1,8 трилиона USD током периода од 15 година [5]. Ово је представљало приближно 0,12% глобалног бруто домаћег производа [5]. У неким земљама, овај удео може достићи чак 3% домаћег производа [1].

Светска здравствена организација је 2020. године званично прогласила епидемију новог коронавируса као ванредну ситуацију у области јавног здравља од међународног значаја [1, 6]. Овом објавом СЗО је позвала међународну заједницу да се уједини и осмисли ефикасне стратегије за заустављање преношења новог корона вируса. У знак одговора, власти су спровеле низ противепидемијских мера, као што су одржавање физичке дистанце, изолација, издавање наређења да људи остану код куће, затварање школа, увођење ограничења путовања и великих окупљања, привремено заустављање услуга које нису од суштинског значаја и спровођење различитих нивоа затварања [1]. Као резултат уведених рестриктивних мера, обрасци саобраћаја су неизбежно измењени, што је довело до смањења потрошње горива и смањења загађења ваздуха у градским срединама [7, 8].

Први документовани случај COVID-19 званично је у Србији забележен 6. марта 2020. године. Дана 15. марта проглашено је ванредно стање што је резултирало затварањем свих образовних институција, укључујући школе, вртиће и универзитете. Након три дана, Министарство унутрашњих послова увело је полицијски час широм земље за особе старије од 65 година, када им је било забрањено да напуштају своје домове. Сви ресторани, кафићи, тржни центри и јавни превоз затворени су 20. марта. Влада Републике Србије донела је 21. марта одлуке о забрани уласка у Србију и привременом обустављању јавног превоза [9]. Ванредно стање је трајало 53 дана, тј. до 6. маја, а током ванредног стања спроведен је низ мера различитог степена озбиљности.

Циљ ове студије био је да се анализира појава опасних ситуација у саобраћају у Србији током 2020. године, са

an individual's capacity to engage in employment, resulting in financial instability [4]. The expenses associated with medical treatment, rehabilitation, and continuous care can be substantial, creating financial strain for individuals and their families. Chen et al. discovered that the combined cost of fatal and non-fatal injuries worldwide amounted to approximately 1.8 trillion USD over 15 years [5]. This accounted for approximately 0.12% of the global gross domestic product [5]. In certain nations, this proportion may reach as high as 3% of domestic product [1].

In 2020, the World Health Organization officially designated the novel coronavirus outbreak as a Public Health Emergency of International Concern [1, 6]. The WHO declaration called upon the international community to unite and devise effective strategies to halt the transmission of the novel coronavirus. As a reaction, authorities enforced a range of anti-epidemic measures, such as maintaining physical distance, isolating individuals, issuing orders to stay at home, shutting down schools, imposing restrictions on travel and large gatherings, temporarily halting non-essential services, and implementing different levels of lockdown [1]. As a result of the imposed restrictions, traffic patterns have unavoidably been altered, leading to a reduction in fuel consumption and a decrease in air pollution in urban areas [7, 8].

The initial documented instance of COVID-19 was officially recorded in Serbia on March 6th, 2020. On March 15th, a state of emergency was announced, resulting in the closure of all educational institutions, including schools, kindergartens, and universities. After three days, the Ministry of Interior implemented a nationwide curfew targeting individuals aged 65 and older, restricting them from leaving their residences. On March 20th, all restaurants, pubs, malls, and public transport closed. The Government of the Republic of Serbia made decisions on March 21st regarding the prohibition of entry into Serbia and the temporary halt of public transportation. [9]. The emergency persisted for 53 days until May 6th, during which a range of measures were implemented with differing levels of severity.

The objective of this study was to analyze the occurrence of RTIs in Serbia throughout the year 2020, with a specific focus on the period of lockdown [9]. The study seeks to determine whether there were any discrepancies in the frequency of injuries and the attributes of those injured during the lockdown. Based on the available data and responses to specific questions, it sought to evaluate the impact of emerging social circumstances on road traffic safety in 2020.

посебним фокусом на период затварања [9]. Студија настоји да утврди да ли је било разлика у учесталости повреда и карактеристика повређених током изолације. На основу доступних података и одговора на конкретна питања, истраживање је настојало да процени утицај новонасталих друштвених околности на безбедност саобраћаја на путевима 2020. године.

Методe

Подаци о дневно пријављеним случајевима заражених и смртним случајевима који се приписују новом корона вирусу и вредности индекса снаге преузети су са сајта *ourworldindata.com*. Ова интернет страница приказује податке које је прикупио Универзитет Џонс Хопкинс, а који су обрађени од стране *Our World in Data*. Извори података укључују СЗО COVID-19 *Dashboard*, као и студије које су спровели Донг и др. и Хале и др. [10, 11]. Индекс снаге је композитна мера од девет метрика, укључујући затварање јавног превоза, захтеве за останак код куће, интерна ограничења кретања и контролу путовања [12]. Виши скор на скали од 0 до 100 указује на строжу реакцију владе.

На сајту Агенције за безбедност саобраћаја Републике Србије дати су подаци о учесталости незгода и карактеристикама повређених лица од 1. јануара 2015. до 31. децембра 2020. године. Повреде су класификоване као смртоносне и несмртоносне. Смртни исходи су класификовани као смрт на лицу места, смрт током транспорта до здравствене установе и смрт у року од 30 дана након удеса.

Карактеристике које се узимају у обзир укључују: пол повређеног, старост (категорисана као млађи од 18, 18–64, или 65 или старији), њихове индивидуалне карактеристике (да ли су били пешаци, путници или возачи), начин превоза (аутобус, бицикл, мотоцикл, путничко возило, камион или друго превозно средство), место незгоде (да ли се догодила унутар или ван градског подручја), тежину повреде, као и конкретан дан и време када се незгода догодила. Такође смо истраживали конкретну локацију на којој се смртни случај догодио у случајевима смртоносних повреда.

Анализирали смо процентуалну промену броја повреда по месецима 2020. године у односу на референтни период према формули која следи у наставку текста. Референтни период је представљао исти месец 2019. године или петогодишњи просек за исти месец. Петогодишњи период обухватао је период од 2015. до 2019. године. За повреде у априлу израчуната је апсолутна разлика између броја повреда по сату у 2020. години у

Methods

The data regarding the daily reported infected cases and deaths attributed to a novel coronavirus and the Strength Index values were acquired from *ourworldindata.com*. This website presents data compiled by Johns Hopkins University, which Our World has processed in Data. The data sources include the WHO COVID-19 Dashboard and studies conducted by Dong et al. and Hale et al. [10, 11]. The Strength Index is a composite measure of nine metrics, including closures of public transport, stay-at-home requirements, internal movement restrictions, and travel control [12]. A higher 0–100 scale score suggests a more stringent governmental response.

The National Agency for Traffic Safety website provided data on accident frequency and the attributes of the injured individuals from January 1st, 2015, to December 31st, 2020. The injuries were classified as fatal and non-fatal. Fatal outcomes were classified as death on the spot, death in transit to a medical facility, and death within 30 days following the incident.

The attributes that are taken into account include the gender of the injured, their age (categorized as below 18, 18–64, or 65 or older), their individual characteristics (whether they were a pedestrian, passenger, or driver), the mode of transportation (bus, bicycle, motorcycle, passenger vehicle, truck, or other means), the location of the accident (whether it occurred inside or outside of an urban area), the severity of the injury, and the specific day and time when the accident took place. We also examined the particular site where the fatality took place in instances of lethal injuries.

We analyzed the percentage change in the number of injuries by month in 2020 compared to the reference period according to the formula given below. The reference period was the same month of 2019 or the five-year average for the same month. The five-year period covered the period from 2015 to 2019. For injuries in April, the absolute difference between the number of injuries per hour in 2020 compared to the same hours during April 2019 was calculated.

$$\text{Percentage of change} = \frac{\text{Number of injured in 2020}}{\text{Number of injured in reference period}} - 1$$

We used the Chi-square test and Fisher's exact test to examine the statistical disparities between attributes of the injured ones in April 2020 compared to April 2019. A p-value less than 0.05 is considered statistically significant.

Data was collected in Microsoft Excel, where all charts

поређењу са истим сатима током априла 2019. године.

$$\text{Процентуална промена} = \frac{\text{Број повређених 2020. године}}{\text{Број повређених у референтном периоду}} - 1$$

Користили смо χ^2 тест и Фишеров егзактни тест да бисмо испитали статистичке разлике између карактеристика повређених у априлу 2020. у односу на април 2019. године. Р-вредност мања од 0,05 сматра се статистички значајном.

Подаци су прикупљани у *Microsoft Excel*-у, где су направљени сви графикони. Статистички тестови су спроведени коришћењем IBM SPSS 21.

Резултати

У Србији је током 2020. године било 323.367 потврђених случајева заразне болести COVID-19 и 2983 смртних случајева изазваних инфекцијом COVID-19. Графикон 1 приказује обрасце кумулативно потврђених случајева COVID-19 (плава линија), кумулативних смртних случајева (наранџаста линија) и Индекс строности, који мери строност мера политике које су примењене током периода затварања (сива линија). Први случај COVID-19 у Републици Србији откривен је 6. марта 2020. године. Ванредно стање је проглашено 15. марта 2020. године, а места за јавна окупљања која нису од суштинског значаја су затворена 20. марта 2020. Индекс строности на слици 1 приказује образац у којем се прво јавља пораст у марту 2020. године, након чега следи опадање у мају када су мере биле ублажене. Два додатна периода владиних рестрикција уведена су у каснијим интервалима, при чему се први догодио у јулу 2020, а други у зиму 2020. године.

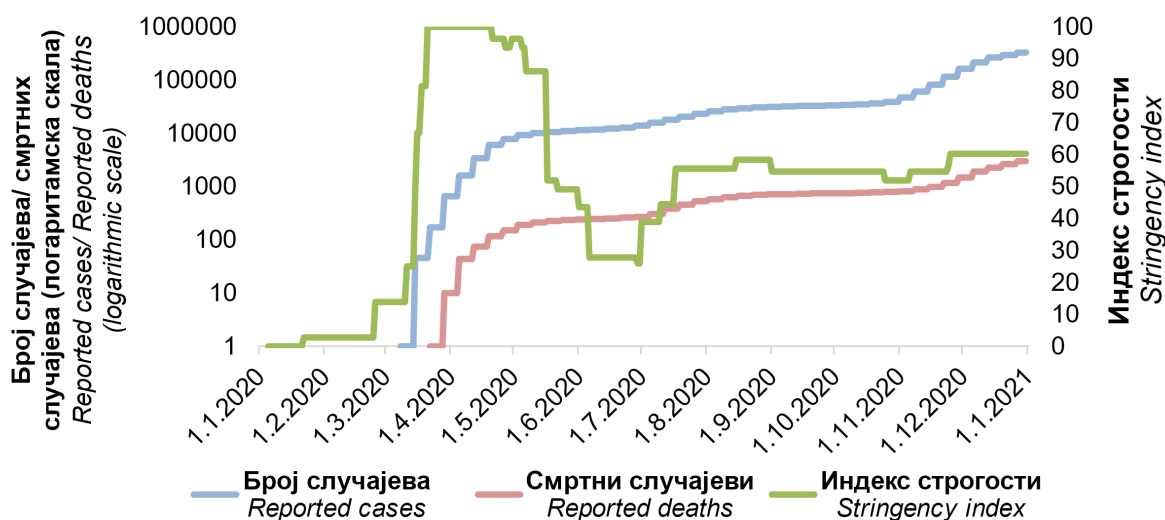
were created. Statistical tests were conducted using IBM SPSS 21.

Results

There were 323,367 confirmed cases of COVID-19 infection and 2,983 COVID-19 deaths in Serbia during 2020. Chart 1 illustrates the patterns of cumulative confirmed COVID-19 cases (the blue line), cumulative deaths (the orange line), and the Stringency Index, which gauges the strictness of the policy measures implemented during the lockdown period (the grey line). The initial COVID-19 case in the Republic of Serbia was identified on March 6th, 2020. A state of emergency was declared on March 15th, 2020, and non-essential and public gathering places were closed on March 20th, 2020. The Stringency index in Chart 1 reflects a pattern where there is an initial increase in March 2020, followed by a decrease in May when the measures were eased. Two additional periods of government-imposed restrictions were implemented at later intervals, with the first occurring in July 2020 and the second in winter 2020.

Графикон 1. Број случајева COVID-19, смртних случајева изазваних инфекцијом COVID-19 и Индекс строности у Србији током 2020. године

Chart 1. Number of COVID-19 cases, COVID-19 deaths, and Stringency index in Serbia during 2020

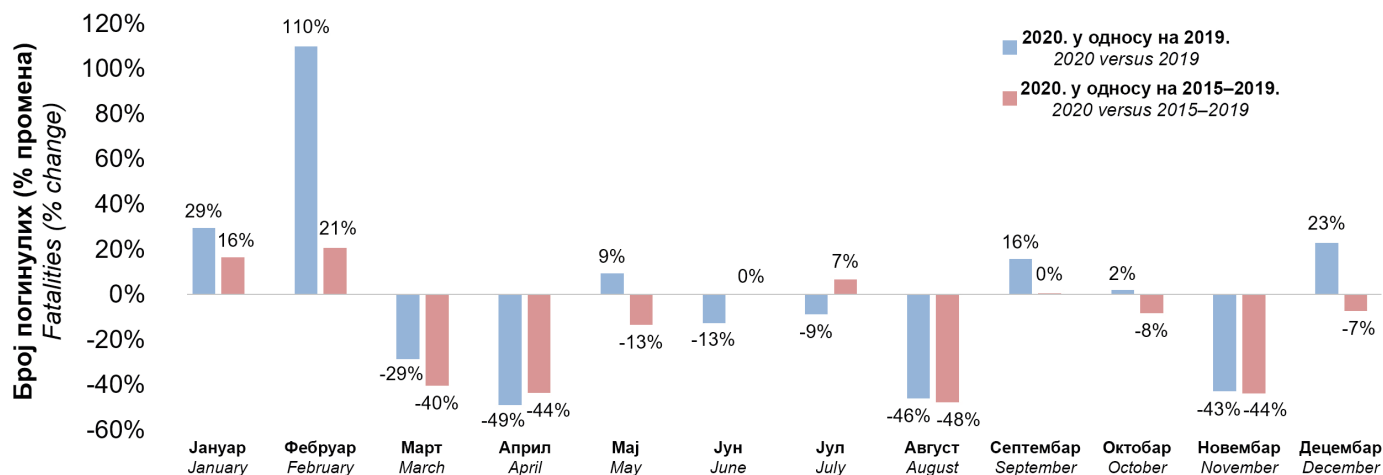


Поређења између 2020. године и референтних периода

Током 2020. године у саобраћајним незгодама у Републици Србији смртно су страдале 492 особе, што представља смањење од 8% у односу на 2019. годину када су погинуле 534 особе. У поређењу са петогодишњим просеком, ово је пад од 14%. У истој години, 17.246 лица је повређено без смртог исхода, што је 16% мање у односу на 20.415 повређених 2019. године. Овај проценат је такође рекордан када је у питању смањење у поређењу са петогодишњим просеком (допунски материјал, табела Д1 и табела Д2).

Графикон 2 приказује процентуалну промену броја погинулих у 2020. години у односу на 2019. и претходни петогодишњи просек, категорисано по месецима. Позитивна вредност промене указује на пораст броја повреда 2020. године у поређењу са истим месецом у референтним периодима. С друге стране, негативна вредност указује на смањење броја повреда. Током прва два месеца 2020. године забележен је већи број смртних случајева у односу на исти период 2019. године. Током марта и априла дошло је до значајног смањења броја смртних случајева од 29% и 49%, респективно. Промене у истом правцу уочене су и приликом поређења података са петогодишњим просеком. Након тога, дошло је до повећања од 9% у мају након смањења првобитних ограничења. У августу и новембру забележено је значајно смањење броја смртних случајева у поређењу са истим месецима 2019. године, са смањењем од 46% и 43%, респективно. Током септембра, октобра и децембра дошло је до пораста опасних ситуација у саобраћају са смртним исходом. У поређењу са петогодишњим просеком, забележено је смањење незгода са смртним исходом у свим месецима након марта, осим у јулу и септембру.

Графикон 2. Процентуална промена у месечном броју погинулих у саобраћајним несрећама током 2020. у односу на 2019. годину и петогодишњи просек



Comparisons between 2020 and reference periods

During 2020, 492 people were fatally injured in traffic accidents in the Republic of Serbia, which is a decrease of 8% compared to 2019, when 534 people died. Compared to the five-year average, this is a decrease of 14%. In the same year, 17,246 people were non-fatally injured, which is 16% less than the 20,415 people injured in 2019. This percentage is also a record when it comes to a decrease compared to the five-year average (Supplementary material, Table S1 and Table S2).

Chart 2 illustrates the percentage change regarding the number of fatally injured in 2020 compared to 2019 and the previous five-year average, categorized by month. A positive change value indicates an increase in the number of injuries in 2020 compared to the same month in reference periods. On the other hand, a negative value indicates a decrease in the number of injuries. During the initial two months of 2020, there was a greater number of deaths compared to the same period in 2019. During March and April, there was a substantial decrease of 29% and 49% in fatalities, respectively. Changes in the same direction were also observed when comparing the data with the five-year average. Subsequently, there was a 9% increase in May after the initial restrictions reduction. In August and November, there were significant reductions in fatalities compared to the same months in 2019, with a decrease of 46% and 43%, respectively. There was a rise in fatal RTIs during September, October, and December. Compared to the five-year average, there was a decrease in fatal accidents in all months after March except July and September.

Chart 2. Percentage change in the monthly fatalities in road traffic accidents during 2020 compared to 2019 and the five-year average

Графикон 3 приказује месечну процентуалну промену повреда без смртног исхода у 2020. у поређењу са 2019. годином и петогодишњим просеком. Слично као код смртоносних повреда, дошло је до пораста у прва два месеца 2020. године, након чега је уследио пад од марта надаље у поређењу са оба референтна периода. Највећи пад догодио се у априлу, марту и новембру.

Графикон 3. Процентуална промена у месечном броју повређених у саобраћајним несрећама током 2020. у односу на 2019. годину и петогодишњи просек

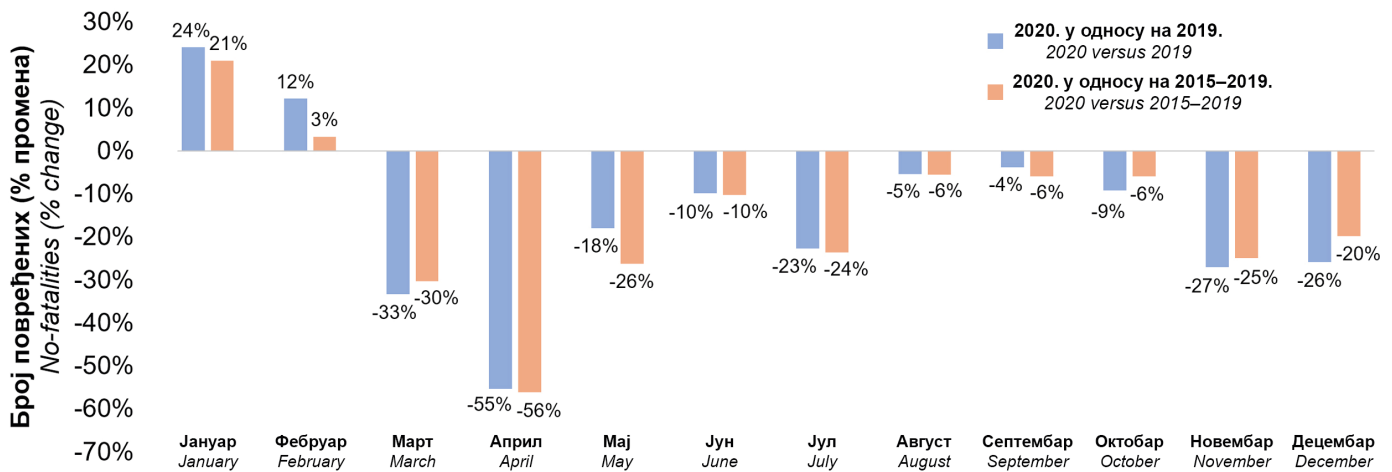


Chart 3 illustrates the monthly percentage change of non-fatal injuries in 2020 compared to 2019 and the five-year average. Similar to fatal injuries, there was a rise in the first two months of 2020, followed by a decline from March onward compared to both reference periods. The largest drop occurred in April, March, and November.

Chart 3. Percentage change in the monthly non-fatalities in road traffic accidents during 2020 compared to 2019 and the five-year average

Поређење између априла 2019. и априла 2020. године

Графикон 4 приказује апсолутну промену броја погинулих у априлу 2020. године у поређењу са априлом 2019. године, по сатима. Позитивна вредност промене указује на пораст броја погинулих током 2020. године, а негативна вредност указује на смањење броја погинулих. Највећи број смртних случајева у 2019. години забележен је током јутарњег шпица у 6.00 часова, поподневног шпица у 15.00 часова и 20.00 часова. Током периода затварања, дошло је до значајног смањења броја смртних случајева у 6.00 и 20.00 часова. Ипак, овај пад је био релативно мање изражен током трајања поподневног шпица. Једини сати у којима је забележен пораст броја погинулих били су 5.00, 13.00 и 17.00 часова. Допунски материјал садржи детаљне податке о броју повреда по сату изражене у апсолутним бројевима (допунски материјал, табела Д3).

Број повреда без смртног исхода по сату био је мањи 2020. године него у 2019. години, осим у 6.00 часова и посебно у 16.00 часова (графикон 4). Штавише, дошло је до значајног пада броја повређених између 20.00 часова и 4.59, са 357 лица 2019. на 12 особа 2020. године. Битно је напоменути да је затварање обично почињало у 20.00 сати, а завршавало се у 5.00 часова следећег јутра.

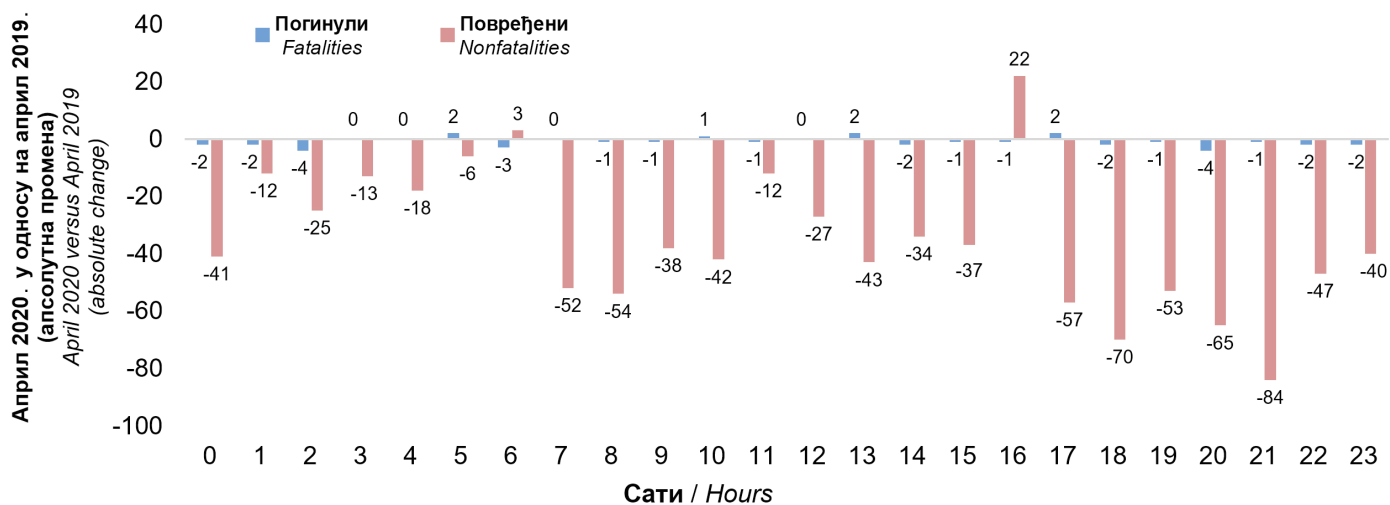
Comparison between April 2019 and April 2020

Chart 4 illustrates the absolute change of the fatally injured in April 2020 compared to April 2019 by hours. A positive change value indicates an increase in the number of fatalities during 2020, and a negative value suggests a decrease in the number of fatalities. The peak number of deaths in 2019 was recorded during the morning rush hour at 6 a.m., the afternoon rush hour at 3 p.m., and 8 p.m. Amidst the lockdown, there was a notable reduction in fatalities at 6 a.m. and 8 p.m. Nevertheless, the decline was comparatively less pronounced during the afternoon peak period. The only hours in which an increase in the number of deaths was recorded were 5 a.m., 1 p.m., and 5 p.m. The Supplement contains detailed data on the number of injuries per hour expressed in absolute numbers (Supplementary material, Table S3).

In 2020, the number of non-fatal injuries per hour was lower than in 2019, except at 6 a.m. and especially at 4 p.m. (Chart 4). Moreover, there was a substantial decline in the number of injured between 8 p.m. and 4:59 a.m., from 357 in 2019 to 12 in 2020. It is essential to mention that the lockdown typically commenced at 8 p.m. and concluded at 5 a.m. the following morning.

Графикон 4. Апсолутна промена броја погинулих и повређених у саобраћајним незгодама током априла 2020. у поређењу са априлом 2019. године, по сатима

Chart 4. Absolute change in the number of fatalities and non-fatalities in road traffic accidents during April 2020 compared to April 2019 by hours.



У табели 1 упоредили смо карактеристике погинулих особа у априлу 2019. и 2020. године. Број смртних случајева је смањен за скоро 50%. Две значајне разлике су откривене након примене Фишеровог теста у погледу превозног средства и времена незгоде. Број смртних случајева је смањен код свих видова транспорта осим у последњој категорији остала превозна средства, као што су пољопривредна возила. Ниједан смртни случај није повезан са коришћењем аутобуса или камиона, а само један смртни случај је забележен међу корисницима бицикла. Најмањи процентуални пад броја смртних случајева је забележен међу корисницима мотоцикла, са падом од само једне трећине. Што се тиче доба дана када је дошло до опасне ситуације у саобраћају, највеће смањење је било у периоду од 20.00 часова до поноћи, а нарочито од поноћи до 4.00.

Table 1 compares the attributes of the fatally injured individuals in April 2019 and 2020. The fatality count has decreased by nearly 50%. Two notable disparities were identified following the implementation of the Fisher test regarding the means of transport and time of the accident. Fatalities decreased in all modes of transport except in the last category of other modes, such as agricultural vehicles. No deaths were associated with bus or truck use, only one among bicycle users. The smallest percentage drop in fatalities was among motorcycle users, down by just one-third. Regarding the time of day when the RTI occurred, the largest decrease was from 8 pm to midnight, especially from midnight to 4 am.

Табела 1. Карактеристике погинулих у друмском саобраћају током априла 2019. у поређењу са априлом 2020. године

Table 1. Characteristics of the fatally injured in road traffic during April 2019 compared to April 2020

	Април 2019. <i>April 2019</i>		Април 2020. <i>April 2020</i>		p-вредност <i>p-value</i>
	n	%	n	%	
Старост / Age					0,649†
0–17	3	6%	0	0%	
18–64	34	72%	18	75%	
65+	10	21%	6	25%	
Пол / Sex					0,739†
Мушки / Male	39	83%	21	88%	
Женски / Female	8	17%	3	13%	
Карактеристике / Characteristics					0,181†
Пешак / Pedestrian	9	19%	4	17%	
Путник / Passenger	12	26%	2	8%	
Возач / Driver	26	55%	18	75%	
Превозно средство / Means of transport					0,026†
Аутобус / Bus	4	9%	0	0%	
Бицикл / Bicycle	6	13%	1	4%	
Мотор/мопед / Motor/moped	6	13%	4	17%	
Путничко возило (аутомобил) / Passenger vehicle (car)	28	60%	13	54%	
Камион / Truck	2	4%	0	0%	
Друго превозно средство / Other	1	2%	6	25%	
У градском подручју / In urban area					0,862‡
Да / Yes	21	45%	12	50%	
Не / No	26	55%	12	50%	
Место смрти / Location of death					0,924†
На лицу места / On the spot	30	64%	14	58%	
Током транспорта у здравствену установу <i>During transportation to a health facility</i>	4	9%	2	8%	
У року од 30 дана након незгоде / Within 30 days of the accident	13	28%	8	33%	
Дан незгоде / Day of the accident					0,258†
Понедељак–петак / Monday–Friday	32	68%	20	83%	
Викенд / Weekend	15	32%	4	17%	
Време незгоде (00–24) / Time of the accident (00–24)					0,049†
08–19	19	40%	14	58%	
20–23	12	26%	3	13%	
00–03	8	17%	0	0%	
04–07	8	17%	7	29%	
Укупно / Total	47	100%	24	100%	

†Фишеров тест

‡ χ^2 тест

†Fisher test

‡Chi square

У табели 2 представљени су број и стопа насталих повреда у априлу 2020. и априлу 2019. године, заједно са бројем повређених особа који је смањен за преко 50% и проценат промене броја погинулих. Статистички тестови су открили статистички значајне разлике у годинама, карактеристикама повређених (пешаци, путници или возачи), дану и времену незгоде. Пад броја повре-

Table 2 presents the count and occurrence rate of non-fatally injured in April 2020 and April 2019, with the number of individuals injured decreasing by over 50% and the percentage change for fatalities. Statistical tests revealed statistically significant differences in age, characteristics of the injured (pedestrians, passengers, or drivers), and day and time of the accident. The decline in the number

ђених посебно је изражен код најстаријих, а у погледу времену незгоде викендом и током ноћних и раних јутарњих часова.

of injured was particularly pronounced among the oldest, and in terms of the time of the accident on weekends and during the night and early morning hours.

Табела 2. Карактеристике повређених у друмском саобраћају током априла 2019. у поређењу са априлом 2020. године

Table 2. Characteristics of non-fatally injured in road traffic during April 2019 compared to April 2020

	Април 2019. April 2019		Април 2020. April 2020		p-вредност p-value
	n	%	n	%	
Старост / Age					< 0,001‡
0–17	173	11%	41	6%	
18–64	1187	78%	615	90%	
65+	168	11%	27	4%	
Пол / Sex					0,070‡
Мушки / Male	936	61%	446	65%	
Женски / Female	592	39%	237	35%	
Карактеристике / Characteristics					< 0,001‡
Пешак / Pedestrian	190	12%	47	7%	
Путник / Passenger	516	34%	184	27%	
Возач / Driver	822	54%	452	66%	
Превозно средство / Mean of transport					0,142†
Аутобус / Bus	63	4%	0	0%	
Бицикл / Bicycle	129	8%	56	8%	
Мотор/мопед / Motor/moped	106	7%	78	11%	
Путничко возило (аутомобил) / Passenger vehicle (car)	1119	73%	492	72%	
Камион / Truck	78	5%	40	6%	
Друго превозно средство / Other	33	2%	17	2%	
У градском подручју / In urban area					0,174†
Да / Yes	1147	75%	494	72%	
Не / No	381	25%	189	28%	
Тежина повреде / Severity of injury					0,135‡
Лакша / Light	1279	84%	554	81%	
Тешка / Severe	249	16%	129	19%	
Дан незгоде / Day of the accident					0,001‡
Понедељак–петак / Monday–Friday	1148	75%	648	95%	
Викенд / Weekend	380	25%	35	5%	
Време незгоде (00–24) / Time of the accident (00–24)					0,001‡
08–19	1053	69%	608	89%	
20–23	245	16%	9	1%	
00–03	94	6%	3	0%	
04–07	136	9%	63	9%	
Укупно / Total	1528	100%	683	100%	

†Фишеров тест

†Fisher test

‡ χ^2 тест

‡Chi square

Дискусија

Discussion

Ова студија указује да су повреде и смртни случајеви у саобраћају у Србији смањени 2020. у односу на 2019. годину. У априлу је забележен приметан пад броја смрт-

This study indicates that traffic-related injuries and fatalities in Serbia decreased in 2020 compared to 2019. In April, there was a notable decrease in deaths, with a de-

них случајева, уз смањење од 49%, као и смањење броја повреда без смртног исхода за 55% у односу на одговарајући месец претходне године. Сличне промене су забележене и у другим земљама у окружењу [13, 14]. У периоду од априла 2019. до априла 2020. године Хрватска је забележила пад од 39% у броју смртних случајева изазваних опасним ситуацијама у саобраћају, Мађарска је забележила пад од 43%, Словенија пад од 42%, а Румунија је забележила пад од 46% [13, 14]. Смањење броја смртних случајева у априлу примећено је као светски феномен, како наводе Катраказас и др, Номура и др, и Алхајасен и др. [13, 15, 16].

Анализом података који обухватају целу 2020. годину, евидентно је да је Србија забележила знатно мањи пад стопе смртности него остале земље. Смањење је износило свега 7,9% у односу на претходну годину, што је знатно ниже од смањења која су уочена у Албанији (20%), Бугарској (26%), Хрватској (20%), Мађарској (25%), Словенији (22%) и Румунији (12%) [14, 17]. Мањи пад броја смртних случајева имале су Босна и Херцеговина и Северна Македонија, са смањењем од 7% и 5%, респективно, у односу на Србију. С друге стране, у Црној Гори је забележен пораст броја погинулих за 2% [14, 17]. На глобалном нивоу, већина земаља је 2020. године доживела пад броја смртних случајева изазваних опасним ситуацијама у саобраћају. Ипак, поједине земље су одступиле од овог тренда [17].

Противепидемијске мере, укључујући затварање, довеле су до смањене мобилности и промена у обрасцу коришћења саобраћаја. Вожња ноћу представља фактор ризика за смртне случајеве изазване у саобраћају [18]. Дакле, ограничавање учешћа у саобраћају током ноћних сати морало је неизбежно утицати на смањење броја смртних случајева. Ипак, број удеса без смртног исхода у 16.00 часова премашио је број из априла 2019. То се десило зато што су људи желели да заврше своје обавезе пре затварања. Претходно обавештење о мерама изолације од стране власти је неопходно како би се ублажила анксиозност грађана. Доживљавање стреса може довести до паничне куповине, упуштања у опасна понашања и доношења исхитрених одлука у вези са путовањем, што може индиректно утицати на безбедност на путевима [19].

Противепидемијске интервенције, као што су затварање у вечерњим сатима и ноћу, као и ограничења јавног превоза, могу објаснити статистички значајне разлике у броју смртних случајева током априла 2020. године [20]. Јавни превоз у Србији престао је са радом 22. марта [20]. Иако су људи били приморани да се пребаце на друге видове превоза, посебно активни

cline of 49%, and a reduction in non-fatal injuries by 55% compared to the corresponding month in the previous year. Similar changes were noted in other surrounding countries [13, 14]. In the period from April 2020 to April 2019, Croatia experienced a decline of 39% in the number of deaths caused by RTIs, Hungary reported a decrease of 43%, Slovenia experienced a drop of 42%, and Romania reported a decline of 46% [13, 14]. The decline in the number of deaths in April was observed as a worldwide phenomenon, as reported by Katrakazas et al., Nomura et al., and Alhajaseen et al. [13, 15, 16].

Upon analyzing the data encompassing the entirety of 2020, it is evident that Serbia experienced a markedly lesser decline in mortality rates than other nations. The decrease amounted to a mere 7.9% in comparison to the preceding year, which is significantly lower than the reductions observed in Albania (20%), Bulgaria (26%), Croatia (20%), Hungary (25%), Slovenia (22%), and Romania (12%) [14, 17]. Bosnia and Herzegovina and North Macedonia had a lesser decline in the number of deaths, with a decrease of 7% and 5%, respectively, compared to Serbia. On the other hand, Montenegro witnessed a 2% increase in the number of deaths [14, 17]. On the global scale, most countries experienced a decline in the number of deaths caused by RTIs in 2020. Nevertheless, certain countries have deviated from this trend [17].

Anti-epidemic measures, including closure, resulted in reduced mobility and changes in the pattern of traffic use. Driving at night is a risk factor for traffic fatalities [18]. So, restricting traffic participation during night hours must have inevitably impacted the decrease in deaths. Nevertheless, the number of non-fatal incidents at 4 p.m. exceeded that of April 2019. This occurred because individuals desired to finish their tasks before the closure. Prior notification of closure measures by authorities is necessary to mitigate citizen anxiety. Experiencing stress could result in panic buying, engaging in dangerous behaviors, and making hasty travel choices, which can indirectly impact road safety [19].

Anti-epidemic interventions, such as evening and nighttime closures, as well as restrictions on public transport, may explain the statistically significant differences in deaths during April 2020 [20]. Public transport in Serbia stopped operating on March 22 [20]. Even though people were forced to switch to other forms of transport, especially active transport, we also documented a large decrease in the number of cyclist deaths [21]. One possible explanation could be the restriction on leaving home for people over 65. It is known that almost half of cyclist deaths in Europe are in people from this age group [22]. The search for alternatives to public transport and restrictions on car

транспорт, такође смо документовали велики пад броја смртних случајева међу бициклистима [21]. Једно од могућих објашњења могло би да буде ограничење кретања за особе старије од 65 година. Познато је да се готово половина броја смртних случајева бициклиста у Европи односи на особе из ове старосне групе [22]. Тражење алтернатива јавном превозу и ограничења у коришћењу аутомобила у вези са максималним дозвољеним бројем људи у возилу повећала је употребу других превозних средстава и број настрадалих током коришћења тих превозних средстава. Упитно је колико су ова алтернативна превозна средства била технички исправна и колико су њихови корисници били обучени и вешти при управљању.

Занимало нас је да ли је смањење броја повреда на путевима последица побољшања безбедности или смањења обима саобраћаја током затварања. У том смислу смо консултовали литературу. Подаци из многих земаља указују на пад безбедности на путевима током 2020. године [23]. Један од могућих разлога је неопрезна вожња. Катраказас и др. открили су да је смањење обима саобраћаја у Грчкој и Саудијској Арабији довело до повећане појаве возача који прекорачују ограничење брзине и користе мобилне телефоне током вожње [13]. Слични резултати примећени су у Бразилу, Индији и САД, како наводе Мицхелараки и др, Донг и др, Тито Переира Собреира и др, и Арун Патхак и др. [10, 24–26]. Ванлар и др. су интервјуисали возаче у Америци и Канади и открили да су амерички возачи показали знатно већу склоност да учествују у ризичном понашању у саобраћају током пандемије [27]. Недавна студија о канадским возачима открила је да је пандемија довела до смањења агресивног понашања у вожњи међу испитаницима. Међутим, испитаници из ове студије приметили су пораст агресивног понашања код других возача на путу [28]. Пандемија је можда довела до погоршања стања безбедности саобраћаја услед смањене будности надзора од стране саобраћајне полиције [16]. На крају, не треба занемарити утицаје вируса на оболеле особе и здраве преносиоце. COVID-19 има потенцијал да изазове озбиљно стање познато као синдром акутног респираторног дистреса, што може довести до когнитивних и емоционалних оштећења која могу потрајати неколико месеци након инфекције [29]. Ово су само нека од могућих објашњења пораста броја погинулих током маја 2020. године, чиме је дошло до првог приметног ублажавања рестриктивних мера.

Национални извештај о стању безбедности на путевима Србије указује да је пад броја погинулих и повређених 2020. године био краткотрајан, барем у погледу повреда са смртним исходом [30]. У 2021. години број

use related to the maximum number of people allowed in a vehicle increased the use of other means of transport and the number of casualties during their use. It is questionable how technically sound these alternative means of transport were and how skilled their users were.

We were interested in whether the reduction in road injuries was a consequence of improved safety or reduced traffic volumes during the lockdown. In this regard, we consulted the literature. Data from many countries suggest a decline in road safety during 2020 [23]. One potential rationale is reckless driving. Katrakazas et al. discovered that the reduction in traffic volume in Greece and Saudi Arabia led to an increased occurrence of drivers exceeding speed limits and using mobile phones while driving [13]. Similar results were observed in Brazil, India, and the USA, as reported by Michelaraki et al., Dong et al., Tito Pereira Sobreira et al., and Arun Pathak et al. [10, 24–26]. Vanlaar et al. conducted interviews with drivers in America and Canada and discovered that American drivers had a notably greater inclination to participate in hazardous traffic behavior during the pandemic [27]. A recent study on Canadian drivers revealed that the pandemic led to a decrease in aggressive driving behavior among respondents. However, respondents from this study observed a rise in aggressive behavior from other drivers on the road [28]. The pandemic may have led to a deterioration in traffic safety due to reduced vigilance from traffic enforcement units [16]. Lastly, the effects of the virus on the sick person and the healthy carrier should not be ignored. COVID-19 virus has the potential to result in a serious condition known as acute respiratory distress syndrome, which can lead to cognitive and emotional impairments that may persist for several months following infection [29]. These are just some of the possible explanations for the increase in the number of fatally injured people during May 2020, as soon as the first noticeable easing of restrictive measures occurred.

The Serbian National Road Safety Report suggests that the decline in fatalities and injuries in 2020 was short-lived, at least in terms of fatal injuries [30]. In 2021, fatalities increased compared to 2020. What's more, the number of deaths in 2022 was above the five-year average [30]. This slight growth existed at the global level, also [31]. On the other hand, the number of non-fatal injuries increased in 2021 compared to the previous year but did not reach the value of 2019 [30]. In 2022, the number of non-fatalities was slightly lower than the five-year average [30].

Finally, it is vital to acknowledge the potential constraints of this research. One group of limitations relates to the quality of data collected by many people from across the country. It is entirely expected that errors might occur in the collec-

смртних случајева се повећао у односу на 2020. Штавише, број смртних случајева 2022. био је изнад петогодишњег просека [30]. Овај благи раст је постојао и на глобалном нивоу [31]. С друге стране, број повреда без смртног исхода је повећан 2021. године у односу на претходну годину, али није достигао вредност из 2019. године [30]. Број повређених је 2022. године био нешто мањи од петогодишњег просека [30].

Конечно, од виталног је значаја признати потенцијална ограничења овог истраживања. Једна група ограничења односи се на квалитет података које прикупљају многи људи из целе земље. Сасвим је очекивано да дође до грешака у прикупљању података о карактеристикама повређених. Друга група ограничења односи се на повезаност између пандемије и учесталости опасних ситуација у саобраћају. Уочени трендови у вези са опасним ситуацијама у саобраћају можда нису повезани са појавом COVID-19, али су свакако последица флукуација у учесталости опасних ситуација у саобраћају.

Закључак

Студија је спроведена у циљу анализе појаве опасних ситуација у саобраћају (енгл. RTI) у Републици Србији током периода затварања услед пандемије COVID-19. Резултати анализе су показали да је 2020. године дошло до укупног смањења укупног броја повређених и погинулих у односу на претходну годину. Најзначајније одступање у учесталости повреда примећено је у априлу, посебно током најригорознијих мера затварања.

Изјава о финансијској добити

Аутори нису добили никаква посебна средства за овај рад.

Доступност података

Подаци ће на разуман захтев бити стављени на располагање од стране аутора за кореспонденцију.

tion of attributes of the injured. The second group of limitations relates to the connection between the pandemic and the frequency of RTIs. The observed trends in RTIs may have no connection with the emergence of the COVID-19 virus but are a consequence of fluctuations in the frequency of RTIs.

Conclusion

A study was conducted to analyze the occurrence of RTIs in the Republic of Serbia during the COVID-19 lockdown period. The findings of the analysis showed that in 2020, there was an overall decrease in the total number of both non-fatally and fatally injured compared to the previous year. The most significant discrepancy in injury incidence was observed in April, particularly during the most rigorous lockdown measures.

Financial Disclosure Statement

The authors received no specific funding for this work.

Data availability

Data will be made available by the corresponding author upon reasonable request.

Допуна

Supplement

Табела Д1. Број погинулих по месецима у периоду од 2015. до 2020. године

Table S1. Number of fatalities by month from 2015 to 2020

Месец Month	#2015	#2016	#2017	#2018	#2019	Свеукупно Grand Total	петогодишњи просек 5-year average	#2020
01_Јануар / January	50	35	32	38	34	189	38	44
02_Фебруар / February	39	38	34	43	20	174	35	42
03_Март / March	33	34	40	33	28	168	34	20
04_Април / April	36	51	49	30	47	213	43	24
05_Мај / May	39	51	42	38	32	202	40	35
06_Јун / June	46	50	47	42	55	240	48	48
07_Јул / July	56	68	57	42	68	291	58	62
08_Август / August	72	63	61	67	63	326	65	34
09_Септембар / September	71	47	63	62	51	294	59	59
10_Октобар / October	58	59	64	56	52	289	58	53
11_Новембар / November	53	59	42	46	49	249	50	28
12_Децембар / December	46	52	48	51	35	232	46	43
Укупно / Total	599	607	579	548	534	2867	573	492

Табела Д2. Број повређених по месецима у периоду од 2015. до 2020. године

Table S2. Number of non fatalities by month from 2015 to 2020

Месец Month	#2015	#2016	#2017	#2018	#2019	Свеукупно Grand Total	петогодишњи просек 5-year average	#2020
01_Јануар / January	1201	1343	1188	1587	1289	6608	1322	1599
02_Фебруар / February	1097	1186	1356	1439	1147	6225	1245	1286
03_Март / March	1269	1385	1601	1493	1519	7267	1453	1012
04_Април / April	1384	1679	1602	1604	1528	7797	1559	683
05_Мај / May	1686	1915	1834	1870	1603	8908	1782	1315
06_Јун / June	1832	1906	2034	1673	1853	9298	1860	1670
07_Јул / July	1921	1946	2228	1933	1981	10009	2002	1530
08_Август / August	1875	1866	2017	1964	1928	9650	1930	1823
09_Септембар / September	1877	1977	1879	1928	1862	9523	1905	1791
10_Октобар / October	1762	2036	1907	1894	1987	9586	1917	1803
11_Новембар / November	1736	1655	1869	1716	1807	8783	1757	1318
12_Децембар / December	1710	1772	1750	1688	1911	8831	1766	1416
Укупно / Total	19350	20666	21265	20789	20415	102485	20497	17246

Табела Д3. Број погинулих и повређених по сатима у априлу 2019. и априлу 2020. године

Table S3. Number of fatalities and non fatalities by the hour in April 2019 and April 2020

Сати / Hours	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Погинули #2019 Fatalities	2	2	4	0	1	0	5	2	1	1	0	3	2	0	2	5	2	0	2	1	5	3	2	2
Погинули #2020 Fatalities	0	0	0	0	1	2	2	2	0	0	1	2	2	2	0	4	1	2	0	0	1	2	0	0
Повређени #2019 Non fatalities	43	12	26	13	18	15	33	70	64	67	82	81	88	111	117	103	85	116	82	57	71	85	49	40
Повређени #2020 Non fatalities	2	0	1	0	0	9	36	18	10	29	40	69	61	68	83	66	107	59	12	4	6	1	2	0

Литература / References

1. World Health Organization. Global status report on road safety 2023 [Internet]. Geneva: World Health Organization; 2023. Available from: <https://iris.who.int/bitstream/handle/10665/375016/9789240086517-eng.pdf>
2. The Lancet. Road safety: more than reducing injuries. *The Lancet*. 2022; 400(10346):73. [https://doi.org/10.1016/S0140-6736\(22\)01209-0](https://doi.org/10.1016/S0140-6736(22)01209-0)
3. Rolison JJ, Regev S, Moutari S, Feeney A. What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records. *Accid Anal Prev*. 2018; 115:11–24. <https://doi.org/10.1016/j.aap.2018.02.025>
4. Gopalakrishnan S. A public health perspective of road traffic accidents. *J Family Med Prim Care*. 2012; 1(2):144–50. <https://doi.org/10.4103/2249-4863.104987>
5. Chen S, Kuhn M, Prettnner K, Bloom DE. The global macroeconomic burden of road injuries: estimates and projections for 166 countries. *Lancet Planet Health*. 2019; 3(9):e390–e8. [https://doi.org/10.1016/S2542-5196\(19\)30170-6](https://doi.org/10.1016/S2542-5196(19)30170-6)
6. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet*. 2020; 395(10223):470–3. [https://doi.org/10.1016/S0140-6736\(20\)30185-9](https://doi.org/10.1016/S0140-6736(20)30185-9)
7. Li J, Tartarini F. Changes in air quality during the COVID-19 lockdown in Singapore and associations with human mobility trends. *Aerosol and Air Quality Research*. 2020; 20(8):1748–58. <https://doi.org/10.4209/aaqr.2020.06.0303>
8. Tian X, An C, Chen Z, Tian Z. Assessing the impact of COVID-19 pandemic on urban transportation and air quality in Canada. *Sci Total Environ*. 2021; 765:144270. <https://doi.org/10.1016/j.scitotenv.2020.144270>
9. Uredba o merama za vreme vanrednog stanja [Regulation on measures during a state of emergency], „Službeni glasnik Republike Srbije”, broj 55/2020 [Official Gazzete of the Republic of Serbia, No. 55/2020].
10. Dong X, Xie K, Yang H. How did COVID-19 impact driving behaviors and crash Severity? A multigroup structural equation modeling. *Accid Anal Prev*. 2022; 172:106687. <https://doi.org/10.1016/j.aap.2022.106687>
11. Thomas H, Boby T, Angrist N, Cameron-Blake E, Hallas L, Kira B, et al. “Variation in Government Responses to COVID19” Version 9.0 [Internet]. Blavatnik School of Government Working Paper. 2020. Available: <https://www.bsg.ox.ac.uk/sites/default/files/2020-12/BSG-WP-2020-032-v10.pdf>
12. Roser M. What is the COVID-19 Stringency Index? [Internet]. Our World in Data; 2021. Available from: <https://ourworldindata.org/metrics-explained-covid19-stringency-index>
13. Katrakazas C, Michelaraki E, Sekadakis M, Yannis G. A descriptive analysis of the effect of the COVID-19 pandemic on driving behavior and road safety. *Transp Res Interdiscip Perspect*. 2020; 7:100186. <https://doi.org/10.1016/j.trip.2020.100186>
14. Yasin YJ, Grivna M, Abu-Zidan FM. Global impact of COVID-19 pandemic on road traffic collisions. *World J Emerg Surg*. 2021; 16(1):51. <https://doi.org/10.1186/s13017-021-00395-8>
15. Nomura S, Kawashima T, Yoneoka D, Tanoue Y, Eguchi A, Gilmour S, et al. Trends in deaths from road injuries during the COVID-19 pandemic in Japan, January to September 2020. *Inj Epidemiol*. 2021; 7(1):66. <https://doi.org/10.1186/s40621-020-00294-7>
16. Alhajyaseen WKM, Almuqdad A, Hussain Q, Almallah M, Al Malki MA, Singaravelu J, et al. Road safety status during COVID-19 pandemic: exploring public and road safety expert's opinions. *Int J Inj Contr Saf Promot*. 2022; 29(2):135–51. <https://doi.org/10.1080/17457300.2021.1962915>
17. Transport Community. Fatalities for 2020: Annual Statistics for Western Balkans. Belgrade: Transport Community; 2021.
18. Sapsirisavat V, Mahikul W. Drinking and Night-Time Driving May Increase the Risk of Severe Health Outcomes: A 5-Year Retrospective Study of Traffic Injuries among International Travelers at a University Hospital Emergency Center in Thailand. *Int J Environ Res Public Health*. 2021; 18(18). <https://doi.org/10.3390/ijerph18189823>

19. Adanu EK, Okafor S, Penmetsa P, Jones S. Understanding the Factors Associated with the Temporal Variability in Crash Severity before, during, and after the COVID-19 Shelter-in-Place Order. *Safety*. 2022; 8(2):42. <https://doi.org/10.3390/safety8020042>
20. Lazić N, Lazić V, Kolarić B. First three months of COVID-19 in Croatia, Slovenia, Serbia and Federation of Bosnia and Herzegovina—comparative assessment of disease control measures. *Hrvatska: Infektološki glasnik*. 2020; 40:43–9. <https://doi.org/10.37797/ig.40.2.1>
21. Glavić D, Trpković A, Milenković M, Jevremović S. Pandemic impact on traffic trends and patterns in the city of Belgrade. *Transport*. 2023; 38(3):165–77. <https://doi.org/10.3846/transport.2023.19375>
22. Evgenikos P, Yannis G, Folla K, Bauer R, Machata K, Brandstätter C. How Safe are Cyclists on European Roads? *Transportation Research Procedia*. 2016; 14:2372–81. <https://doi.org/10.1016/j.trpro.2016.05.269>
23. Rajput K, Sud A, Rees M, Rutka O. Epidemiology of trauma presentations to a major trauma centre in the North West of England during the COVID-19 level 4 lockdown. *Eur J Trauma Emerg Surg*. 2021; 47(3):631–6. <https://doi.org/10.1007/s00068-020-01507-w>
24. Michelaraki E, Sekadakis M, Katrakazas C, Ziakopoulos A, Yannis G. A four-country comparative overview of the impact of COVID-19 on traffic safety behavior. 10th International Congress on Transportation Research, Future Mobility and Resilient Transport: Transition to Innovation ICTR; 2021.
25. Tito Pereira Sobreira L, Luna M, Cunto F, Hellinga B. Impact of COVID-19 on Traffic Volume, Violations, and Crashes in Fortaleza, Brazil. *Journal of Transportation Engineering*. 2022; 148. <https://doi.org/10.1061/JTEPBS.0000738>
26. Arun Pathak A, Chandrasekaran S, Annamalai B. Analysis of motor vehicle accidents: Comparison between before and during the COVID-19 lockdown in Maharashtra, India. *Transportation research record*. 2023; 2677(4):503–16. <https://doi.org/10.1177/03611981221089936>
27. Vanlaar WGM, Woods-Fry H, Barrett H, Lyon C, Brown S, Wicklund C, et al. The impact of COVID-19 on road safety in Canada and the United States. *Accid Anal Prev*. 2021; 160:106324. <https://doi.org/10.1016/j.aap.2021.106324>
28. Lopetrone E, Biondi FN. On the Effect of COVID-19 on Drivers' Behavior: A Survey Study. *Transp Res Rec*. 2023; 2677(4):742–50. <https://doi.org/10.1177/03611981221103866>
29. do Carmo Filho A, van Duinkerken E, Tolentino JC, Schmidt SL. Attention profile of physically recovered COVID-19 inpatients on the day of discharge. *J Psychiatr Res*. 2022; 150:189–96. <https://doi.org/10.1016/j.jpsychires.2022.03.047>
30. Agencija za bezbednost saobraćaja. Statistički izveštaj o stanju bezbednosti saobraćaja u Republici Srbiji 2022 [Statistical report on the state of traffic safety in the Republic of Serbia 2022]. Beograd: Agencija za bezbednost saobraćaja. 2023.
31. World Health Organization. Global status report on road safety 2023 [Internet]. Geneva: World Health Organization; 2024. Available from: <https://www.transport-community.org/wp-content/uploads/2024/06/Global-status-report-on-road-safety-2023.pdf>



Примљено / Received

6. 12. 2024.

Ревидирано / Revised

13. 12. 2024.

Прихваћено / Accepted

13. 12. 2024.

Кореспонденција / Correspondence

Александар Медаревић – Aleksandar Medarević

aleksandar_medarevic@batut.org.rs

КАТЕГОРИЈА М60

ELEKTROMETRIJSKA ISPITIVANJA VODOIZVORIŠTA MAKIŠ

Filip Arnaut¹, Branislav Sretenović¹, Vesna Damnjanović¹, Vesna Cvetkov¹

¹*Univerzitet u Beogradu, Rudarsko-geološki fakultet, Dušina 7 Beograd, g804-21@rgf.bg.ac.rs*

Izvod

U radu su prikazane mogućnosti metode elektrometrijskog skeniranja, testirane u zoni beogradskog vodoizvorišta „Makiš“, sa ciljem procene njene pouzdanosti u monitoringu zagađenja vode.

Ključne reči: elektrometrijska metoda, zagađenje voda, dimetil-benzen

Uvod

Geofizičko elektrometrijsko skeniranje (ERT) je nedestruktivna metoda koja se zasniva na merenju specifične električne otpornosti [1]. Mogućnost brzog merenja, praktičnost, kao i ekonomska pogodnost, glavni su razlozi zašto se ERT primenjuje u različitim oblastima inženjerstva i nauke kao što su geotehnička i građevinska istraživanja, fundamentalna geološka istraživanja, hidrogeološka istraživanja [2], i dr.

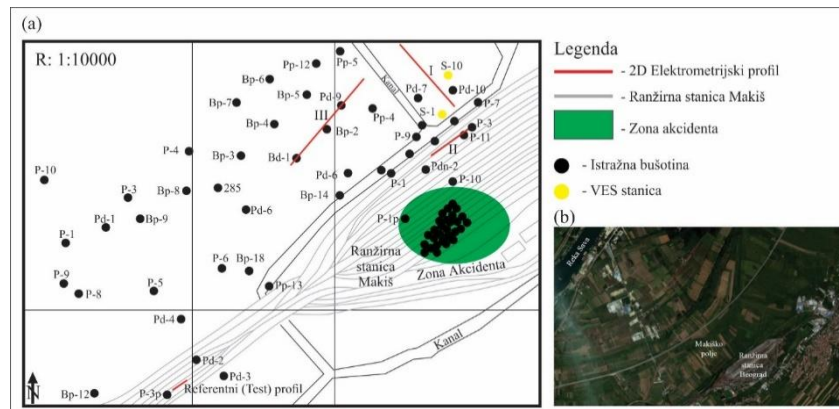
Akcident prolivanja dimetil-benzena koji je doveo je do ozbiljnog zagađenja podzemnih voda trebalo je najpre detektovati a zatim i osmatrati zbog moguće migracije zagađenja prema bunarima za snabdevanje pijaćom vodom, kao i prema glavnom hidrogeološkom akviferu „Makiš“. Akcident je iskorišćen da se testira ERT metoda. Za monitoring zagađenja vode izabrana je navedena metoda jer može da se bez ikakvih rizika primeni u zoni samog izvora zagađenja. Od još većeg značaja je to što metoda pruža podatke i o vertikalnom rasprostranjenju zagađenja ka većim dubinama. „Bušotine“ takođe daju tačne informacije o vertikalnom rasprostranjenju zagađenja, ali ne pružaju, kao ERT metoda, mogućnost dobijanja dovoljno „gustih“ podataka o lateralnom širenju zagađenja. Čak šta više, duboke bušotine u kontaminiranim zonama mogu da prouzrokuju širenje zagađenja naniže, ka akviferu. Osnovni zadatak ispitivanja u slučaju vodoizvorišta Makiš, sačinjenom od peskova i šljunkova na dubinama od 15 do 25 metara [3], bio je određivanje zaštitnih karakteristika pokrivača, kao i hidrogeoloških karakteristika akvifera, koji se nalazi u okviru rečne terase reke Save.

2D geofizička elektrometrijska merenja u zoni zagađenja u oblasti vodosnabdevanja i diskusija rezultata merenja

Sedimenti aluvijona reke Save izgrađeni su od peskova i peskovitih i prašinstih glina zbog čega su visokopropusni, što omogućava migriranje zagađenja naniže [4]. Prirodno tlo je od ranžirne stanice odvojeno drenažnim kanalom koji samo delimično sprečava širenje zagađenja u horizontalnom pravcu prema bunarima za vodosnabdevanje (Slika 1a).

Rezultati plitkog ispitivanja ERT metodom u zoni relativno homogenog peščanog pokrivača koji je nasut u zoni ranžirne stanice (profil II na Slici 2), upoređeni su sa rezultatima koji su dobijeni na nezagađenom test području (test profil na Slici 1). Lokaliteta su izabrani na osnovu rezultata prethodnih ispitivanja u plitkim bušotinama i na osnovu položaja izvora zagađenja u odnosu na regionalni tok podzemnih voda (test profil postavljen je na višoj koti za oko 2 m). Ispitivanja na test području izvedena su sa 32. elektrode koje su na međusobnom rastojanju od 2 m na 5 dubinskih nivoa. Poređena 2D interpretacija, duž test profila i profila koji su postavljeni u zoni zagađenja, jasno pokazuju mogućnost detekcije zagađenja, jer su vrednosti specifične električne otpornosti ρ bile značajno niže u zagađenoj zoni. Kako pokrovni veštački

sloj (nasip od peska i tucanika) nema značajnih lateralnih litoloških promena, ovakav pad vrednosti ρ nedvosmisleno je povezan sa promenom kvaliteta podzemnih voda.



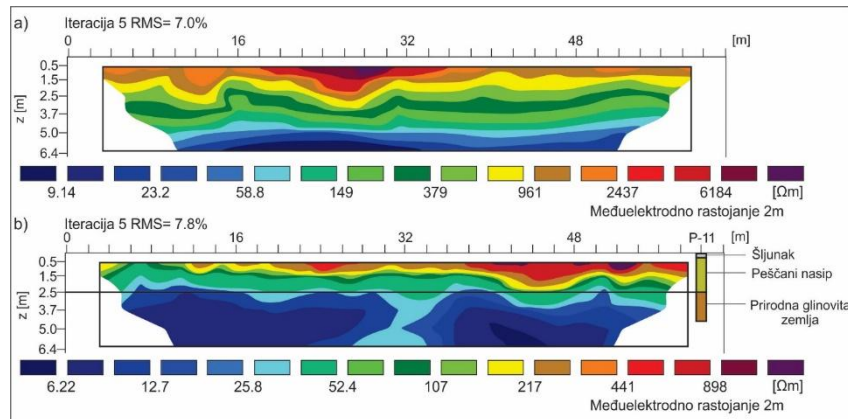
Slika 1. (a) Lokacija: istražnih bušotina, elektrometrijskih 2D profila, i VES stanica (vertikalno električno sondiranje); (b) Satelitski snimak Makiškog polja [5]

Rezultati ispitivanja duž nultog test profila predstavljeni su na slici 2a. Na osnovu poređenja merene i teorijske prividne ρ u 115 tačaka duž nultog test profila, utvrđeno je da srednje kvadratno odstupanje (RMS) iznosi 7%. Vrednost RMS odstupanja je zadovoljavajuća s obzirom na to da je merna nesigurnost 2-3%. Jasno uočljivo prisustvo visokootporne električne nehomogenosti u površinskoj zoni je najverovatnije posledica 3D karaktera. Pored nehomogenosti, nasuti materijal karakterišu i visoke vrednosti ($\rho \approx 10^3 \Omega\text{m}$) bliže površini terena, i nešto manje ($\rho \approx 400 \Omega\text{m}$) na dubini od 3 m. Raspodela ρ odražava prisustvo najpre (po dubini) krupnog tucanika bez zapune, zatim tucanika sa peskovitom zapunom, i konačno, peskovitog materijala. Debljina nasipa varira u rasponu od oko 2,5-3,5 m. Rezultati bušenja na obližnjoj bušotini P-3p potvrđuju da je nasip izgrađen od 0,3 m tucanika, potom od 0,3 m tucanika sa refuliranim peskom, i 1,9 m refuliranog peska, što daje ukupnu debljinu nasutog materijala od oko 2,5 m.

Na Slici 2- dole, predstavljeni su rezultati interpertacije duž prvog dela Profila II, na isti način kao i rezultati duž test profila. RMS je približno isto - 7,8%, međutim, kako u pogledu geometrije, tako i u pogledu raspodele vrednosti ρ , uočava se bitna razlika u izdvojenom 2D modelu poluprostora. Električno viskootpornom pokrivaču (nasuti pesk), odgovaraju generalno niže vrednosti ρ (maksimalno $1,1 \cdot 10^3 \Omega\text{m}$) i manje je debljine (0,5-2,5 m). Već na dubini od oko 2,5-3 m, vrednost ρ naglo opada i pokazuje izraženije lateralne promene u odnosu na test profil. Generalno, uz malo odstupanje, niske vrednosti ρ zadržavaju se i na dubinama od preko 6 m. Na osnovu podataka dobijenih „bušenjem“ u bušotinama P-11 i Pdn-2, koje se nalaze u zoni profila, debljina nasutog materijala kreće se od 1,9 m na P-11 (na 60-om m prvog dela profila II) do 2,3 m na Pdn-2 (na 40-om m od kraja Profila II). Nasip je na bušotini P-11 izgrađen od 0,1 m tucanika i 1,8 m refuliranog peska koji je delimično šljunkovit, dok na bušotini Pdn-2 tucanika nema a nasip je izgrađen od sitnog refuliranog peska sa malo šljunka. Poređenjem vrednosti ρ na prvom delu Profila II, sa odgovarajućim vrednostima ρ na test profilu, uočava se da su one, očigledno, na istim dubinskim nivoima ispitivanja višestruko niže, naročito na dubinama od 2,5-5 m.

S obzirom na veću debljinu tucanika i tucanika sa peskom na nultom test profilu, kao i na nešto veću debljinu nasipa u zoni nultog profila, razumljiv je viši nivo ρ na manjim dubinama. Međutim, prisustvo 0,6 m tucanika i tucanika sa refuliranim peskom u zoni test profila, ne može da prouzrokuje znatniju razliku u ρ na dubinskim nivoima koji odgovaraju refuliranom pesku (do ≈ 3 m) kao i u podinskim glinovitim i glinovito-peskovitim sedimentima (na

dubinama 3-5 m). Usvajajući pretpostavku o relativno sličnom sastavu nasipa od refuliranog peska i približno istom nivou podzemne vode, može da se pretpostavi da je na različite vrednosti ρ presudno uticala znatno snižena vrednost ρ slojne vode u zagađenoj zoni.



Slika 2. (a) Rezultati 2D inverzije duž referentnog test profila;
(b) Rezultati 2D inverzije duž profila u zoni zagađenja podzemnih voda

Ispitivanja ERT metodom izvedena su i u zoni van nasutog peščanog sloja duž profila I (dužine 300 m) u pravcu vodoizvorišta (Slika 3a). Primenjen je *Wenner*-ov raspored elektroda [6]. Minimalno rastojanje između elektroda iznosilo je 10 m a merenja na šest dubinskih nivoa omogućila su da se dosegne dubinski zahvat ispitivanja od oko 30 m, i da se detektuje električno visokootporni sloj (od $\rho \approx 70 \Omega\text{m}$), koji je verovatno izgrađen od šljunkovitog peska (Slika 3a- beli pravugaonik).

Da bi se bolje uočila građa ispitivanog poluprostora rezultati 2D interpretacije koji su prikazani na slici 3 dati su sa uvećanjem vertikalne ose (dubine z). Nižeotporni sloj raspona $\rho \approx 18\text{-}25 \Omega\text{m}$ (glinoviti pesak) oplicava u pravcu vodoizvorišta (od $\approx 13\text{-}7,5$ m), što ukazuje na postepenu lateralnu promenu snižavanjem vrednosti ρ na 110-om metru profila (Slika 3a, crveni pravugaonik). Ovo snižavanje vrednosti ρ odražava povećanje udela glinovite komponente u sedimentima što poboljšava zaštitna svojstva od prodiranja mogućeg zagađenja sa površine u dubljazaležuci akvifer. Smanjenje debljine ovog pokrovnog sloja, sa druge strane, smanjuje i vrednost uzdužne električne provodljivosti (S), koja jednoznačno ukazuje na zaštitna svojstva pokrivača bez uticaja višeznačnosti rezultata interpretacije (princip ekvivalencije). Činjenica da nisu deteknovane vrednosti $\rho < 18 \Omega\text{m}$, potvrđuje da nema prodora zagađenja iz zone ranžirne stanice u široj zoni, ni po dubini, ni prema bunarima vodoizvorišta.

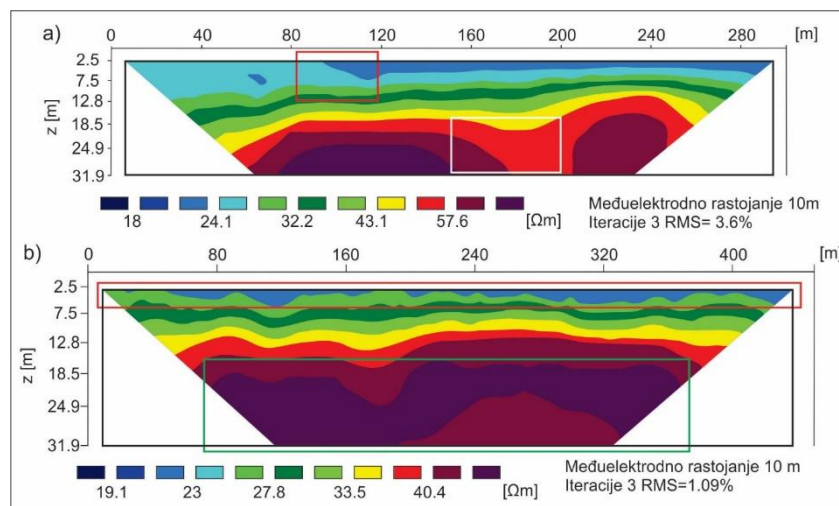
Primenom *Wenner*-ovog rasporeda elektroda sa manjom vrednošću minimalnog rastojanja elektroda ($a < 10$ m), ovaj sloj može da se ispita detaljnije i samim tim bolje definišu zaštitna svojstva pokrivača. U slučaju ovih ispitivanja, izborom $a = 10$ m napravljen je kompromis između detaljnosti izdvojenih sredina u pličoj zoni i maksimalno ostvarenog dubinskog zahvata ispitivanja u cilju definisanja građe dubljih slojeva. Sa povećavanjem dubine uočava se rast vrednosti ρ ($\approx 25\text{-}70 \Omega\text{m}$) što ukazuje na povećanje udela peskovite komponente, odnosno na prisustvo akvifera.

Dvodimenzionalni model (Slika 3a), takođe, otkriva postojanje izrazitih lateralnih litoloških promena na većim dubinama (u zoni između 150-og i 200-tog metra), što je važna informacija za karakterisanje akvifera, odnosno o bočnim promenama udela nižeotporne glinovite i višeotporne peskovite komponente.

Profil III, dužine 450 m (Slika 3b), postavljen je približno normalno na pravac Profila I, na udaljenosti od ranžirne stanice od oko 200-250 m, pri kraju Profila I. Primenjen je *Wenner*-ov dispozitiv sa minimalnim rastojanjem elektroda od 10 m, na šest dubinskih nivoa ispitivanja.

Električno nižeotporni pokrovni sloj (Slika 3b- crveni pravugaonik) sa vrednostima ρ u rasponu od 19-23 Ωm (glinoviti pesak) znatno je tanji u odnosu na Profil I (7,5 m i manje), a njegova zaštitna svojstva od mogućeg zagađenja sa površine terena su znatno manja. Naime, pri istoj vrednosti ρ_1 , zbog manje debljine pokrovnog sloja h_1 , umanjena je vrednost uzdužne električne provodljivosti ($S_1=h_1/\rho_1$). U ovom slučaju primena *Wenner*-ovog rasporeda elektroda sa manjom vrednošću minimalnog rastojanja elektroda ($a=2$ m, na minimum 10 nivoa ispitivanja) je nužna da bi pokrovni sloj bio ispitan detaljnije, tj. da bi zaštitna svojstva pokrivača bila bolje definisana. Uočava se da i duž Profila III nisu detektovane vrednosti $\rho < 19 \Omega\text{m}$, što potvrđuje da nema prodora zagađenja iz zone ranžirne stanice, ni po dubini, ni prema bunarima u široj zoni vodoizvorišta.

Dvodimenzionalni model ukazuje na postojanje kontinualnog električno višeotpornog peskovitog sloja ($\approx 50 \Omega\text{m}$) (Slika 3b, zeleni pravugaonik), koji je na Profilu I detektovan samo u zoni od 70-150 m na dubini 25-30 m. Pored toga, primetno je sniženje vrednosti ρ ispod ovog sloja, što otkriva povećan udeo glinovite komponente u okviru akvifera.



Slika 3. (a) Rezultat 2D inverzije podataka elektrometrijskog skeniranja duž Profila I; (b) Rezultat 2D inverzije podataka geoelektričnog skeniranja duž Profila III

Zaključak

Geofizička elektrometrijska ispitivanja 1D, 2D i 3D pružaju mogućnost brze, efikasne, i ekonomične procene zaštitnih svojstva nižeotpornog pokrivača, kao i eventualnih promena izazvanih mogućim zagađenjem u zoni akvifera u široj zoni vodoizvorišta. 1D elektrometrijsko sondiranje po mreži 50×50 m ukazuju na svojstva zaštitnog pokrivača i akvifera „pre i posle“ izgradnje i eksploatacije budućeg metroa i naselja. U zonama u kojima se na osnovu 1D metodologije otkrije da su osetljivije u pogledu mogućeg zagađenja sa površine terena neophodno je primeniti 2D ispitivanja po pravilnoj mreži 2D profila, pri čemu međusobno rastojanje profila ne bi trebalo da prevazilazi vrednost dvostrukog minimalnog rastojanja između elektroda [7]. Na taj način bi se obezbedila 2D inverzija duž profila, kao i 3D inverzija na osnovu većeg broja paralelnih profila skeniranja.

Navedena metodologija elektrometrijskih ispitivanja redukovala bi obim neophodnog bušenja i omogućila bi da se razreše eventualne neodumice o mogućem uzročniku nižeotpornih zona u okviru akvifera vodoizvorišta “Makiš” (npr. litološke promene ili promene kvaliteta podzemne

vode). Pored toga, ponovljenim ispitivanjima u određenim vremenskim intervalima moglo bi da se dođe do saznanja o eventualnom širenju zagađenja, kao i o promeni nivoa zagađenja.

Literatura

- [1] Bhattacharya, B.B., & Shalivahan, S. (2016). *Geoelectric methods: Theory and Application*. McGraw Hill Education (India).
- [2] Reynolds, J. M. (2011). *An introduction to applied and environmental geophysics*. John Wiley & Sons.
- [3] Tošović, S., Sarić, Z., Gburčik, V., Dopović, D., Matić-Besarabić, S., Tanasković, M., Majstorović, B., Stojanović, R., Marčetić, Lj., Jelača, P., Milutinović, M., Agramović, R., Mandić, M., Mladenović, S., & Pajić, D. (2002). Ekološka valorizacija područja generalnog plana Beograda- Ekološki atlas Beograda. *Direkcija za građevinsko zemljište i izgradnju Beograda. Gradski zavod za zaštitu zdravlja Beograda*.
- [4] Marković, B., Veselinović, M., Anđelković, J., Stevanović, P., Roglić, Č. & Obradinović, Z. (1984). Osnovna geološka karta Jugoslavije 1:100 000, Tumač za list Beograd (K34-113), *Savezni geološki zavod, Beograd*.
- [5] GeoSrbija. (2021). *Nacionalna infrastruktura geoprostornih podataka*, [Online]: <https://geosrbija.rs/>
- [6] Bhattacharya, P.K. & Patra, H.P. (1968). *Direct current geoelectric sounding: Principles and interpretation*. Elsevier publishing company.
- [7] Loke, M.H. (2004). Tutorial: 2-D and 3-D Electrical Imaging Surveys, *Geotomo Software, Res2dinv Software Manual*.



ISBN: 978-86-7031-604-1

PROGNOZIRANJE IZNADPROSEČNIH VREDNOSTI KVALITETA VAZDUHA U NOVOM SADU KORIŠĆENJEM RANDOM FOREST MODELA

Filip Arnaut, Vesna Cvetkov, Dragana Đurić

Univerzitet u Beogradu, Rudarsko-geološki fakultet, Đušina, 7 11000 Beograd

filip.arnaut@rgf.rs

Izvod

Kvalitet vazduha je značajan parametar kvaliteta života i životne sredine celokupno, pa je praćenje emisija suspendovanih čestica u vazduhu od izuzetnog značaja za očuvanje životne sredine i unapređenje kvaliteta života. U ovom radu se istražuje primena mašinskog učenja za prognoziranje iznad ili ispodprosečnih vrednosti zagađujućih materija u vazduhu, sa fokusom na koncentraciju PM_{2.5} čestica. Korišćeni Random Forest model se pokazao kao adekvatno rešenje za problem klasifikacije, sa tačnošću i preciznošću od 77%, čak i bez optimizacije modela. Analiza pokazuje da parametar SO₂ ima zanemarljivu korelaciju sa ostalim parametrima kvaliteta vazduha (PM_{2.5}, PM₁₀ i NO₂), što omogućava izostavljanje ovog parametra u narednim modelima i smanjenje utrošenog računarskog vremena.

Ključne reči: Klasifikacija, Mašinsko učenje, Statistička analiza kvaliteta vazduha, Novi Sad.

Uvod

Povišene vrednosti zagađujućih materija u vazduhu predstavlja sve veći problem u urbanim sredinama širom sveta, a posebno u gradovima Srbije poput Beograda i Valjeva koji se često nalaze na listi najzagađenijih gradova na svetu (Air Visual). Potrebno je napomenuti da platforma Air Visual za rangiranje gradova po kvalitetu vazduha koristi podatke sa zvaničnih državnih mernih stanica. Ovo je posledica zavisnosti Republike Srbije od fosilnih goriva za proizvodnju energije, neregulisane industrije i povećanja broja motornih vozila. Kvalitet vazduha se najčešće meri kroz nivo suspendovanih čestica koje mogu negativno uticati na zdravlje ljudi, posebno na respiratorni sistem i kardiovaskularne funkcije. To su zagađujuće materije suspendovane u vazduhu u formi čvrstih čestica ili tečnih kapljica koje su značajan polutant u vazduhu sa snažnim uticajem na ljudsko zdravlje. Stoga, bilo kakvi pokušaji razumevanja uzročnika zagađujućih materija u vazduhu, kao i mogućnost njihovog prognoziranja mogu se smatrati značajnim po šire društvo.

Za prognoziranje koncentracije zagađujućih materija u vazduhu sve više se koriste metode prognoziranja vremenskih serija i mašinskog učenja. Primer toga je Prophet model [1-7], kao i određene hibridne metode [8].

U ovom radu je predstavljen pristup koji kombinuje metode mašinskog učenja i podatke o kvalitetu vazduha u Republici Srbiji kako bi se predvideo nivo zagađujućih materija u vazduhu u budućnosti. Za ovaj pristup korišćena je klasifikaciona metoda, a ne regresiona. Rezultati dobijeni ovim pristupom mogu predstavljati kvalitetnu uvodnu studiju za buduća opsežnija istraživanja prognoziranja kvaliteta vazduha u Srbiji. Ova studija se razlikuje od drugih pristupa koji koriste metode prognoziranja vremenskih serija ili hibridne metode. Primena mašinskog učenja na podatke o kvalitetu vazduha može biti korisna za razumevanje uzroka povišenih vrednosti zagađujućih materija u vazduhu i za dalje istraživanje prognoziranja koncentracija zagađujućih materija u vazduhu u Republici Srbiji.

Metodologija i podaci

Deskriptivna statistika predstavlja standardnu proceduru prilikom statističkih istraživanja. Parametri koji spadaju pod deskriptivnu statistiku pomažu prilikom „upoznavanja sa podacima“. Deskriptivnu statistiku mogu označavati vrednosti minimuma i maksimuma podataka, kao i modaliteta, medijane i prosečne vrednosti. Mere koje kvantifikuju disperziju podataka, takođe, pripadaju koracima deskriptivne statistike, kao što su koeficijent varijacije, standardna devijacija i varijansa. Konstrukcija raspodela (histograma) podataka i prikazivanje koeficijenta asimetrije i koeficijenta spljoštenosti spadaju takođe pod deskriptivnu statistiku. U radu sa vremenskim serijama potrebno je poznavati ukupan broj podataka i broj nedostajućih (preskočenih) opservacija da bi se odredio kompletan set podataka i odabrala adekvatna metoda imputacije ili amputacije podataka.

Koeficijenti korelacije predstavljaju kvantitativnu meru korelacije između dve grupe podataka. Pirsonov koeficijent korelacije (r) meri intenzitet i smer linearnog odnosa između dve grupe podataka, dok Spirmanov koeficijent korelacije (r_s) kvantitativno ocenjuje intenzitet i smer korelacije između dve grupe podataka bez pretpostavki o raspodeli ili vrsti povezanosti. Opseg u kom se Pirsonov koeficijent korelacije nalazi je od -1 (negativna korelacija) do 1 (pozitivna korelacija). Vrednost od 0 kod Pirsonovog koeficijenta korelacije predstavlja ne postojanje korelacije između dve grupe podataka. Pretpostavka kod Pirsonovog koeficijenta korelacije je da su podaci u linearnoj zavisnosti i da su obe grupe podataka u normalnoj raspodeli.

Spirmanov koeficijent korelacije (r_s) kvantitativno ocenjuje intenzitet i smer korelacije između dve grupe podataka. Slično Pirsonovom koeficijentu korelacije, opseg u kome se prikazuje korelacija se nalazi od -1 do 1, gde vrednost od 0 se smatra nepostojanjem korelacije. Prilikom istraživanja, korišćena su oba koeficijenta korelacije sa tim da je veći akcenat stavljen na Spirmanov koeficijent korelacije u slučaju ne-normalne raspodele podataka.

Random Forest model prvi put je prikazan 2001. godine [9], i od tada predstavlja jedan od najprimenljivijih metoda mašinskog učenja. Random Forest model svoju primenu je našao u klasifikacionim kao i u regresorskim problemima. Random Forest model zasniva se na kombinaciji većeg broja stabala time da se međusobnim glasanjem stabala dolazi do jedinstvenog rešenja koje je predstavljeno određenom klasom [9-10]. Kao hiperparametar ili kao regularizacioni parametar se može posmatrati broj stabala, sa tim da su veće vrednosti uglavnom bolje ukoliko postoji adekvatno računarsko vreme na raspolaganju [11].

Podaci za ovo istraživanje preuzeti su od strane Agencije za zaštitu životne sredine Republike Srbije. Baza podataka za 2021. godinu sadrži merene parametre kvaliteta vazduha kao što su SO_2 , NO_2 , $PM_{2.5}$, PM_{10} , O_3 , CO i druge za sve merne stanice u Republici Srbiji. Podaci za mernu stanicu Novi Sad- Rumenačka su preuzeti iz baze podataka i korišćeni za ovo istraživanje.

Da bi podaci bili adekvatni za klasifikacioni problem mašinskog učenja, merene vrednosti PM_{10} , $PM_{2.5}$, SO_2 i NO_2 su prebačene u vrednosti klasa (*TRUE* ili *FALSE*) ukoliko je merena vrednost za dati sat veća od prosečne vrednosti tog parametra. Parametar koji je bio prognozirani je koncentracija $PM_{2.5}$, dok su PM_{10} , SO_2 i NO_2 korišćeni kao parametri za klasifikaciju $PM_{2.5}$ vrednosti. Time se omogućila primena Random Forest klasifikacionog modela, odnosno, omogućuje se klasifikacija budućih $PM_{2.5}$ vrednosti na osnovu iznad ili ispodprosečnih vrednosti PM_{10} , NO_2 i SO_2 . Takođe, da bi podaci bili prikladni za Random Forest model, podeljeni su na tri grupe: podaci za treniranje modela „*Train*“, podaci za validaciju modela „*Validation*“ i podaci za testiranje modela, odnosno „*Test*“. Podela su izvršene prema 64%- 16%- 20% odnosu. Test grupa podataka korišćena je za verifikaciju i statističku evaluaciju Random Forest modela.

Rezultati i diskusija

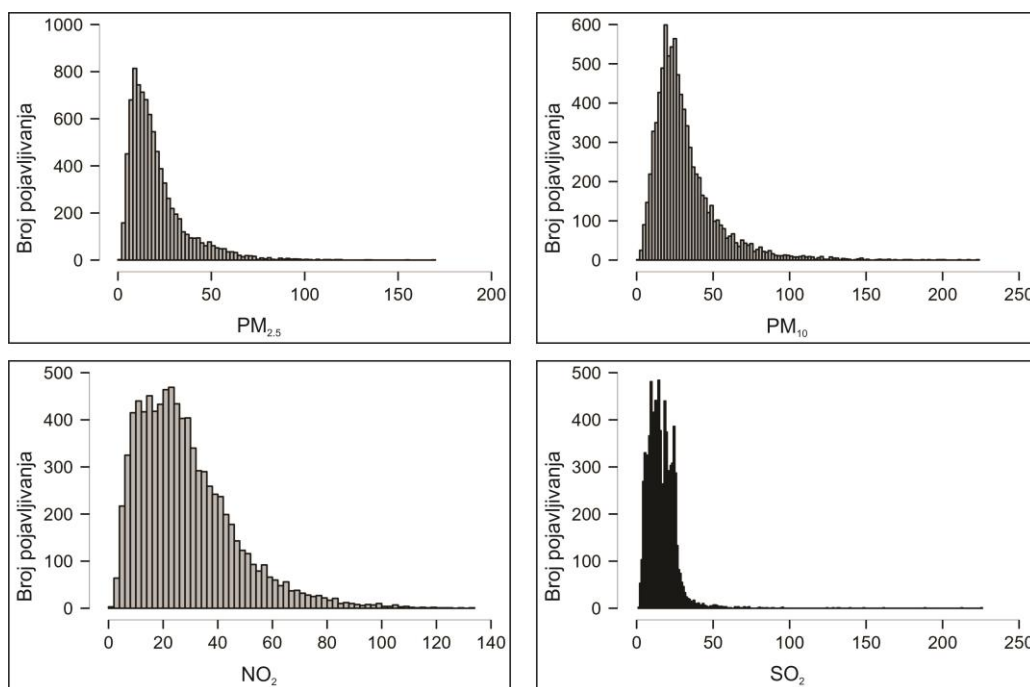
Statistička analiza podataka

U tabeli 1 prikazani su parametri deskriptivne statistike sračunate za koncentracije PM₁₀, PM_{2.5}, SO₂ i NO₂ u vazduhu merenoj na mernoj stanici Rumenička u Novom Sadu u 2021. godini. Ukupan broj podataka koji može biti meren u godini (u slučaju da godina nije prestupna) iznosi 8760 podataka, pod pretpostavkom da su podaci mereni sa rezolucijom od 1h, kontinualno. Procenat nedostajućih podataka sa merne stanice Novi Sad- Rumenička nalazi se u opsegu od 1.4% (PM_{2.5} i PM₁₀) do 1.59% (NO₂), što se za analizu primenom metoda mašinskog učenja u ovom istraživanju može smatrati kontinualnim zapisom podataka. Takođe, maksimalan broj uzastopno prekinutih opservacija se nalazi u redu veličine od 100 podataka (sati), zbog toga se, set podataka može smatrati uslovno kontinualnim za svrhe ovog istraživanja, a metode imputacije i amputacije podataka neće biti primenjene.

Tabela 1. Deskriptivna statistika parametara kvaliteta vazduha.

	PM _{2.5} [µg/m ³]	SO ₂ [µg/m ³]	NO ₂ [µg/m ³]	PM ₁₀ [µg/m ³]
Broj validnih podataka	8639	8625	8623	8639
Broj nedostajućih podataka	121	135	137	121
Procenat nedostajućih podataka	1.40	1.57	1.59	1.40
Modalitet	10.4	14.5	10.5	24.7
Medijana	16.3	14.8	24.9	26.1
Prosečna vrednost	20.487	16.287	28.489	31.688
Standardna devijacija	15.304	10.064	17.883	21.566
Koeficijent asimetrije	2.153	5.211	1.321	2.397
Koeficijent spljoštenosti	7.332	71.721	2.486	9.356
Minimum	1.81	1.52	1.68	1.83
Maksimum	170	226	134	223

Podaci iz tabele 1 ukazuju na desno-asimetričnu raspodelu svih parametara kvaliteta vazduha, gde je vrednost modaliteta (najčešće vrednosti) manja od vrednosti medijane podataka, koja je pak manja od prosečne vrednosti. Koeficijenti asimetrije za sve četiri koncentracije ukazuju na desno-asimetričnu raspodelu, kao i izrazito velike vrednosti maksimuma u odnosu na prosečnu vrednost, što ukazuje na prisustvo ekstremnih vrednosti.



Slika 1. Raspodela parametara kvaliteta vazduha na mernoj stanici Novi Sad- Rumenička za 2021. godinu

Bitan parametar pri većini statističkih istraživanja predstavlja koeficijent korelacije, odnosno mera korelacije između podataka. U tabeli 2 su dati izračunati koeficijenti korelacije primenom Pirsonovog i Spirmanovog koeficijenta korelacije. Sa obzirom da ni jedan prethodno prikazan parametar kvaliteta vazduha nije predstavljen normalnom raspodelom, težinski će se više uzimati značajnost Spirmanovog koeficijenta korelacije. Parametar SO₂ prikazuje zanemarljive koeficijente korelacije sa NO₂ i PM₁₀, dok je sa PM_{2.5} prikazana slaba korelacija. Sa druge strane, najveću (jaku) korelaciju prikazuju parametri PM₁₀ i PM_{2.5}, dok su NO₂ i PM₁₀ i PM_{2.5} predstavljeni srednjom korelacijom.

Tabela 2. Koeficijenti korelacije parametara kvaliteta vazduha.

		PM _{2.5} [µg/m ³]	SO ₂ [µg/m ³]	NO ₂ [µg/m ³]
SO ₂ [µg/m ³]	Pirson (r)	-0.084*		
	Spirman (r _s)	-0.283**		
NO ₂ [µg/m ³]	Pirson (r)	0.518***	0.018*	
	Spirman (r _s)	0.46***	-0.012*	
PM ₁₀ [µg/m ³]	Pirson (r)	0.834****	0.078*	0.64***
	Spirman (r _s)	0.793****	-0.019*	0.59***

*Zanemarljivo; **Slabo; ***Srednje; ****Jako; *****Vrlo jako
 opsezi koeficijenata korelacije dati prema [12]; p-vrednosti su manje od 0.001

Random Forest klasifikacija

Random Forest klasifikacija je urađena sa podelom podataka tako da podaci za treniranje modela uzimaju ukupno 5606 sati iz godine (oko 64% od ukupnog seta podataka), podaci za validaciju modela uzimaju 1402 sata (16% celog seta podataka), dok podaci za testiranje uzimaju preostalih 1752 sati, odnosno 20% celog seta podataka. Random Forest model je imao ukupno 73 stabla.

Tabela 3 prikazuje matricu konfuzije koja je sačinjena nakon treniranja i validacije modela sa podacima za testiranje. Iz tabele 3 se može videti da model u 47% slučajeva prognozira da će nova merena vrednost $PM_{2.5}$ koncentracije biti manja od srednje vrednosti $PM_{2.5}$ koncentracije za godinu. Tačnu prognozu model daje u 35% slučajeva dok u ostalih 12% istinita vrednost $PM_{2.5}$ koncentracije je veća od prosečne vrednosti. Sa druge strane, model u 53% slučajeva prognozira da će nova merena vrednost $PM_{2.5}$ koncentracije biti veća od prosečne vrednosti. Tačnu prognozu model pravi u 42% slučajeva, dok u preostalih 11% pravi pogrešnu prognozu. Ukupno, u 77% slučajeva model pravi tačnu prognozu, a u 23% slučajeva pogrešnu prema podacima iz Tabele 3.

Tabela 3. Matrica konfuzije za Random Forest klasifikacioni model.

		Prognozirano			
		FALSE [%]	FALSE	TRUE [%]	TRUE
Istinito	FALSE	35	615	11	187
	TRUE	12	209	42	741

Da bi se dobila kompletnija slika o sposobnostima modela prikazani su dodatni parametri evaluacije Random Forest modela (Tabela 4). Ukupna tačnost modela (0.774) prikazuje da je model relativno adekvatan, ali sa prostorom za njegovo unapređenje i poboljšanje tačnosti. Parametar preciznosti za „TRUE“ klasu uzima veće vrednosti nego za „FALSE“ klasu. Drugim rečima, kada model napravi prognozu da novi mereni podatak $PM_{2.5}$ koncentracije neće biti veći od prosečne vrednosti, ima preciznost od 74.6%, dok kada model napravi prognozu da će novi mereni podatak biti veći od prosečne vrednosti njegova preciznost će biti veća sa 79.8%.

Vrednosti F1 mere za obe klase su relativno bliske, što govori o tome da model ima balansiran kompromis između preciznosti i odziva za obe klase. Stopa lažno pozitivnih rezultata prikazuje da model netačno prognozira iznadprosečne vrednosti kvaliteta vazduha u 23.3% vremena kada je stvarna vrednost ispod prosečne, i obrnuto (stopa lažno negativnih rezultata). Statistički paritet za model prikazuje malu pristranost ka „TRUE“ klasi, odnosno da model ima malo veću verovatnoću prognoziranja novih merenih vrednosti kao iznadprosečnih.

Tabela 4. Evaluacija Random Forest klasifikacionog modela.

	FALSE	TRUE	Prosek/Ukupno
Broj podataka	802	950	1752
F1 mera	0.756	0.789	0.774
Tačnost	0.774	0.774	0.774
Preciznost	0.746	0.798	0.775
Odziv	0.767	0.78	0.774
Stopa lažno pozitivnih rezultata	0.22	0.233	0.227
Stopa lažno negativnih rezultata	0.233	0.22	0.227
Statistički paritet	0.47	0.53	1

Bitan parametar nakon treniranja i evaluacije modela predstavlja parametar prosečnog smanjenja tačnosti koji u ovom slučaju prikazuje da na vrednosti prognoze iznadprosečnih ili ispodprosečnih vrednosti $PM_{2.5}$ koncentracije najviše utiče parametar koncentracije PM_{10} , nakon čega parametar koncentracije NO_2 i na kraju SO_2 .

Drugim rečima, na kvalitet dobijene prognoze najviše doprinose vrednosti koncentracije PM_{10} , a najmanje vrednosti koncentracije SO_2 . U velikoj meri se parametar prosečnog smanjenja tačnosti poklapa sa koeficijentom korelacije dobijenim u tabeli 2, gde je prikazano da parametar koncentracije SO_2 ima zanemarljive do slabe korelacije sa ostalim parametrima kvaliteta vazduha prema Spirmanovom koeficijentu korelacije. Sa druge strane, najveću korelaciju parametar $PM_{2.5}$ ima upravo sa parametrom PM_{10} , što u najvećoj meri objašnjava i veliki uticaj parametra PM_{10} na prosečno smanjenje tačnosti. U budućim istraživanjima i primeni Random Forest klasifikacionog modela, moguće je izbaciti parameter SO_2 iz analize zbog njegovog zanemarljivog doprinosa tačnosti. Time bi se, pri analizi velikog broja podataka (na primer istraživanje sa podacima od više godina i parametara), smanjilo računarsko vreme koje je potrebno za analizu.

Zaključak

Kvalitet vazduha predstavlja bitan pokazatelj stanja životne sredine, naročito u urbanim sredinama. Povišene vrednosti zagađujućih materija u vazduhu mogu imati nepovoljne uticaje na respiratorne i kardiovaskularne bolesti kod stanovništva. Poznavanje budućih vrednosti koncentracija zagađujućih materija u vazduhu može biti značajna informacija za ugrožene grupe (decu, stariju populaciju i osobe sa hroničnim bolestima) kao i za celokupno stanovništvo.

U radu je prikazan alternativni postupak primene mašinskog učenja za prognoziranje budućih vrednosti kvaliteta vazduha. Problemu klasifikacije iznadprosečnih vrednosti parametara $PM_{2.5}$ pristupljeno je korišćenjem parametara PM_{10} , SO_2 i NO_2 . Rezultati ukazuju da i modeli koji nisu u potpunosti optimizovani mogu dati relativno kvalitetne rezultate. Prikazan Random Forest model ostavlja prostor za dalju optimizaciju. Takođe je utvrđeno da, je moguće i korisno primeniti dodatna istraživanja sa optimizovanim modelima, većim brojem podataka i parametara radi evaluacije prognoze budućih vrednosti kvaliteta vazduha.

Literatura

- [1] Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72, 37-45.
- [2] Zhao, N., Liu, Y., Vanos, J. K., & Cao, G. (2018). Day-of-week and seasonal patterns of $PM_{2.5}$ concentrations over the United States: Time-series analyses using the Prophet procedure. *Atmospheric environment*, 192, 116-127.
- [3] Samal, K. R., Babu, K. S., Das, S. K., & Acharaya, A. (2019). Time series-based air pollution forecasting using SARIMA and Prophet model. *Proceedings of the 2019 international conference on information technology and computer communications*, 80-85.
- [4] Ye, Z. (2019). Air pollutants prediction in Shenzhen based on ARIMA and Prophet method. *E3S Web of Conferences*. EDP Sciences.
- [5] Shen, J., Valagolam, D., & McCalla, S. (2020). Prophet forecasting model: A machine learning approach to predict the concentration of air pollutants ($PM_{2.5}$, PM_{10} , O_3 , NO_2 , SO_2 , CO) in Seoul, South Korea. *PeerJ*, 8.
- [6] Zhou, L., Chen, M., & Ni, Q. (2020). A hybrid Prophet-LSTM model for prediction of air quality index. *2020 IEEE Symposium Series on Computational Intelligence*, 595-601, IEEE
- [7] Tejasvini, K. N., Amith, G. R., & Shilpa, H. (2020). Air pollution forecasting using multiple time series approach. *Proceedings of the Global AI Congress 2019*, 91-100. Springer Singapore.

- [8] Ejohwomu, O. A., Shamsideen Oshodi, O., Oladokun, M., Bukoye, O. T., Emekwuru, N., Sotunbo, A., & Adenuga, O. (2022). Modelling and forecasting temporal PM_{2.5} concentration using ensemble machine learning methods. *Buildings*, *12*, 46.
- [9] Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.
- [10] Nosek, T., Brkljač, B., Despotović, D., Sečujski, M. & Lončar-Turukalo, T. (2020). *Praktikum iz mašinskog učenja*. Univerzitet u Novom Sadu, Fakultet Tehničkih Nauka, Katedra za telekomunikacije i obradu signala.
- [11] Nikolić, M., & Zečević, A. (2019). *Mašinsko učenje*. Beograd: Matematički fakultet.
- [12] Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia*, *126*, 1763-1768.



ISBN: 978-86-7031-623-2

PRIMENA JEDNOSTAVNIH METODA PREDVIĐANJA NA PODATKE VERTIKALNOG ELEKTROMETRIJSKOG SONDIRANJA

Filip Arnaut, Vesna Cvetkov, Dragana Đurić

Univerzitet u Beogradu, Rudarsko-geološki fakultet, Departman za geofiziku, Beograd, Srbija

E-mail:G804-21@rgf.bg.ac.rs

Ključne reči: Metoda proseka, Naive metoda, drift metoda, ekstrapolacija elektrometrijskih podataka

Napredak geofizike iziskuje integraciju različitih metoda koje se koriste u drugim naučnim oblastima, kao što je analiza i predviđanje vremenskih serija. Glavni cilj ovog istraživanja je razvijanje praktične, stabilne i precizne metode ekstrapolacije elektrometrijskih podataka. Pogodnosti takve metode ogledaju se u kraćoj akviziciji, odnosno istom utrošenom vremenu, ali sa više prikupljenih podataka tokom procesa obrade. Ekstrapolacija podataka sa nepristupačnih terena, gde nije bilo mogućnosti da se prikupe podaci, takođe, predstavlja pogodnost primene prediktivne metode.

Prva iteracija je usmerena na prikupljanje vrednosti ekstrapolacije, koje će u kasnijem istraživanju biti korišćene kao referentna vrednost za kompleksnije metode ekstrapolacije. U tom pogledu, tri metode su korišćene za referentno testiranje: Metoda proseka, Naive metoda i drift metoda.

Za testiranje je korišćen uzorak od 93 krive vertikalnog elektrometrijskog sondiranja, od kojih su 63 sintetički napravljene korišćenjem softvera, dok su 30 bile realne, merene krive sondiranja. Test je izveden tako što se obriše poslednja merena vrednost, ekstrapoluje se vrednost sa prethodno pomenutim metodama i uporedi se sa istinitom vrednosti. Ovakav vid istraživanja se može nazvati ekstrapolacija u uzorku sa dužinom horizonta od jednog podatka. Poređenje među izabranim metodama ekstrapolacije je vršeno u odnosu na srednje-kvadratno-odstupanje (RMSE), apsolutnu grešku (ABSE) i relativnu grešku (RELE).

Metoda proseka prikazala je najveće RMSE, ABSE i RELE vrednosti, sa maksimalnim vrednostima reda veličine 10^3 , zbog toga što uzima prosek svih prethodnih merenja. Detaljnije poređenje je urađeno između Naive i Drift metode. Obe metode imaju slične vrednosti RMSE od 14.2 i 12.79 kao i slične vrednosti ABSE od 20.08 i 18.08 Ωm , ali divergirajuće vrednosti RELE parametra sa 16.48 i 24.11 procenata. Drift metoda je pokazala preciznije vrednosti od Naive metode kod RMSE i ABSE parametra, ali je imala veću vrednost relativne greške zbog jednog podatka sa izrazito većom vrednosti, koja je iskrivila srednju vrednost RELE parametra. Analiza iskrivljenosti za RMSE i ABSE parametar prikazala je da Drift metoda ima manje vrednosti oba sračunata parametra u odnosu na Naive metodu (1.57 i 2.16). Ovo ukazuje da je raspodela za Drift metodu više simetrična i ima kraći rep raspodele. Analiza kurtoze prikazala je takođe manje vrednosti za Drift metodu sa 2.12 u poređenju sa 5.96 za Naive metodu.

U određenim slučajevima, najbolja metoda predviđanja je najjednostavnija metoda (ekonomičnost modela predviđanja). To neće uvek biti slučaj, ali ovakva istraživanja proizvode vredne podatke u vidu referentnih vrednosti za razvijanje budućih modela i algoritama za predviđanje. Ako kompleksan model ne može da prikaže bolje vrednosti od jednostavnog Drift ili Naive modela, onda se taj model može zanemariti i potraga za novim modelom nastaviti.

Razvijanje elektrometrijske metode zahteva dublje razumevanje prikupljenih podataka, kao i potrebu da se izvuku sve moguće informacije iz malog broja dostupnih podataka. Takav vid istraživanja implicira da bi obrada podataka trebalo da bude orijentisana ka savremenim postupcima i metodama njihove analize. Analiza i predviđanje vremenskih podataka može biti moćan alat, koji u takvim situacijama može pomoći, ali da bi se primenio potrebno je da se obavi njegovo detaljno istraživanje.

Ovaj rad finansiran je po „Ugovoru o realizaciji i finansiranju naučnoistraživačkog rada NIO u 2022. godini“, br. 451-03-68/2022-14/ 200126.

APPLICATION OF SIMPLE FORECASTING METHODS ON VERTICAL ELECTRICAL SOUNDING DATA

Filip Arnaut, Vesna Cvetkov, Dragana Đurić

University of Belgrade, Faculty of Mining and Geology, Department of geophysics, Belgrade, Serbia
E-mail:G804-21@rgf.bg.ac.rs

Key words: Average method, Naïve method, drift method, extrapolation of electrometry data

Multidisciplinary approach and using of methods applied in different scientific fields are of major importance for improving in geophysics. The time series analysis and forecasting are one of them. The main goal of this research was to create a practical, stable, and accurate electrometry data extrapolation method. The shorter the acquisition time, i.e., the same time spent but with more data obtained during the processing stage would be the major benefit of developing such a method, as well as extrapolating data from inaccessible terrain, with a lack of data.

The first iteration was focused on obtaining benchmark extrapolation values, supposed to be used as a reference point for future extrapolation methods that are more complex. Considering previous, three methods have been chosen: the Average method, the Naïve method, and the drift method.

A testing sample size included 93 sounding curves, of which 63 have been curves synthetically created and 30, calculated based on terrain measurements. Testing was conducted by deleting the last measured data point, extrapolating based on previous measured values and by comparing the extrapolated value to the true value. This can be denoted as an in-sample extrapolation with the forecasting horizon of one data point. The comparison of those methods has been done based on the root-mean-square-error (RMSE), absolute error (ABSE), and relative error (RELE).

The Average method showed the highest values of RMSE, ABSE, and RELE, with maximums of 10^3 , since it takes the average of all previous measurements. A more detailed comparison has been made for the Naïve and Drift method. Both have similar mean RMSE values of 14.2 and 12.79 and ABSE values of 20.08 and 18.08 Ωm , but diverging RELE values of 16.48 and 24.11 percent, respectfully. The drift method outperformed the Naïve method in the RMSE and ABSE error categories, despite having a higher RELE value due to one sounding curve having an erroneously high RELE value and distortion of the overall mean. Analysis of skewness for the RMSE and ABSE parameter revealed that the drift method had lower values than the Naïve method (1.57 and 2.16 respectfully). This indicates that the distribution for the drift method has a shorter tail and is more symmetric. Kurtosis analysis, also, displayed lower values for the drift method (2.12) compared to the Naïve method (5.96).

In some cases, the best forecasting method is the simplest one available (forecasting model parsimony). Even if this is not always the case, this type of research provides valuable reference values for developing and testing more complex forecasting models and algorithms. If a more complex model fails to outperform a simple drift or Naïve method, the model can be ignored and a new one found.

The advancement of electrometry method requires a deeper understanding of the acquired data, as well as the need to extract all the possible information from the sparse data available. This implies that data processing should be oriented towards modern procedures and methods of data analysis. Time series analysis and forecasting methods, in this view, can be a powerful tool that can help in such situation. However, to apply the method, it requires a detailed and thorough investigation.

This paper has been financed by the „Contract on realisation and financing of scientific research of SRI in 2022“, Nr. 451-03-68/2022-14/ 200126

ОСТАЈЕ ПУБЛИКАЦИЈЕ

SOME EXAMPLES OF ELECTRICAL IMAGING (ERT) AND GROUND PENETRATING RADAR (GPR) IN SOLVING CIVIL ENGINEERING PROBLEMS

Filip Arnaut, Branislav Sretenović

University of Belgrade, Faculty of Mining and Geology, Department of Geophysics, Belgrade, Serbia
G601-20@rgf.bg.ac.rs

Abstract: A few engineering geology and geotechnical case studies are seen in this research paper that are solved using two dimensional (2D) electrical imaging surveys (ERT) and, in four cases, ground penetrating radar (GPR). Case studies that demonstrate the versatility of the ERT method as well as other methods that complement obtained ERT data will be presented. Surveys of flotation tailings, detection of old mines, communal waste dumps, rock quality and lakebed surveys, and so on are among the case studies. Two-dimensional electrical scanning method proved its efficiency and cost effectiveness, especially when the survey needed a large surface area to be investigated.

Key words: electrical imaging (ERT); ground penetrating radar (GPR); engineering geology

INTRODUCTION

In many geological situations, 2-D electrical imaging surveys may provide valuable information that complements that obtained by other geophysical methods. Seismic methods, for example, are excellent at mapping undulating interfaces, but they fail (without using sophisticated data processing techniques) to map discrete bodies like boulders, cavities, and pollution plumes (Loke, 2012).

In certain cases, 2D electrical imaging proved to be the most effective form of investigation when addressing concerns like overflowing urban landfills and flotation tailings. This technique may also

be used to investigate the consistency of rock in quarries. This work includes a number of examples that back up this claim.

In areas with conductive unconsolidated sediments, such as clayey soils, ground penetrating radar (GPR) surveys may provide more accurate images, but they have limited depth of investigation. (Loke, 2012). Electrical surveys in two dimensions should be used in combination with seismic or GPR surveys since they provide additional knowledge about the subsurface. (Loke, 2012).

1. GEOPHYSICAL SURVEY OF THE FLOTATION TAILING IN BOR

Geotechnical and geophysical surveys were performed over the ore body "H" in order to survey the flotation tailing in Bor, Serbia. The exhausted exploitation of the ore body "H" resulted in the formation of a depression, which was filled with sand dams and used as a flotation tailing. To determine the thickness of the surface (filling) layer around the depression, the depth to the bedrock and the consistency of the bedrock and filled material, geoelectrical surveys and refraction seismic surveys were performed.

The results of 1D inversion of geoelectrical sounding data acquired using the Schlumberger array (blue line) and 1D refraction seismic data anal-

ysis (red line) at the north dam of the tailings are shown in Figure 1 (top). Refraction seismic surveys are conducted with two profiles of relatively short lengths, each with three ignition points, which does not provide enough data to detect lateral changes in elastic properties and wave velocity. As a result, only two boundaries have been identified, the second of which is located at a depth of 20 meters and has a gradual dip toward the end of the second seismic profile. Throughout the length of the geoelectrical profile, this boundary is deeper than the boundary defined by geoelectrical sounding, with the exception of the end of the geoelectrical profile, where both boundaries overlap.

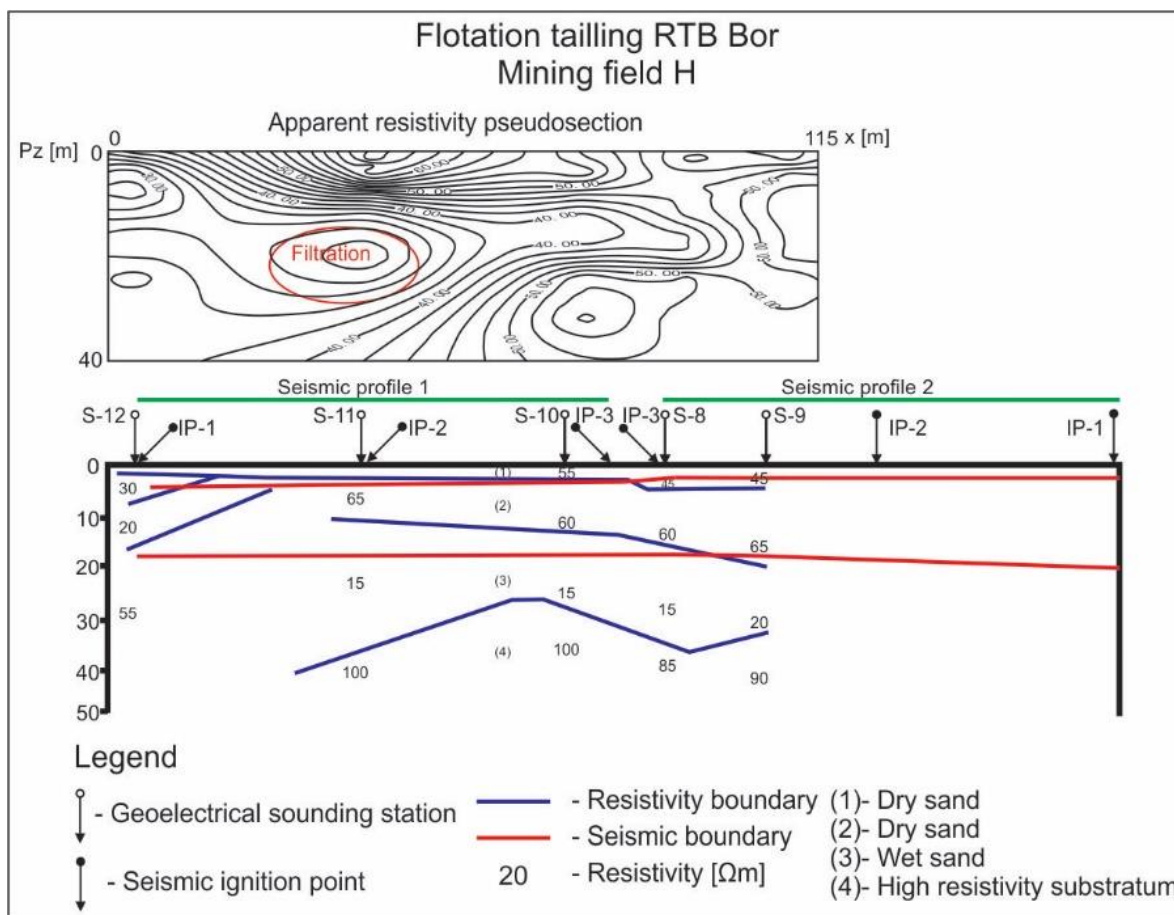


Fig. 1. Apparent resistivity section and results of 1D inversion of geoelectrical sounding diagrams and 1D refraction seismic inversion results along the sand dam axis

Since further separation of current electrodes during the Schlumberger array sounding along the north tailing dam was not feasible, the geoelectrical profile covered a smaller portion of the surveyed half space.

The apparent resistivity pseudosection (Figure 1, top) shows broad differences from a 1D surface model since apparent resistivity isolines are horizontal with small mutual gradient. Because of the divergence from a homogeneous, horizontal subsurface model between sounding stations S-11 and S-12, the results of 1D inversion of geoelectrical sounding diagrams in that region are unreliable. Since there were no lateral effects at smaller electrode spacings, 1D interpretation allowed for more accurate boundary depth determination. The subsurface model does not imply lateral changes in elastic properties at lower depths of investigation, as defined by the refraction seismic survey method.

During the survey, water began to accumulate in the depression behind the sand dam, which may be the result of water movement through the dam.

The apparent resistivity section shows a region of low resistivity values, which can be attributed to the filtration of a low resistivity electrolyte from the tailings through the dam. The filtration elutriates the sand material that can lead to perforation of the dam and subsequent ecological catastrophes. When the flotation tailing dam ruptured near Majdanpek (Serbia) in 1997, large areas of arable land and one village were contaminated with flotation sludge, permanently contaminating the soil and groundwaters in the accident region.

An approximate 2D subsurface model was constructed based on a 1D geoelectrical model and the apparent resistivity section (Figure 2). The apparent resistivity section with the Schlumberger array was calculated for that subsurface model (obtained by direct modeling in RES2DMOD) (Figure 2a). Theoretically calculated section for the 2D model of the dam body generally corresponds to the section of apparent resistivity (resistivity values and overall geometry) which was constructed based upon the measured geoelectrical sounding curves (Figure 1).

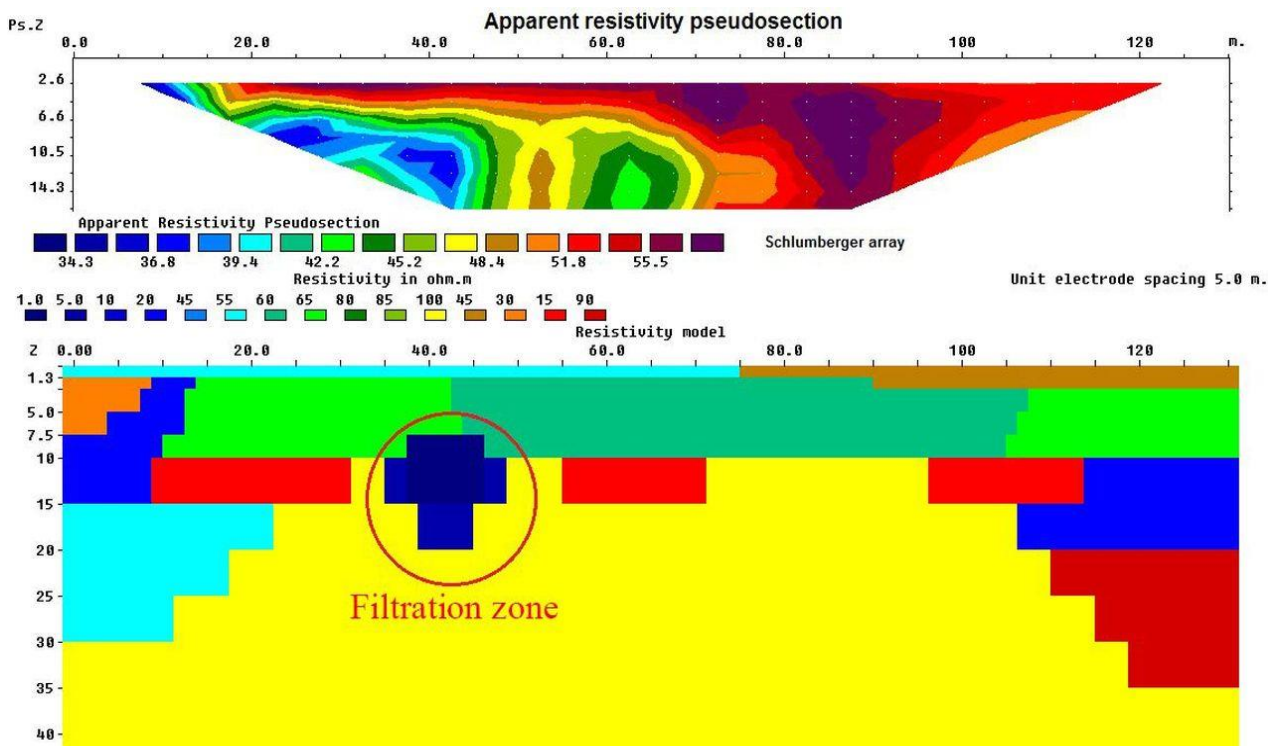


Fig. 2. 2D model constructed based upon 1D sounding curve inversion and the apparent resistivity section along the north dam axis with the Schlumberger array

A low resistivity zone (1–5 Ωm) was inserted into the 2D model to represent the electrolyte filtration zone in the flotation tailing. The RES2DINV software allowed 2D inversion of synthetic apparent resistivity sections, resulting in a 2D model (Figure

3) that better defines the subsurface architecture and true resistivity values that are closer to those described in the apparent resistivity section from Figure 2.

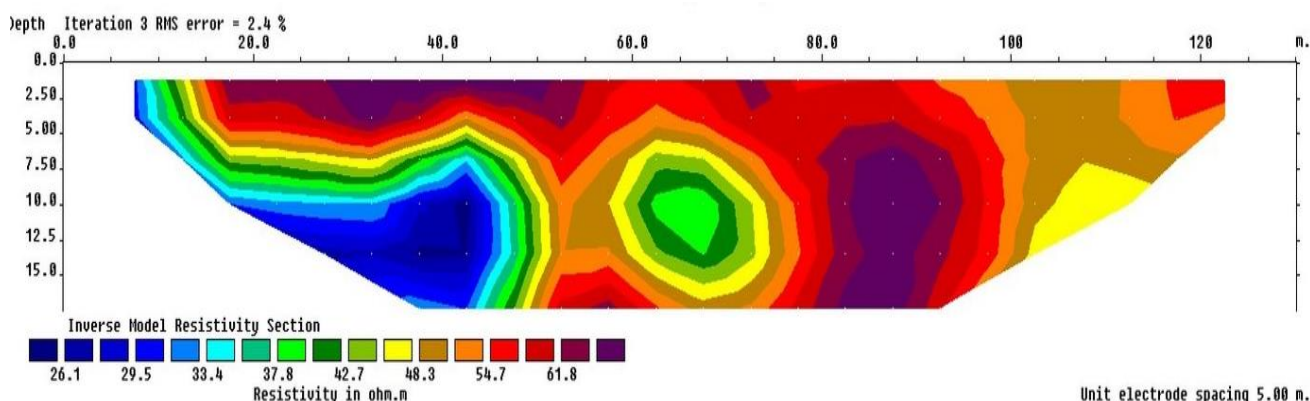


Fig. 3. 2D inversion of synthetic data for the 2D model along the north dam axis with the Schlumberger array

It is possible to create geoelectrical monitoring on all of the dams in the flotation tailing RTB Bor using the ERT method. Geoelectrical sounding surveys at the flotation tailing mining field "H" showed very low resistivity values of the tailing and electrolytes it protects (a few Ωm) allowing for detection

of filtration through the dam's body. A fully automated system, described as a part of the 2D geoelectrical survey method in Loke (2012), could be used for data acquisition in order to monitor the possible contamination of ground water.

2. DETECTION OF OLD MINES IN KOLUBARA – FORMER MINE “JUNKOVAC”

Coal extraction in the Kolubara (Serbia) open cast mine advanced to the region of old mine underground workings, which had been abandoned since 1971 and could not be continued due to the risk of gravitational collapses and cavities. During dredging machine operations, the existence of cavities was confirmed, while the presence of collapsed zones was suggested by intermittent small depressions in the terrain over old mine workings.

The exploitation of the former mine "Junkovac" was carried out using the roof collapse process, i.e. chamber excavation method, according to the review of existing documents (Jovanović et al., 1991). The chambers were $12 \times 12 \times 12$ meters in size, with each chamber holding 1700 m^3 of coal. Given that coal degree of utilization was about 50%, an empty chamber of $800\text{--}850 \text{ m}^3$ could be expected at the place where a chamber previously existed or in the space above the former chamber to the surface. Following work at the mine site, empty volumes were discovered in the region of consolidated lithological formations, which, due to their firm-

ness, prevented the soil from collapsing and subsiding. As a result, the location of the empty volumes cannot be accurately determined based on current mine exploitation documents.

Future excavation activities were deemed extremely dangerous without additional investigations aimed at detecting and defining cavities and disturbed zones. Since drilling alone was ineffective in solving the problem, two geophysical methods were tried (GPR and 2D electrical imaging). Nine nearly parallel old mine tunnels were crossed by 2D electrical imaging profiles (and the 10th perpendicular tunnel connecting them). To confirm the results of the ERT investigations, fifteen control drill holes were planned.

Experimental investigations with the ERT method over an old mine tunnel in Junkovac proved that such empty spaces can be detected successfully (Figure 4). Since the geological setting is straightforward (gently dipping clay and sandy clay strata) (Sretenović et al., 1992), observed anomalies can be correlated to disturbed zones formed by gravitational collapses of old mine tunnels and cavities.

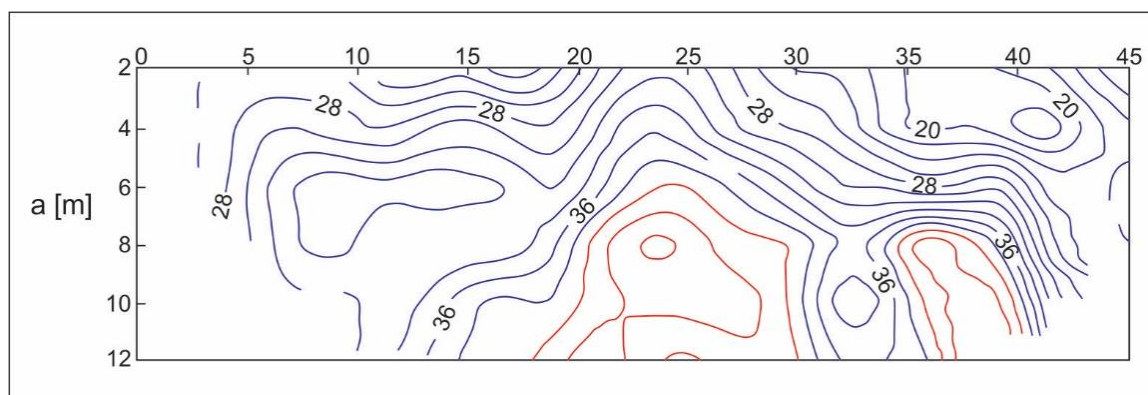


Fig. 4. Apparent resistivity section over a tunnel in Junkovac

2D electrical imaging investigations were performed along a 300-meter-long grid of parallel profiles with a 6-meter spacing between them. Twenty-five to thirty-five electrodes were used, with a minimum unit electrode spacing of ten meters, allowing for a maximum depth of investigation of 30 to 50 meters.

The apparent resistivity map (Figure 5), which was created using a grid of parallel ERT profiles with an electrode spacing of $a = 40$ meters (approximate depth of investigation of 20 meters), reveals

the existence of two anomalous high resistivity zones (A and B). Cavities or gravitational collapses in lower resistivity clay strata may trigger these zones. Drilling data showed that zone "B" corresponds to the presence of a high-resistance sand lens embedded in a lower-resistance sandy clay. Three drill holes in zone "A" confirmed the existence of the disruption in the relatively simple regional geological setting of gently dipping clay strata overlying coal strata as follows: yellow sandy clay, grey clay, grey clay with organic lamination, and finally coal.

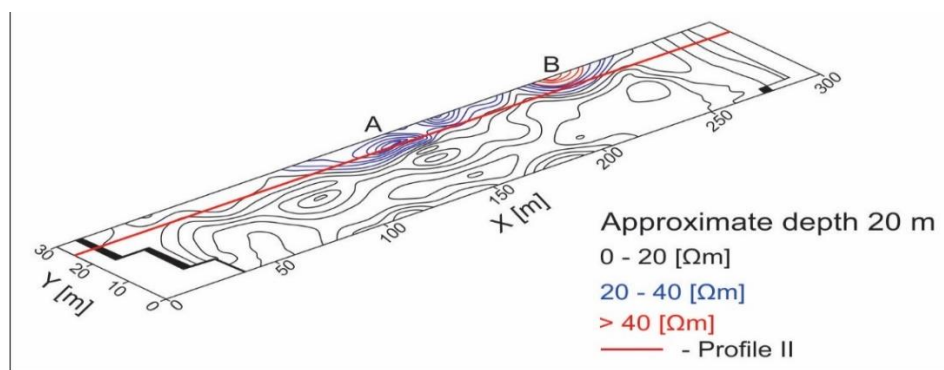


Fig. 5. Apparent resistivity map constructed for approximate depth of 20 meters ($a = 40$ m)

The apparent resistivity pseudosection along ERT profile II, as well as the 2D inversion model, were used to further investigate the anomalous zones (A and B) (Figure 6). Broad near-surface inhomogeneity spreads through the section, masking deeper structures in the apparent resistivity pseudosection. In the model, 2D inversion had to be achieved with a cell width of half unit electrode spacing (a) and 17 out of 147 data points had to be omitted (by trimming data set).

The border between grey clay and clay with organic lamination forms an anticlinal structure in the anomalous zone "A," indicating the sudden shallowing of this boundary. The depth to this boundary was determined by three control drills drilled at the 130th, 140th, and 150th meters of Profile II, respectively. Resistivity measurements on core samples

indicated a strong resistivity contrast at this boundary (1:3 to 1:5), explaining the apparent resistivity anomaly.

The third hole was drilled to a depth of about 60 meters, exposing the presence of two cavities, one at 44 – 48 meters depth and the other at 54–58.5 meters depth. The second cavity in the coal strata is most likely a section of the old mine chamber that has been partly buried by collapsed strata above it. This resulted in the upper strata subsiding and the forming of the anticlinal structure that was discovered (and upper laying cavity). As a result of the high resistivity contrast between grey clay and clay with organic lamination, the collapsed zone could be detected.

The Kolubara former mine detection was also investigated in Jovanović et al., (1991).

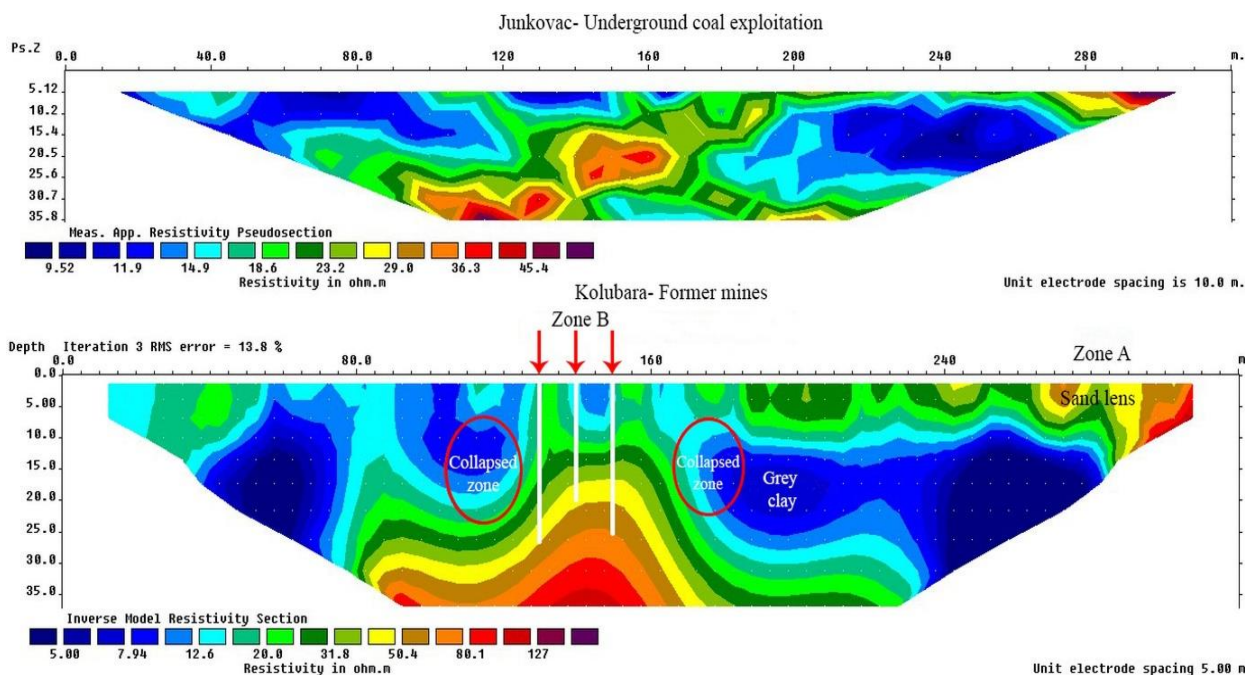


Fig. 6. Apparent resistivity pseudosection and 2D inversion model along Profile II

3. SURVEY AT THE COMMUNAL WASTE DUMP SITE "ADA HUJA"

Zones with filled artificial material or abandoned communal and industrial waste dumps can be detected using ground penetrating radar and geoelectrical surveys. In certain cases, there is inadequate knowledge of the distribution and composition of this material, which has unique geotechnical conditions. Figure 7 depicts the findings of the geoelectrical scanning survey conducted over the former communal waste dump at Ada Huja. The former communal waste dump was to be transformed into a train track and station known as "Karaburma".

A total of 17 electrodes were used in the geoelectrical surveys, with a unit electrode spacing of 5 meters. The survey was conducted on five depth levels, allowing for a total depth of investigation of 12.5 meters. The landfill's geometry is clearly evi-

dent, with low resistivity zones containing communal waste (dark clay, vehicle tires, building and organic material, and so on), and high resistivity zones containing marl and clay. The resistivity of gray sand with organic material is relatively low. Decomposition of organic material creates low-resistivity liquids (leachates), which can reduce the resistivity of the surrounding sand. The deepest part of the profile is made up of compact dark grey marl, which resistivity-wise is nearly identical to the filed material.

In comparison to the geotechnical profile, which was constructed with three boreholes, the 2D geoelectrical model can better define the subsurface. By reducing the unit electrode spacing in a multi-electrode array, the horizontal resolution of geoelectrical surveys can be enhanced, allowing for a more accurate image of subsurface lateral resistivity changes.

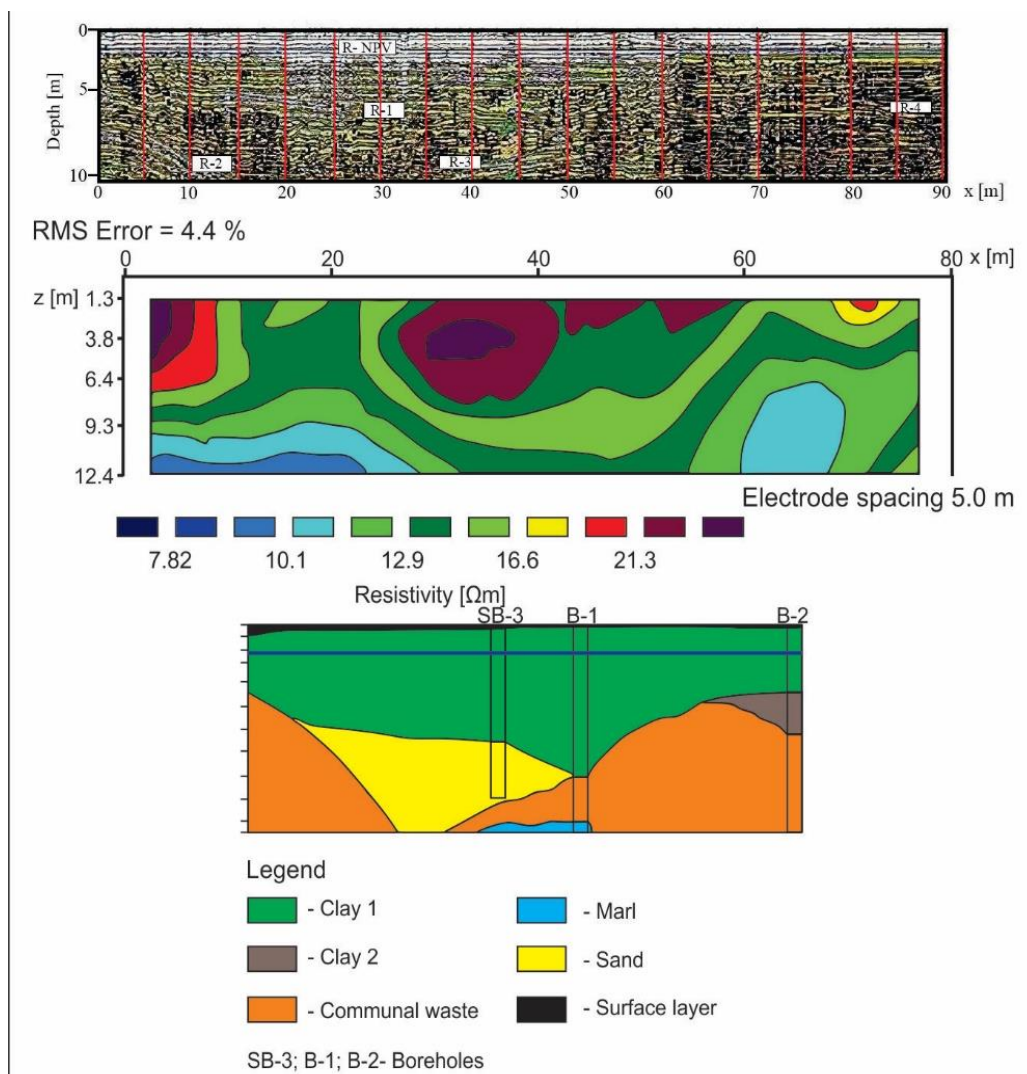


Fig. 7. GPR section (top), ERT section (middle) and the geotechnical profile (bottom)

Ground penetrating radar was used as part of a survey to determine the feasibility of constructing the railway plateau without excavating the communal waste complex. The GPR survey successfully demonstrated the ability to identify the spatial relationship of the communal waste dump, its base and surface layers, as well as the groundwater level and different composition zones within the communal waste dump. Figure 7 (top) shows the GPR survey results, which were collected using a 300 MHz antenna along the same route as the geoelectrical profile. At shallower depths, distinct amplitude reflections can be seen as a result of the surface gravel layer, which has a thickness of a few decimeters; the next reflection R-npv corresponds to the groundwater level determined by borehole results. For each GPR reflection one thinner and two thicker stripes ("triples") that correspond to one sinusoidal signal with a one-color amplitude (peak), a zero line (thinner line), and then a negative amplitude (trough) can be seen. The boundary of dark grey clays and sand corresponds to reflection R-1, the boundary of communal waste-sand to reflection R-2, the boundary of marl-communal waste to reflection R-3, and the boundary of dark grey clays – marl clays to reflection R-4.

Given the extremely low resistivity values found in the surveyed subsurface, which can attenuate the electromagnetic signal, the obtained results are excellent in terms of both depth of investigation (around 10 meters) and resolution. The series of resistivities found in the subsurface was advantageous in this survey; the surface layer has a resistivity of

around 30 Ωm , which is accompanied by lower resistivities found deeper in the subsurface with a resistivity of around 10 Ωm . The detected marl cannot be distinguished from communal waste based on resistivity, but it does contain a significant amount of water, which can cause a contrast in the GPR section. In comparison to the geotechnical profile, the GPR section has more details and a higher resolution than the geoelectrical profile. Figure 7 shows the original GPR section with the attenuated zones (darker zones) as well as the 2D geoelectrical imaging model and the geotechnical model.

Geoelectrical scanning could be done with more electrodes and a smaller unit electrode spacing (less than 5 meters), which would improve horizontal resolution but not vertical resolution. Following the test survey's findings, it was determined that in future detailed surveys, GPR and geoelectrical methods would be used instead of the seismic method. Depth conversion, or the transition from time domain to depth domain, can be difficult with GPR surveys. Given that distinct lateral changes can be assumed with the communal waste dump, depth conversion would be difficult with longer profiles where borehole data would be scarce. Another issue would be a low resistivity surface layer at certain points along the profile, which would attenuate the GPR signal and thus restrict the depth of investigation. Signal attenuation and lack of reflection continuity are strong indicators of potential contamination, and they can help map contaminated areas.

An extensive investigation into the communal waste dump "Ada Huja" is given in Sretenović et al. (2019).

4. ROCK QUALITY SURVEYS WITH THE GEOELECTRICAL SCANNING METHOD

In order to investigate the rock quality at the granitoid rock quarry " Pločnik " in Brajkovac near Lazarevac, complex geophysical, geotechnical, and remote sensing surveys were conducted. In order to achieve the most reliable and cost-effective combination of geophysical methods, GPR, geoelectrical scanning, and seismic refraction with one geophone (single geophone refraction profiling – SGRP) were used. Five boreholes were available during the geophysical data acquisition, three of which revealed unusually high surface layer thicknesses (up to 15 meters). The top layer is made up of a thin layer of clay and disintegrated material. In addition, distinct lateral changes in thickness and physical properties were estimated. Due to signal attenuation of the low

resistivity clay surface layer, GPR and SGRP were unable to provide useful information.

The Single Geophone Refraction Profiling (SGRP) method was used along the electrical scanning profiles to investigate the lateral variations in rock masses using another physical property, lateral velocity variations, which revealed a relatively thinner covering layer. The tests were carried out with two separate offsets and maximum penetration depths of around 5–8 meters. This relatively simple and fast geophysical qualitative approach had a strong connection with the results of the 2D analysis of electrical imaging data sets (increased arrival time denoting lower velocity and resistivity in Figure 8) and could be used as a supplement to confirm the existence of sharp lateral lithological variations.

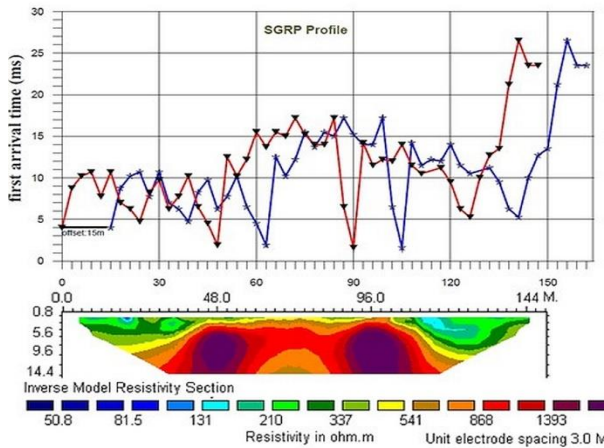
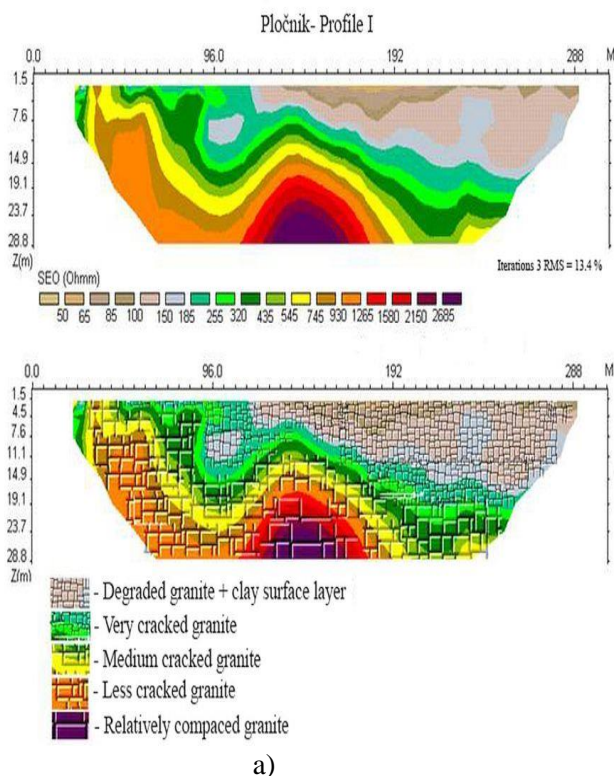


Fig. 8. Qualitative refraction survey with the SGPR method (measurements are attached to the position of the ignition point – red line, and to the position of the geophone – blue line)

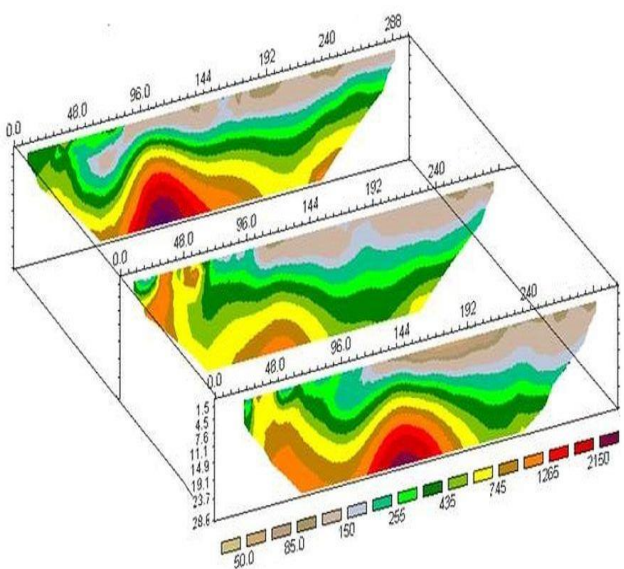
Figure 9a demonstrates how geoelectrical scanning was useful to a depth of over 30 meters, enabling disintegrated material with lower resistivity values to be separated from more compact rocks with higher resistivity values. On a series of parallel profiles, the lateral changes were correlated, suggesting the 2D geometry of subsurface structures (Figure 9b). The geoelectrical survey was carried out on ten depth levels, with a varying unit electrode spacing controlled by the thickness of the surface layer, allowing for different depths of investigation and horizontal resolutions. Bigger unit electrode spacings were imposed by thicker surface layers, lowering horizontal resolution.

The Pločnik quarry has also been investigated in Sretenović et al. (1999).



a)

Fig. 9. (a) 2D model obtained after inversion of profile I at “Pločnik”
(b) Correlation of three parallel profiles at the rock quarry “Pločnik”



b)

5. 2D GEOELECTRICAL SURVEY FOR THE SAINT VASILJE OSTROŠKI TEMPLE FUNDING

It is important to assess the condition of the soil and bedrock during the construction of a building in order to fund and determine the risk of subsidence. 2D geoelectrical surveys will provide detailed information on the type of bedrock and the overall subsurface where a potential structure will be constructed. Figure 10 shows a 2D electrical imaging survey along two perpendicular profiles on

the site of “Saint Vasilje Ostroški” proposed temple in Banjica (Belgrade). Profile I has an east-west heading and indicates a steady rise in the thickness of high resistivity sediments (humus, marl like deluvium, clays, etc.) from the 76th to the 100th meter on the profile, with a west heading.

Local low resistivity inhomogeneities (most likely dusty clays) can also be seen in the sediment

region between the 12th and 16th meters, the 20th and 24th meters, and the 52nd and 60th meters. The 3.2 percent RMS difference between observed and theoretically calculated data indicates that local inhomogeneities have no significant impact on 2D data inversion. Profile II, which has a north-south heading and a lower RMS value of 2.7 percent, can be approximated graphically as a simple 1D subsurface architecture. Profile II measurements were taken with a unit electrode spacing of 3 meters, allowing for greater horizontal resolution and a shallower depth of investigation of 9.5 meters compared to profile I (13 meters).

Due to signal attenuation and ambient noise, refraction seismic surveys that were also conducted, which identified the near surface layer with a thickness of a few meters but had limited depth of investigation. Borehole data allowed for the geological identification of geoelectrical layers, as well as the investigation of the drilling core through geomechanical laboratory experiments. Since there are no discernible lateral changes in the geoelectrical survey area, the data from the geomechanical laboratory may be extended to the entire survey area, lowering the survey cost.

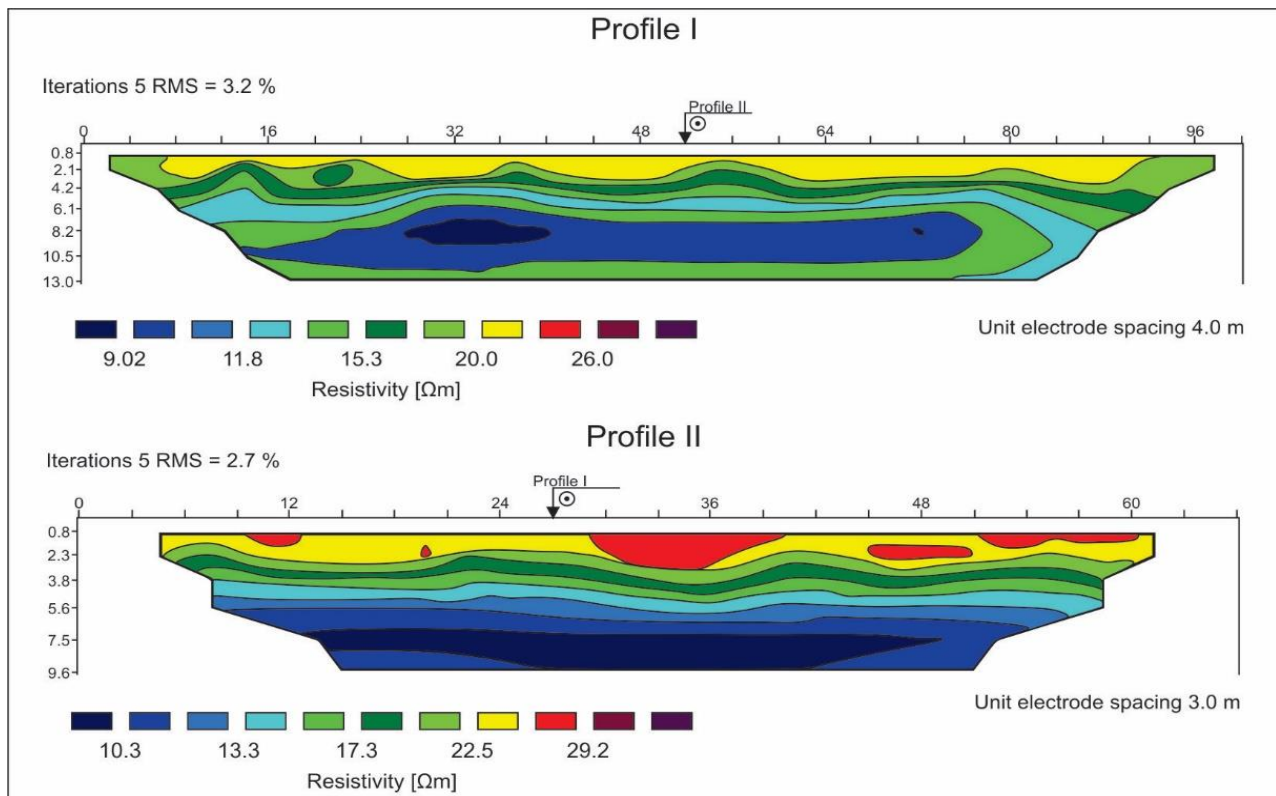


Fig. 10. 2D geoelectrical survey for the funding of the temple of Saint Vasilije Ostroški

6. "ADA CIGANLIJA" LAKEBED GPR SURVEY

The artificial lake "Ada Ciganlija" is a popular tourist, leisure, and excursion destination in Belgrade (Serbia), attracting hundreds of thousands of visitors each day during the summer season. At the same time, since this lake is close to the groundwater resource "Makiš," it should be given special attention as a form of ecological protection.

Divers and an eco-sounder were used to assess the thickness of the sedimented mud and sludge during the monitoring process. Without knowing the

structure of the lakebed, those measurements could only reflect its morphology. The use of ground penetrating radar allowed data on lakebed thickness and morphology to be collected. A plastic boat was used to conduct the GPR survey (plastic material does not alter the propagation of electromagnetic waves), with an 80 MHz antenna. The maximum depth of investigation was about 8 meters, which is similar to the maximum depth of investigation that an 80 MHz antenna in water would achieve.

Figure 11 depicts the lake's morphology clearly, with just a few reflections from the mud and clay bedrock. Figure 11 shows that the mud thickness is very low, so the pond succession process is not considered dangerous. Because of the lower slope angle, the east side produces better reflections. In the presence of higher slope angles, reflective methods are known to be limited. When it comes to the slope angle, GPR surveys have a limit of 30–35 degrees; values above that limit represent the elec-

tromagnetic signal towards the surface, i.e. the receiving antenna. Over the winter, while the lake is frozen, the same survey may be done more effectively. Since ice is essentially an insulator, it does not interfere with electromagnetic wave propagation. Since water is homogeneous and has a constant electromagnetic wave propagation velocity of 3.3 cm/ns, GPR surveys over water bodies are among the easiest to interpret.

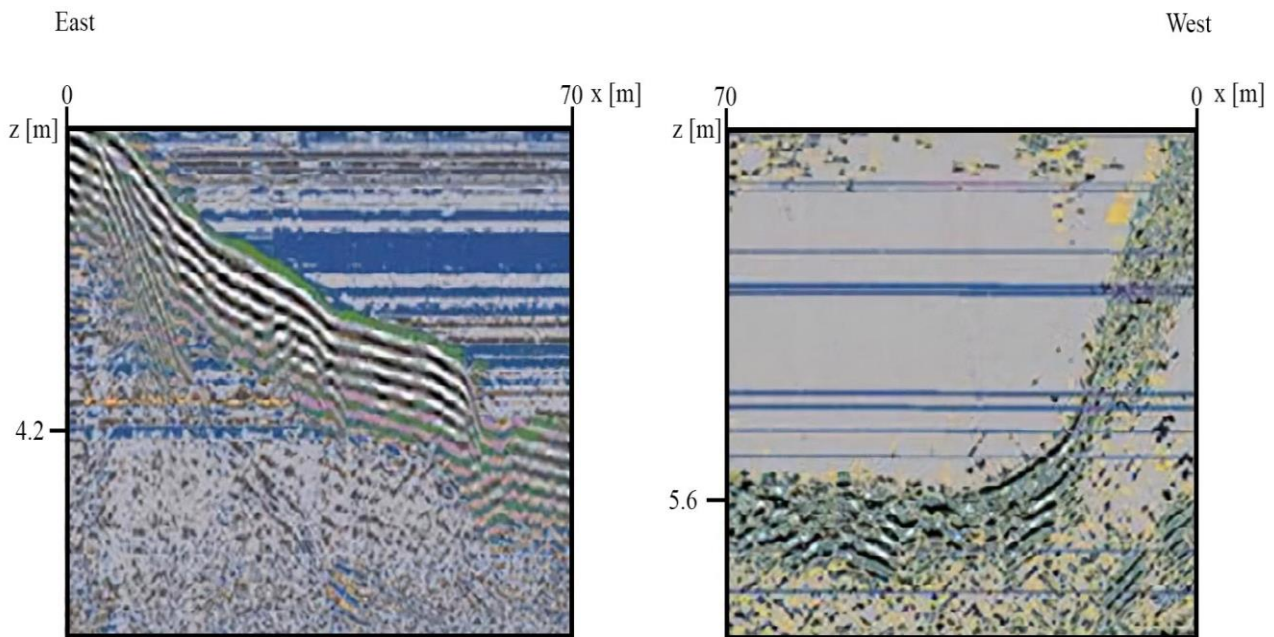


Fig. 11. GPR survey over the east and west coast of the Ada Ciganlija lake

7. BELGRADE AIRPORT RUNWAY GPR INVESTIGATION

The Belgrade Airport's runway needed immediate reconstruction in order to extend its operational lifetime, improve protection when in use, and install visual navigation instrumentation. GPR surveys were conducted during the preliminary phase of runway reconstruction to ensure that the data acquired was of high quality. The main aim of the test was to look for potential linear cavities at the intersection of the concrete slab and the gravelly buffer layer, as well as the buffer zone's boundary with the lost bedrock.

Because of the runway's layout, anomalies in the zone of metal joints between concrete slabs are probable. Since the same anomaly can be observed in cavity zones, this was just an observation. A 500 MHz antenna was used to fulfil the preliminary project goals; the antenna acted as a compromise between the depth of investigation, which was about

1.5 meters, and the resolution needed to detect the 0.5×0.1 -meter cavities. The 40-ns length of the GPR radiogram was sufficient for obtaining a satisfactory depth of investigation; however, depending on the antenna velocity, a total of 12–25 impulses per second ("scans") were used. The maximum speed of the acquisition vehicle was 20 km/s, and each profile (scan) contained 512 digital samples.

Multiple high-amplitude reflections at the metal joints between concrete slabs that are uniformly spaced at 8 meters is clearly visible in the raw GPR subsurface image (Figure 12a). The reflections from the boundary concrete-gravel buffer layer, as well as the reflections from the gravel-loess bedrock boundary, are partly hidden by these high reflections. The antenna noise is visible in the raw data at 30 ns, as well as some high frequency noise. The

noise effects were minimized by data processing that included horizontal and vertical data filtering with a low bandwidth filter. A minor deflection on the profile's left side that can lead to concrete slab failure. Increased water content can be indicated by an increase in dark coloured amplitude reflections at the contact of gravel and loess in the non-disturbed

region, which increases the contrast between dielectric constants in two mediums.

Figure 12^b shows the most remarkable example of GPR surveys' powerful resolution. With both reconstructed and non-reconstructed runway areas, the profile traverses above the field.

The Belgrade Airport runway survey can also be seen at Sretenović et al. (1997)

GPR Section over the Belgrade Airport runway

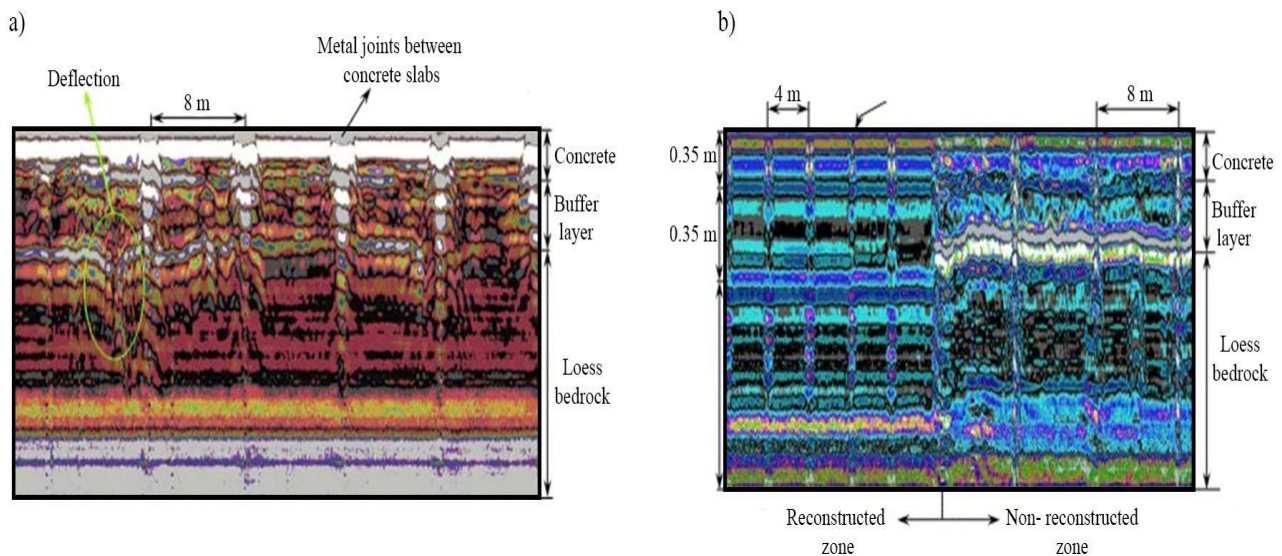


Fig. 12. (a) Raw GPR image, (b) GPR image over the reconstructed and non-reconstructed zone

8. CONCLUSION

In comparison to simpler refraction seismic surveys with inadequate number of ignition points, geoelectrical surveys that are based on the 2D imaging method provide better results. With a large enough number of ignition points, refraction seismic surveys could better assess lateral changes and the velocity of elastic wave propagation in the subsurface.

Where there is a need to cover a wider surface area, which borehole data cannot provide effectively, the geoelectrical scanning method shows its benefits. The ERT method provided useful surface knowledge about plausible cavities in the “Junkovac” example. Two anomalies were subjected to boreholes, one of which was discovered to be a sand lens and the other to be a cavity. Borehole data should resolve the uncertainty between the high resistivity material and cavities.

In the case of the “Ada Huja” communal waste dump, the low resistivity communal waste is easily

detected and differentiated from the higher resistivity marl bedrock using the ERT method. The geoelectrical model, as in the previous case, revealed more information about the subsurface than the geotechnical model. Signal attenuation and a lack of reflection continuity during GPR surveys at the “Ada Huja” communal waste dump site were strong indicators of subsurface contamination.

The SGRP method, in conjunction with the ERT method, provided valuable information about subsurface lateral changes at the “Pločnik” granitoid quarry. Furthermore, the ERT method provides accurate data on the degree of rock disintegration.

When using the GPR method to survey a lakebed, a coast with a lower slope value, usually between 30 and 35 degrees, should be favoured. Since water is homogeneous and has a constant electromagnetic wave propagation velocity, GPR data interpretation over a water body is relatively simple.

At the Belgrade Airport runway, GPR surveys demonstrated its high-resolution capabilities.

REFERENCES

- Jovanović, B., Cojić, M., Sretenović, B. B. (1991): Tehnologija otkopavanja rotornim bagerima SR 1200 24/4 u zoni starih jamskih radova. In: *VIII jugoslovenski simpozijum o površinskoj eksploataciji mineralnih sirovina*, Herceg Novi, 1991, 211–218 pp.
- Loke, M. H. (2012): Tutorial: 2-D and 3-D electrical imaging surveys.
- Sretenović, B., Miličević, B., Marčetić, D. (1992): Detection of disturbed zones over old mine workings using electrical tomography. In: *54th Meeting of the EAEG*, Paris, France, 1992, 632–633 pp.
- Sretenović, B. B., Miletić, V., Djordjević, J. (1997): The application of combined GPR and resistivity investigations in solving different civil engineering, engineering geophysics and environmental problems. In: *3rd EEGS Meeting*, 1997, 447–450 pp.
- Sretenović, B., Vukadinović, M., Vasić, R., Sretenović, N., Bosnjaković, I. (1999): Electrical scanning and SGPR methods for the detection of lateral velocity and resistivity changes in the "Pločnik" quarry. In: *EEGE-ES abstracts*.
- Sretenović, B., Arnaut, F., Vasiljević, I., Cvetkov, V. (2019): 2D geoelectrical resistivity tomography application at the former city waste dump "Ada Huja": Eco-geological problem. *Podzemni radovi*, **34**, 59–76.

Резиме

НЕКОИ ПРИМЕРИ НА ПРИМЕНАТА НА МЕТОДОТ НА ЕЛЕКТРИЧНО СКЕНИРАЊЕ (ЕРТ) И ГЕОРАДАР (ГПР) ВО РЕШАВАЊЕ ПРОБЛЕМИ ОД ГРАДЕЖНОТО ИНЖЕНЕРСТВО

Филип Арнаут, Бранислав Сретеновиќ

Универзитет во Белград, Факултет за рударство и геологија, Катедра за геофизика, Белград, Србија
G601-20@rgf.bg.ac.rs

Клучни зборови: електричен метод (ЕРТ); георадар (ГПР); инженерска геологија

Во овој истражувачки труд се сретнуваат неколку инженерско-геолошки и геотехнички истражувања на случаи, кои се решени со помош на дводимензионален (2D) електричен метод (ЕРТ), а во четири случаи со георадар (ГПР). Прикажани се случаи кои ја демонстрираат разновидноста на ЕРТ-методот, како и други методи кои ги надополнуваат добиените ЕРТ-податоци. Извршени се истра-

жувања на флотациска јаловина, детекција на стари рудници, депонии на комунален отпад, истражувања на квалитет на карпа и корито и сл. Методот на дводимензионално електрично скенирање ја докажа својата ефикасност и ефективност во поглед на трошоците, особено кога истражувањето бараше да биде истражена поголема површинска зона.

Inadequacy of 1D, 2D and 3D Resistivity Inverse Modelling in the Presence of Electrical Anisotropy

Branislav Sretenovic, Filip Arnaut

University of Belgrade, Faculty of Mining and Geology, Department of Geophysics, Belgrade, Serbia

Email address:

branislav.sretenovic@rgf.bg.ac.rs (B. Sretenovic), G94-16@rgf.bg.ac.rs (F. Arnaut)

To cite this article:

Branislav Sretenovic, Filip Arnaut. Inadequacy of 1D, 2D and 3D Resistivity Inverse Modelling in the Presence of Electrical Anisotropy. *Earth Sciences*. Vol. 8, No. 2, 2019, pp. 102-116. doi: 10.11648/j.earth.20190802.14

Received: February 11, 2019; **Accepted:** April 9, 2019; **Published:** April 29, 2019

Abstract: Universal property of matter is the variation of a certain physical characteristic in different direction. The Substances that do not display this property are an exception. Anisotropy, as this property is named, is also notable in electrical conductivity of minerals, ores, rocks and geological formations. In order to properly define a geological model of an investigated area, it is necessary to account for electrical anisotropy and lateral effects of different origin that are almost always present phenomena. The degree of knowledge of these phenomena determines the quality of interpretation. The effects of electrical anisotropy on 1D, 2D and 3D inversion of apparent resistivity data were examined and the way to detect, quantify and analyse electrical anisotropy in 3D case is proposed. The results of this analysis showed that electrically anisotropic models have led to the totally erroneous results of 3D inversion while the effects of electrical anisotropy on 1D and 2D inversion were less pronounced. Among several existing ways of collecting 3D apparent resistivity data the complete Pole-pole array data set is the only one suitable for detection of electrical anisotropy. Pole-pole resistivity data enable calculating of corresponding resistance or apparent resistivity values for all other collinear or square arrays. This fact makes possible to use standard grid of electrodes in performing 3D Pole-pole apparent resistivity measurements for calculating square array apparent resistivity data. In the special case of vertical stratification (fracturing, schistosity..) this calculated square array data were used for electrical anisotropy analysis thus determining values of coefficient of anisotropy (λ) and mean geometric resistivity (ρ_m). The simple two and three layer 1D synthetic anisotropic models, were used to determine parameters of electrical anisotropy by using 3D forward modelling to calculate Pole-pole and then square array apparent resistivity values. The mean geometric resistivity data obtained by using square array, which are orientation-independent, were used for 1D inversion leading to more realistic results. In the case of oblique stratification (lamination, fracturing, schistosity or karstification) crossed square resistivity data can be reconstructed from 3D Pole-pole data set and then can be used to get parameters of electrical anisotropy, namely apparent coefficient of anisotropy (n), mean geometric resistivity (ρ_m) and apparent electrical strike (θ). These parameters provide an indication that investigations of electrical anisotropy should be conducted (by using square and crossed square array) in order to avoid erroneous results of 3D inversion.

Keywords: Electrical Anisotropy, Oblique Lamination, Rock Fracturing, Schistosity, Coefficient of Anisotropy, Forward Resistivity Modelling, Inverse Resistivity Modelling

1. Introduction

Different sedimentological environments (fluvial, lacustrine, glacial, glacio-fluvial, glacio-lacustrine, coastal sand dunes, river dunes, beach ridges, near coastal marine sands, estuarine deposits) lead to depositing sediments characterized by various kinds of oblique lamination which can be resolved by using high resolution ground penetrating radar (GPR) or reflection seismic methods. Electrical DC or

low frequency (several Hz) investigation can't resolve such fine lamination inside layers. Various kinds of oblique lamination can cause phenomenon of electrical anisotropy which in return greatly affects 1D, 2D and 3D resistivity inverse modelling regarding interpreted depths to layer boundaries as well as "true" resistivity of such anisotropic media. The same is the case when dealing with rock fracturing, schistosity or karstification.

Many authors have been dealing with the electrical

anisotropy caused by fracturing. One can find very comprehensive description of the phenomenon as well as previous developments in the recent paper by

N. Inyang Udosen and N. Jimmy George [1]. These authors performed field electrical resistivity tomography (ERT) investigations along two perpendicular lines and at the crossing of these profiles they conducted azimuthal square array resistivity measurements (changing square array orientation from 0° to 180° in increments of 15°) and determined electrical anisotropy. To compute the fracture strike analytically, the crossed square array was used [2]. These authors report that: "At small square array spacing of 5 and 7.1 m, slight anisotropic responses may be present on the plots but lateral resistivity variations exceed that due to anisotropy. At an electrode spacing of 10 m, the plot became elliptical, indicating anisotropy. These plots became more elliptical as the electrode spacing increased from 10 to 50 m. The anisotropic coefficients of the limestone layer were found to range from 1.37 ± 0.075 to 1.44 ± 0.075 , indicating moderate anisotropy". They also report that: "ERT data was acquired to generate a 2D image of the subsurface, determine depth to bedrock, and ascertain the lithological variations within the subsurface." These goals are hard to achieve in the presence of electrical anisotropy by using ERT data. Although these authors claim that "Ignoring the effects of electrical anisotropy within the subsurface while taking field measurements gives unsatisfactory results" [1], they do ignore the effects of electrical anisotropy on 2D inversion of ERT data claiming that "The inversion model was considered reasonably correct due to the small root mean square error of 0.67%". Electrical anisotropy probably affected the results of 2D inversion of ERT data along the two perpendicular profiles but authors didn't take it in consideration. The aim of this paper is to show how inadequate 1D, 2D or 3D resistivity inversion can be in the presence of electrical anisotropy but also to suggest the possibility of detecting and quantifying electrical anisotropy in 3D case.

Colossal advance in the field of DC resistivity method took place in the last thirty years, regarding instrument improvements, automatization of data acquisition and especially in forward and inverse modelling techniques. All improvements that had been achieved allows us to try to improve limited resolving power of DC resistivity method by analysing electrical anisotropy and by using its elements, namely the mean geometric apparent resistivity (ρ_m), apparent coefficient of anisotropy (n) and apparent electrical strike (θ).

Bhattacharya and Patra [3] solved analytically problem of a potential distribution over homogeneous anisotropic subsurface model (Figure 1). Model is made under the assumption that the plane of stratification, lamination or bedding is tilted by an angle (α) in respect to the surface and (ϕ) is the angle between layering strike and current electrodes (in the case of square array).

Property of a homogeneous anisotropic subsurface is dependence of measured resistivity upon the direction. Value of true resistivity parallel to the plane of stratification (ρ_1) is

smaller than the value perpendicular to that plane (ρ_2). These two values determine coefficient of anisotropy $\lambda = \sqrt{\rho_2/\rho_1} > 1$. On the other hand, measured apparent resistivity varies with array orientation to the electrical strike, being minimal when the orientation is perpendicular to it and maximal in the case of parallel array orientation. This is well known paradox of anisotropy which is characteristic of all collinear arrays.

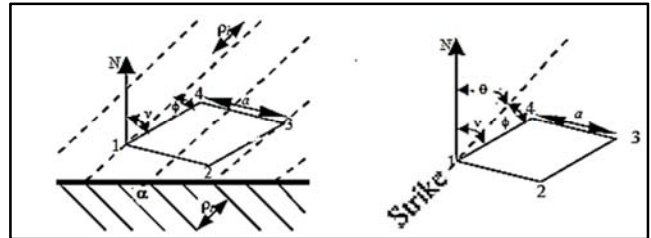


Figure 1. Homogeneous anisotropic subsurface model [3].

For modelling purposes software created by M.H. Loke [4-6] as well as 1D resistivity sounding inversion software created by Vander Velpen [7] were used. The three layer 1D anisotropic model was created by using 2D forward modelling (RES2DMOD). For this three layer 1D model, with the highly anisotropic middle layer, the Wenner apparent resistivity cross-section was calculated. The calculated values were used to reconstruct Wenner alpha sounding curve and then to use it for 1D inversion. There is an excellent option offered by RES2DMOD software to save calculated apparent resistivity data in the format used by inversion RES2DINV software and to simulate measured data by adding a certain level of random noise to the calculated data ($\pm 2\%$ in this case). Theoretically calculated data with $\pm 2\%$ random noise added to it was then inverted by using RES2DINV software and as a result an inverse model resistivity section was created. The same approach of combining forward and inverse modelling was applied for more complex 2D anisotropic models as well as for 3D anisotropic models.

2. The Results of Analysing Anisotropic Models

2.1. One-Dimensional Anisotropic Models

Figure 2 (lower part) shows 1D anisotropic three layer model, the middle layer being highly anisotropic, while the relatively thin covering layer (2.5m, 50 Ohmm) and the third layer (3000 Ohmm) are homogenous and isotropic. The middle layer consists of steeply dipping thin layers with alternating true resistivity values of 300 and 10 Ohmm respectively, which gives rise to strong electrical anisotropy. This anisotropy is detectable by resistivity measurements conducted from the surface due to inclination of alternating high and low resistivity thin layers.

The upper part of Figure 2 shows the calculated Wenner array apparent resistivity cross-section which indicates horizontal layering where the middle layer (blue zone of the

section) appears as an isotropic and relatively low resistive (28-32 Ohmm) layer. The third high resistivity layer (purple zone of the section) is indicated by maximum apparent resistivity values of about 95 Ohmm. After 2D inversion of this apparent resistivity cross-section the 2D inverse model resistivity section is obtained with RMS error of 1.88% (Figure 3). The middle anisotropic layer is characterized by even lower “true” resistivity value (16.5-27.5 Ohmm) and it appears to be an isotropic layer. Two-dimensional inversion clearly indicates presence of the covering layer (with the resistivity of 45-50 Ohmm) and high resistivity substratum (over 650 Ohmm). The depth to the third layer (indicated by

solid red line) is underestimated by at least 15-20% due to the fact that the middle layer is anisotropic. Furthermore there is no indication of electrical anisotropy of the second layer. In the case of horizontal stratification or lamination, anisotropy is not detectable by using measurements conducted from the surface and interpreted depth to substratum would be overestimated by the factor equal to the coefficient of anisotropy $\lambda = \sqrt{\rho_t / \rho_l} > 1$ [8]. Inclined stratification or lamination leads to apparent coefficient of anisotropy (n) which is always smaller than (λ), except in the case of vertical stratification when it is equal to (λ).

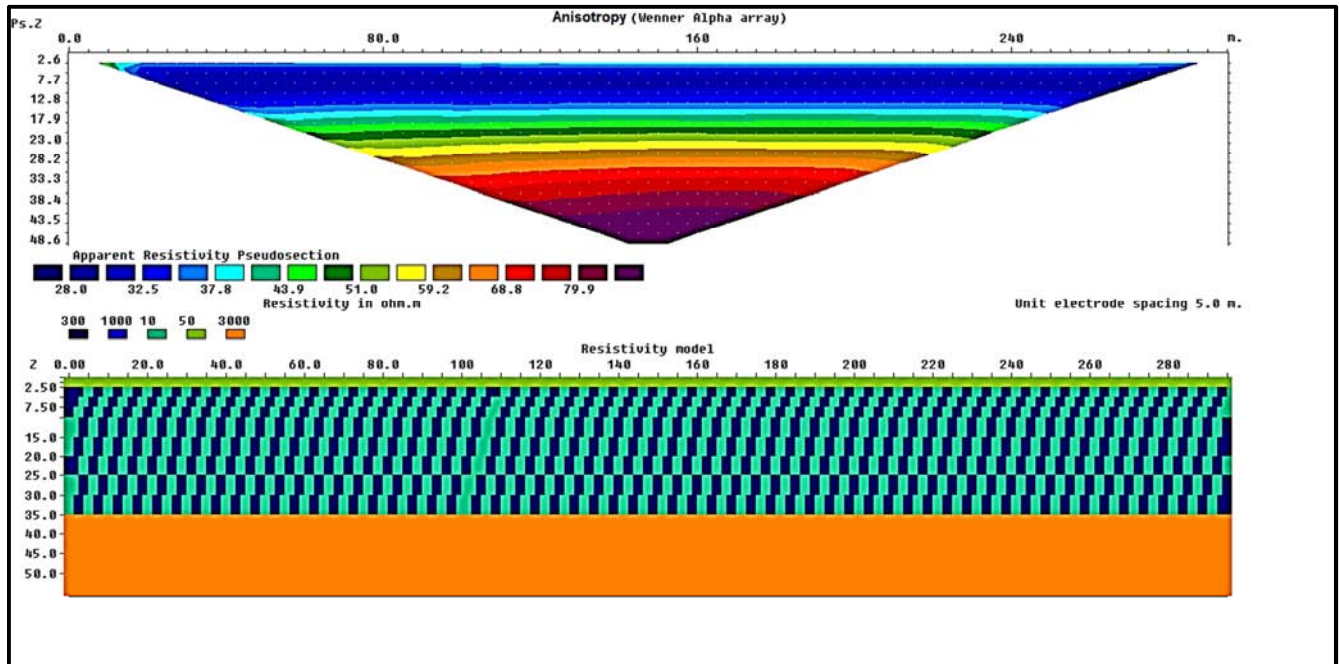


Figure 2. 1D anisotropic three layer model and corresponding apparent resistivity cross-section.

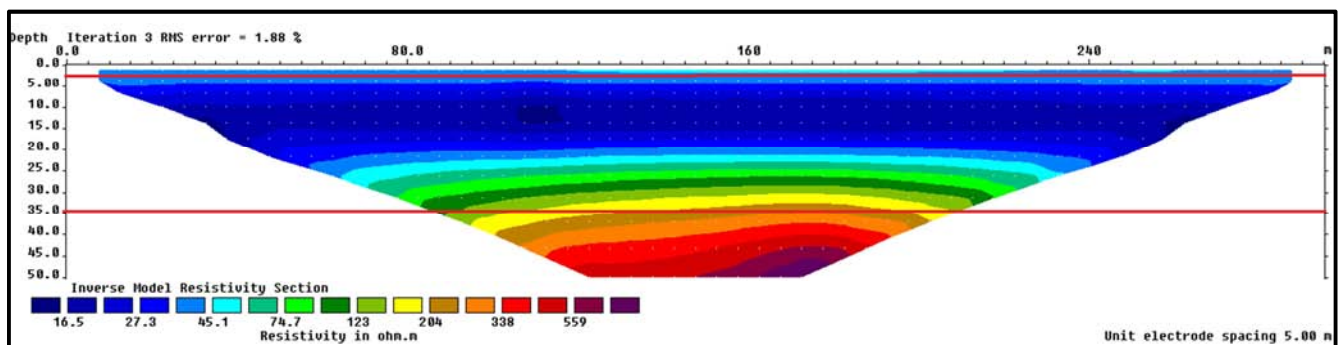


Figure 3. 2D Inverse model resistivity section.

Two-dimensional modelling was performed by using Wenner alpha array oriented perpendicularly to the strike of thin inclined layers (stratification or lamination) inside the middle layer in the 1D anisotropic model. Measured apparent resistivity varies with array orientation to the electrical strike, being minimal when the orientation is perpendicular to it and maximal in the case of parallel array orientation. Therefore one could expect higher values of apparent resistivity with array orientations other than perpendicular and consequently

different results of 2D inversion.

In order to determine “true” resistivity of the middle layer it is possible to perform 1D interpretation by fixing the values of all parameters known from the model (i.e. thicknesses of layers and resistivity of the covering layer and the third layer). Figure 4 shows that the value of root mean square error is of an order defined by the level of random noise added to the theoretically calculated data (RMS Error =2.1%) and that the resistivity of the middle layer is 28.5

Ohmm. It is clearly visible that theoretically calculated resistivity sounding curve (---) deviates the most from the measured data (++) in the zone of curve minimum which corresponds to the middle layer. Figure 5 shows the results of 1D interpretation with lower value of the second layer resistivity (25 Ohmm). Keeping the RMS error value around 2% and not allowing changing of the parameters of the first and the third layer, further lowering of the second layer

resistivity was not possible. Thickness of the second layer was reduced from 32.5 meters to about 27 meters and the depth to substratum from 35 meters to 29.4 meters, which is closer to the results of 2D inversion in Figure 3. The range of equivalence in this case was very narrow since the H-type apparent resistivity sounding curve which corresponds to the 1D model almost reaches the “true” resistivity of the middle layer (the second to the first layer thickness ratio being 13:1).

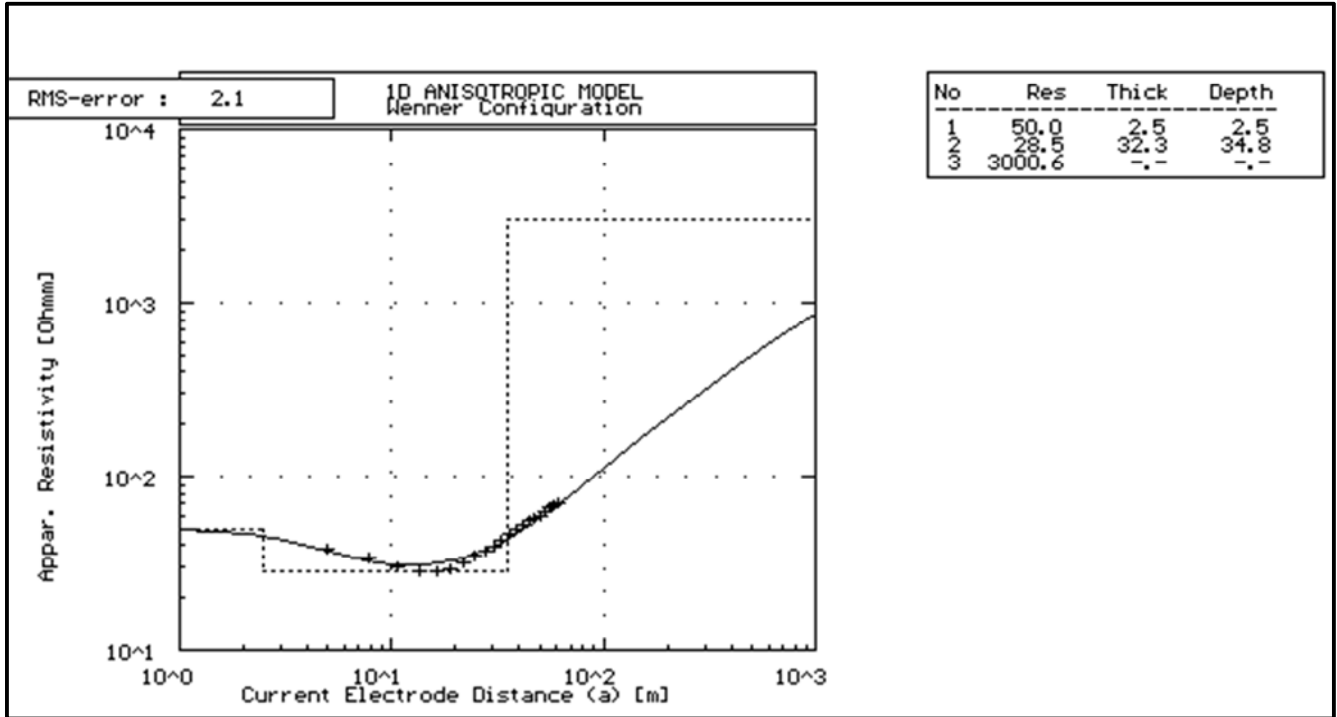


Figure 4. 1D interpretation of Wenner sounding data over the three layer 1D anisotropic model.

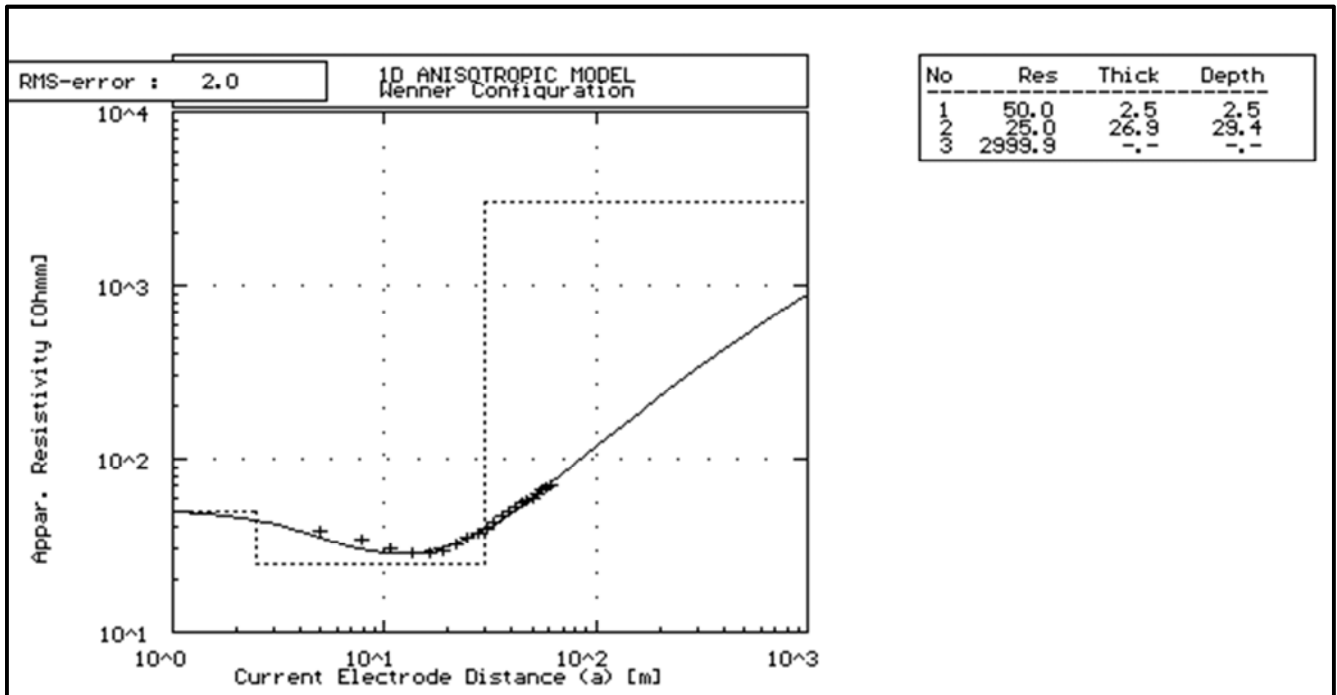


Figure 5. 1D equivalent interpretation of Wenner sounding data over the three layer 1D anisotropic model.

2.2. Two-Dimensional Anisotropic Models

The two-dimensional anisotropic model consists of the covering layer (5m, 50 Ohmm), anisotropic middle layer with the same internal structure as in the case of 1D anisotropic model (resistivity contrast between thin inclined layers amounts 30:1) and high resistivity substratum (3000 Ohmm) in the form of trench structure dipping from 35 to 55 meters in the middle of profile (Figure 6 - lower part). Apparent resistivity cross-section (Figure 6 - upper part) was calculated for Wenner alpha array with 60 electrodes and unit electrode spacing of 5 meters. It reflects the presence of the trench

structure and low resistivity zone (30-35 Ohmm) which corresponds to the middle anisotropic layer. Two-dimensional inversion was carried out by using calculated apparent resistivity data set with $\pm 2\%$ random noise level added to it, so that RMS error of 2.5% is quite satisfactory. The result of the 2D inversion reveals the fact that anisotropic layer is characterized by even lower “true” resistivity value (14-23 Ohmm) and with underestimated depth to substratum of about 15-20%, while general trench-like structure is well defined. Red solid lines in Figure 7 depict the true depths to boundaries in the 2D model.

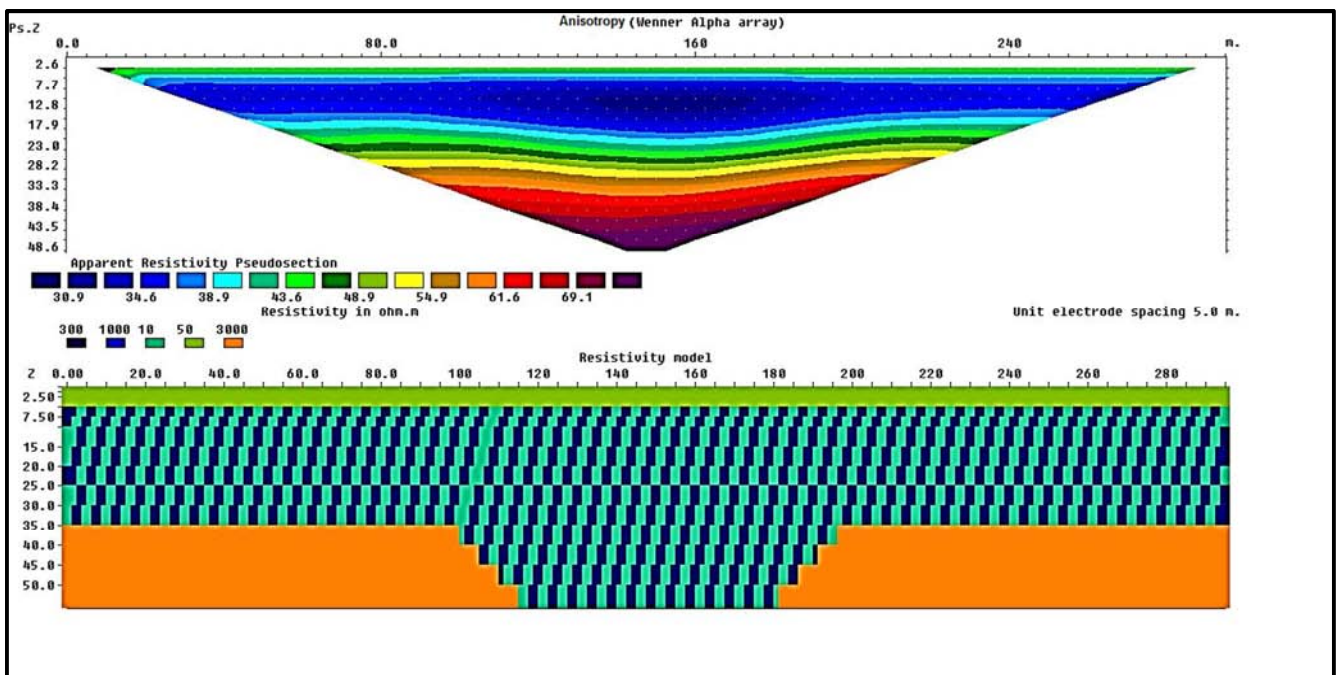


Figure 6. 2D anisotropic model and Wenner apparent resistivity cross-section.

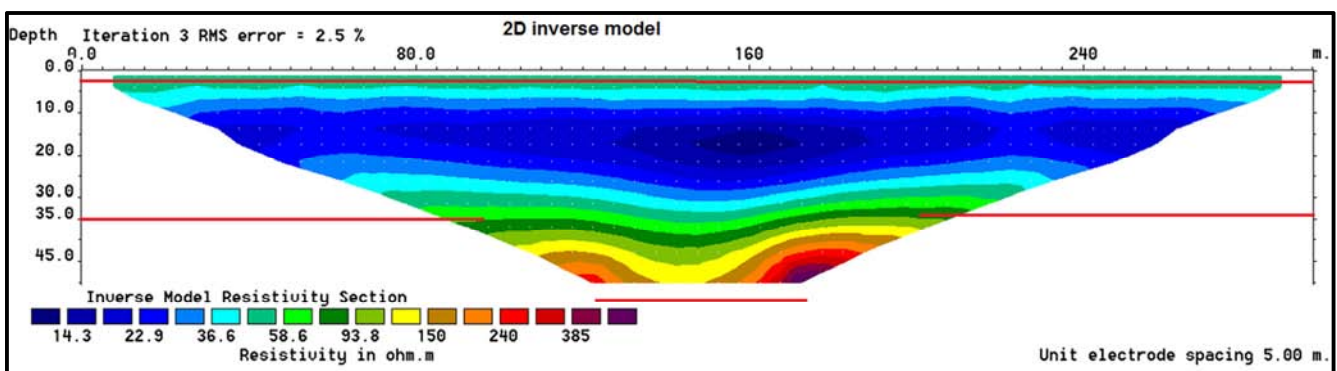


Figure 7. 2D inverse model resistivity section.

In order to investigate 3D forward and inverse modelling ten parallel 2D Wenner array profiles were created with the same anisotropic layer present and by using the same covering layer and substratum parameters. The depth to the substratum and to the bottom of the trench is getting shallower from profile 1 towards profile 10, while the trench structure is getting narrower. Figure 8 shows three out of ten

parallel profiles (Profiles 1, 6 and 10). These ten calculated 2D apparent resistivity data sets with $\pm 2\%$ random noise level added to it were initially inverted independently to give a series of 2D cross-sections, the three of them are given in Figure 9.

“The measured apparent resistivity values from all the lines can also be combined into a 3D data set and inverted with

RES3DINV to give a 3D picture. While the quality of the 3D model is expected to be poorer than that produced with a complete 3D survey, such 3D data set could reveal major resistivity variations across the survey lines. While it is possible to combine a number of 2D lines into a 3D data set, it

might not always be worthwhile to do so. Firstly, it is recommended that there should be at least 5 parallel lines. Furthermore, the separation between the lines should not be more than twice the unit electrode spacing along the lines” [4].

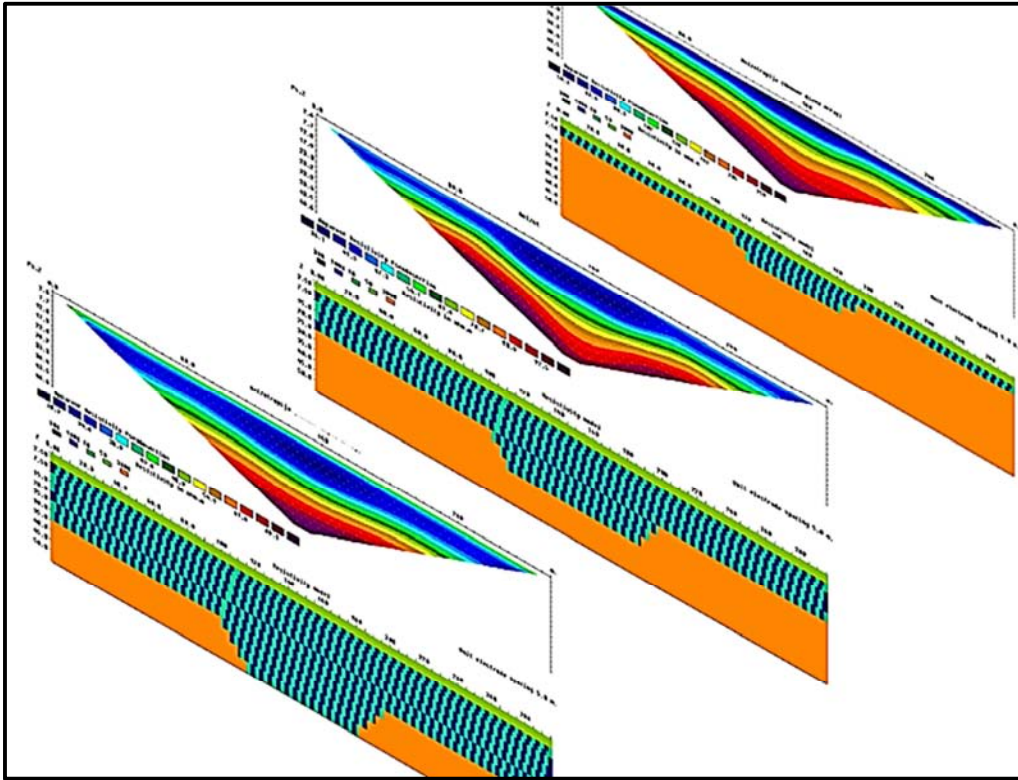


Figure 8. Three parallel 2D profiles together with corresponding Wenner apparent resistivity cross-sections.

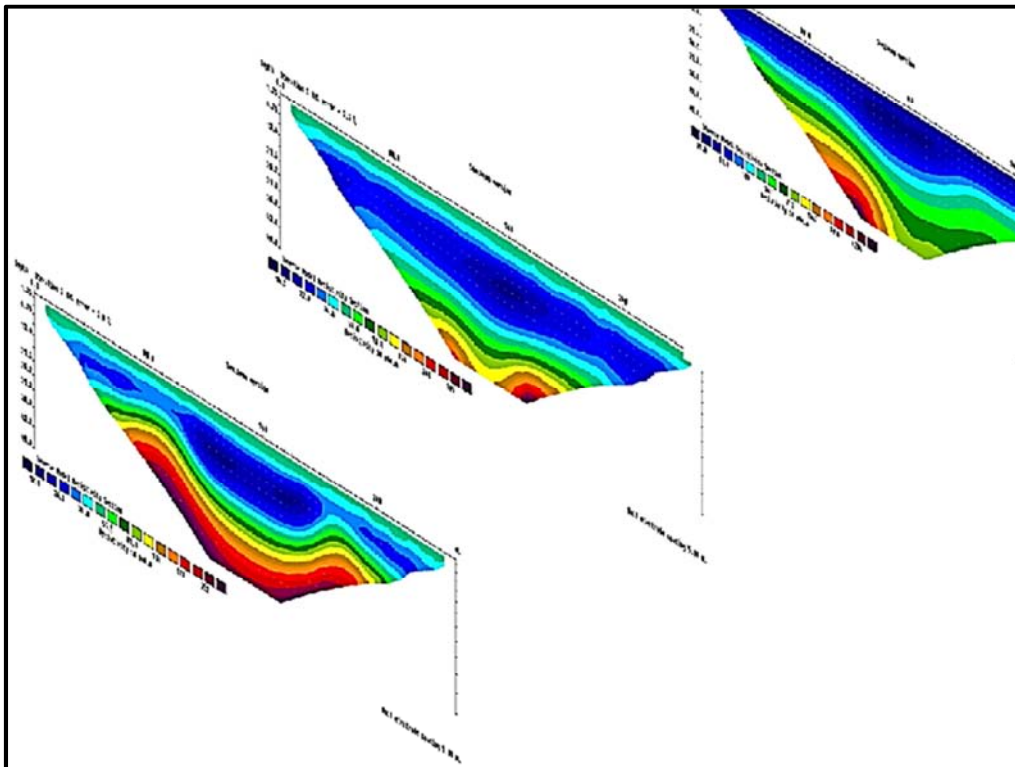


Figure 9. The three out of ten 2D inversion models along parallel profiles.

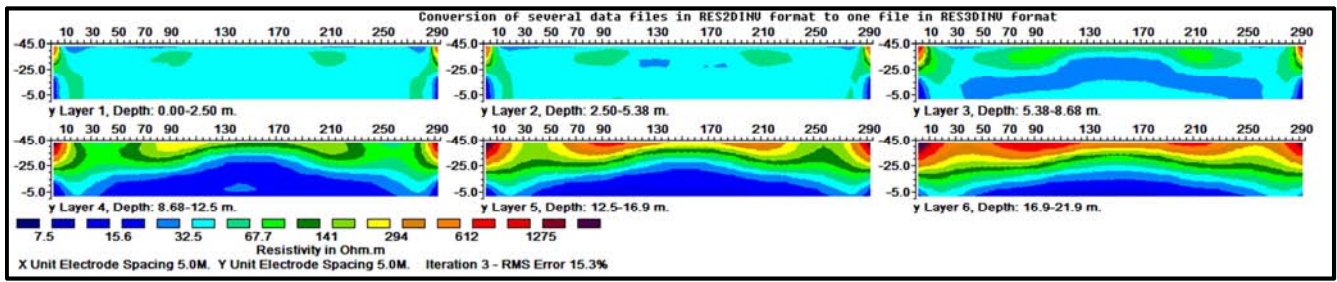


Figure 10. 3D inversion based on the ten inverted 2D apparent resistivity data sets (Y:X axes=2:1).

Three-dimensional inversion based on 2D apparent resistivity data sets along ten parallel profiles (5 meters apart) revealed that the anisotropic layer filling up the 2D trench structure shows very low “true” resistivity values ranging from 7.5 to 15 Ohmm, which is much lower than its values based on 1D and 2D inversions (Figure 10). The covering layer (5meters thick and with resistivity of 50 Ohmm) is well defined as well as the trench structure which is getting wider and deeper from profile 10 towards profile 1. True resistivity of substratum is very well determined (2550 Ohmm compared to about 600 Ohmm in the 2D resistivity inverse model in Figure 7). The results of 3D inversion are presented with six layers with the maximum depth of about 22 meters which is relatively shallow compared to 50 meters in 2D anisotropic model. Root mean square error which is achieved in just three iterations is 15.3%

2.3. Three-dimensional Anisotropic Models

“For relatively small grids of less than 12 by 12 electrodes, the pole-pole array has a substantially larger number of possible independent measurements compared to other arrays. The loss of data points near the sides of the grid is kept to a minimum, and it provides better horizontal data coverage compared to other arrays. This is an attractive array for small survey grids with relatively small spacing (less than

5 meters) between the electrodes. However, it has the disadvantage of requiring two “remote” electrodes that must be placed at a sufficiently large distance from the survey grid. Due to the large distance between the two potential electrodes, this array is more sensitive to telluric noise” [4].

Pole-pole array was used not just for the advantages quoted from [4] but for an additional advantage of Pole-pole array which isn’t mentioned by the author [4]. Namely, resistance measurements conducted by using this array make it possible to calculate the corresponding resistance values of other arrays [9] among which the most interesting for the purposes of this paper is square array.

For 1D anisotropic two-layer model (Figure 11) the apparent resistivity 3D data set was calculated and random noise of ± 2% was added to simulate Pole-pole measured data set. Figures 12 and 13 confirm paradox of anisotropy since apparent resistivity values measured in x-direction, perpendicularly to the strike of vertical thin layers with alternating resistivity values (10 and 300 Ohmm), are of lower magnitude than with Pole-pole array orientation parallel to the electrical strike (y-direction). Both sets of Pole-pole apparent resistivity cross-sections (X and Y orientated) show serious deviation from the 1D anisotropic model, since they indicate shallowing of the lowest resistivity values towards surface.

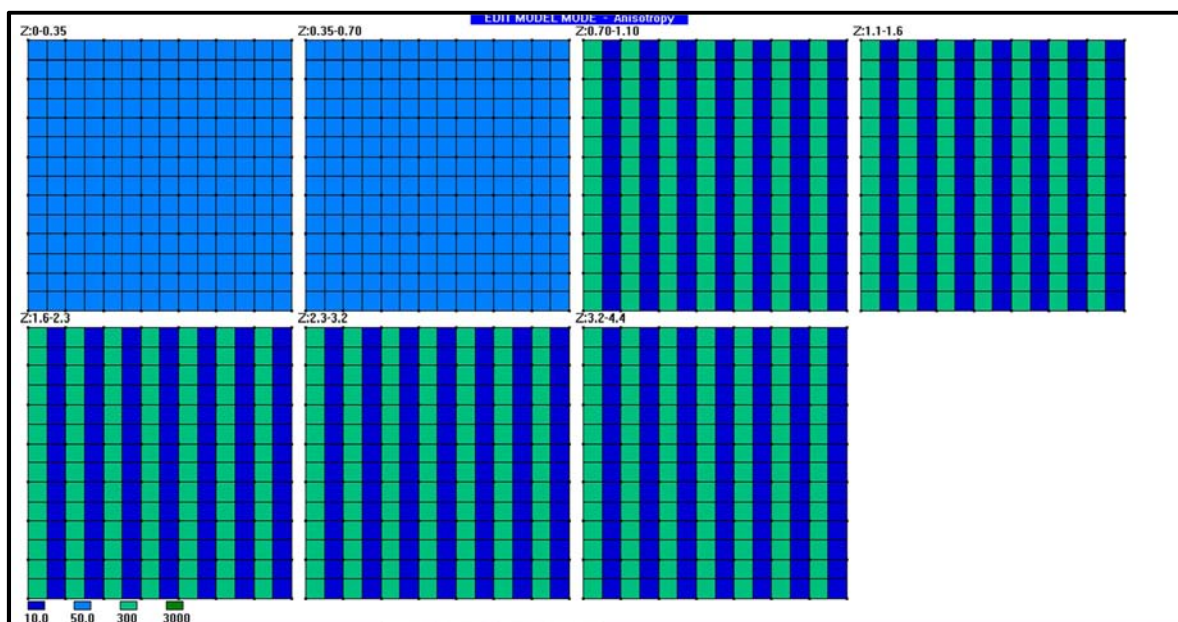


Figure 11. Two layer 1D anisotropic model.

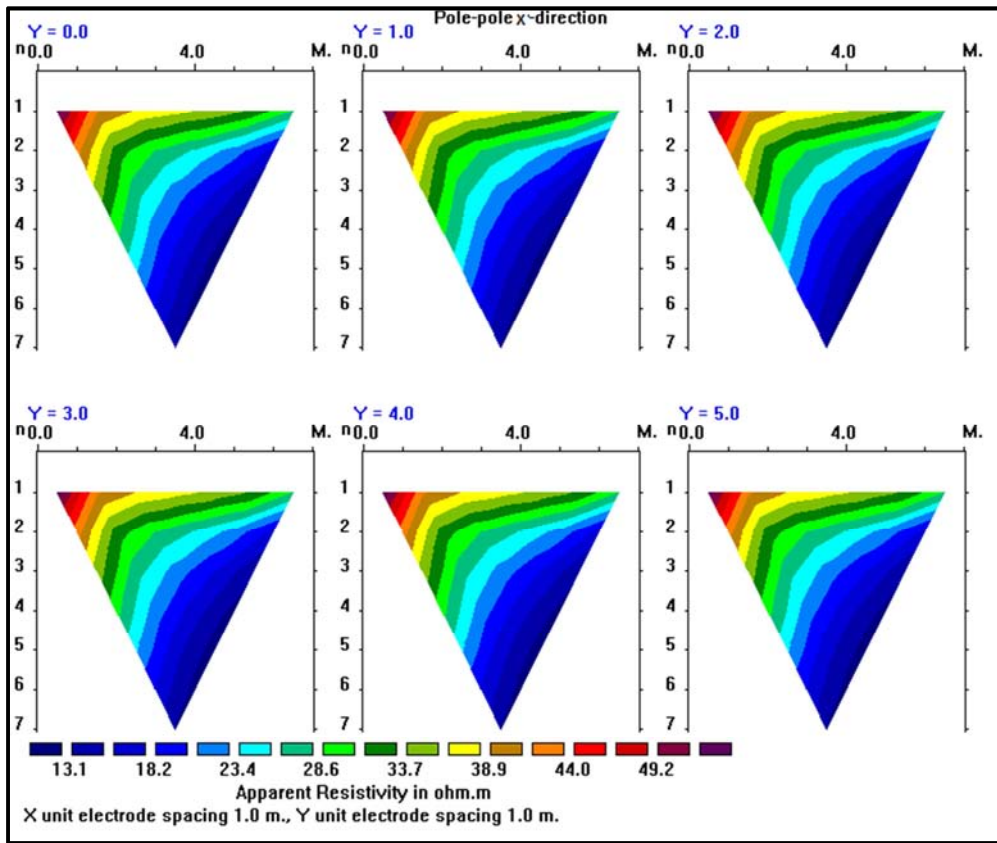


Figure 12. Pole-pole array apparent resistivity cross-sections - X direction.

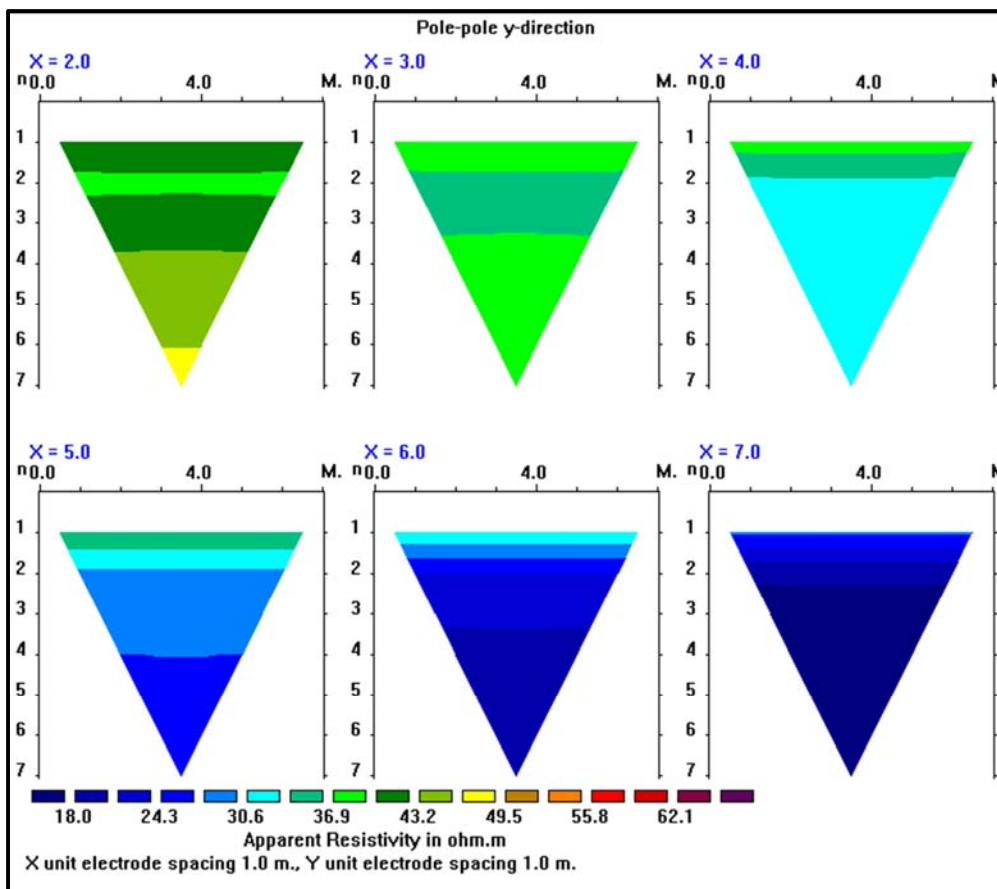


Figure 13. Pole-pole array apparent resistivity cross-sections – Y direction.

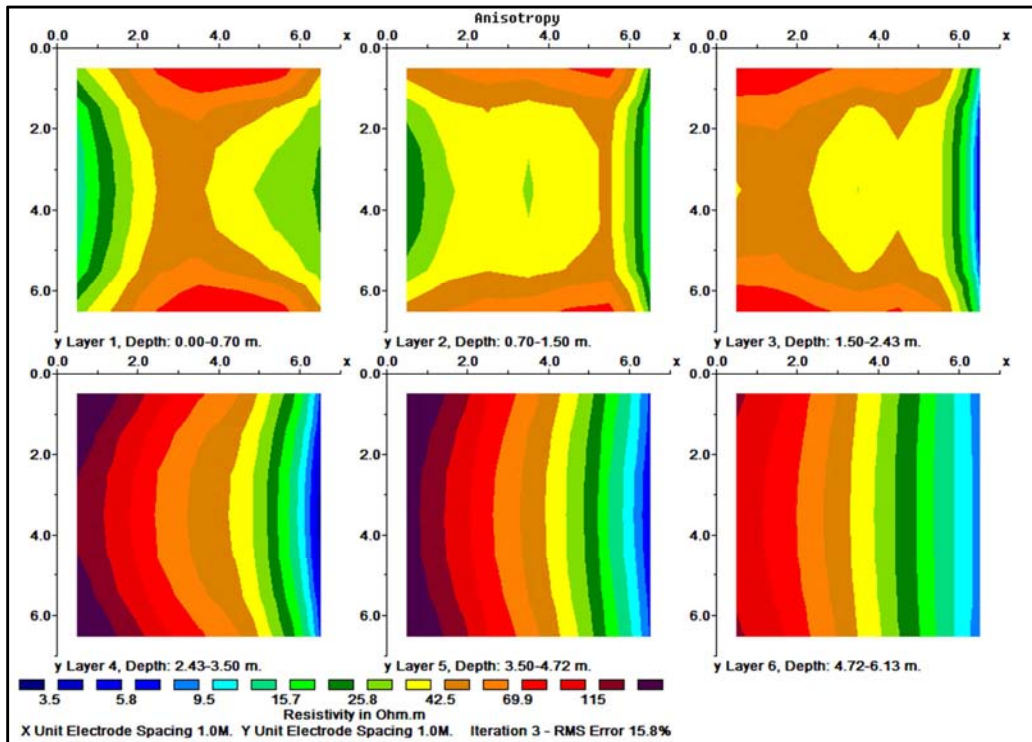


Figure 14. The results of 3D inversion-horizontal sections.

The results of 3D inversion of Pole-pole apparent resistivity complete data set (Figures 14 and 15) are totally erroneous having in mind very simple 1D anisotropic model which was used for this analysis (Figure 11).

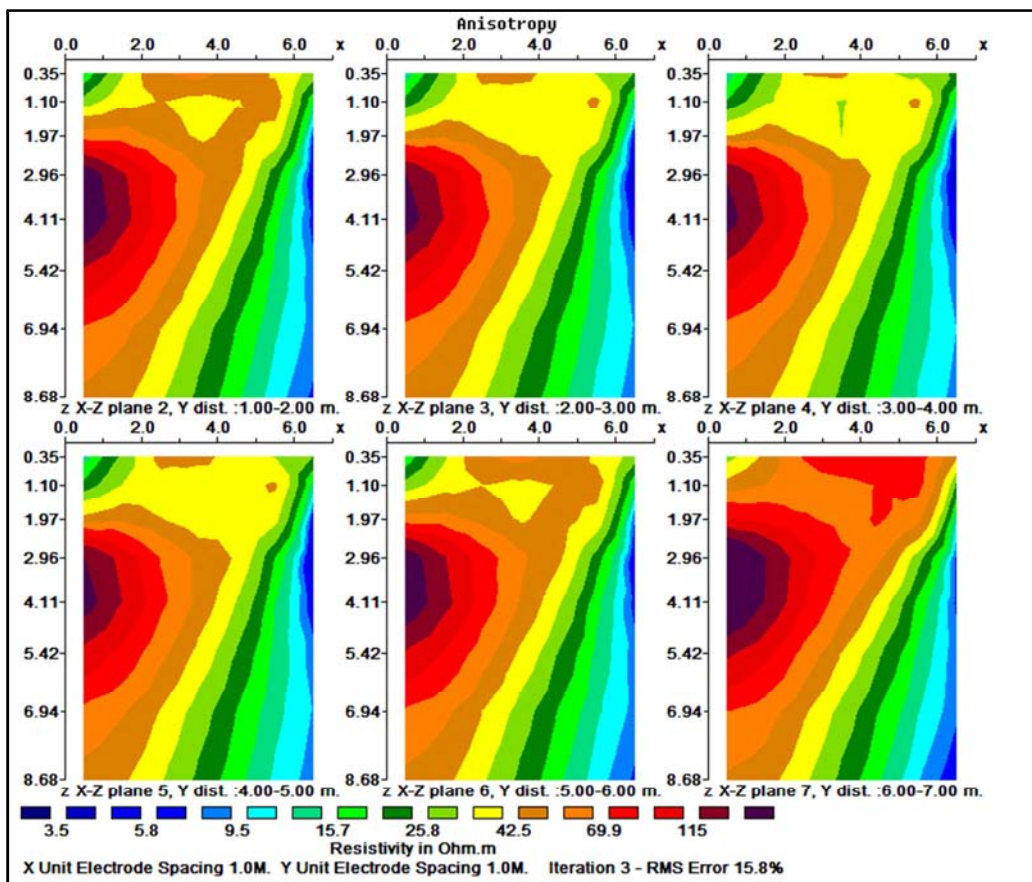


Figure 15. The results of 3D inversion-vertical sections.

3. Reconstructing of Square Array Apparent Resistivity Data from the Complete 3D Pole-Pole Measurements

3.1. Square and Crossed Square Array

Investigation of electrical anisotropy by using square array has many advantages compared to investigation conducted with collinear arrays. It requires less surface area than collinear arrays of similar depth of investigation. This array is less sensitive to lateral effects of local heterogeneities, and much more sensitive to presence of electrical anisotropy than collinear arrays [10]. Also, square array doesn't produce paradox of anisotropy. The location of a measurement is assigned to the central point of the square and the array size (a) is defined by the length of the square side. The array is expanded symmetrically about the center point, so that a resistivity sounding curve can be interpreted as a function of depth [11].

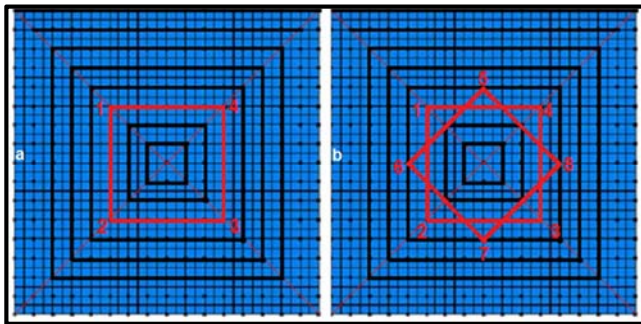


Figure 16. (a) Square array (b) crossed square array.

Since 1D, 2D and 3D inversions of apparent resistivity data sets acquired over an anisotropic half-space lead to erroneous results and give no indication of electrical anisotropy there is a need to reconstruct square array resistivity data from Pole-pole 3D measured apparent resistivity data. Figure 16a illustrates this possibility showing a number of expanding square arrays by using electrodes in a grid (not mesh lines between electrodes). Already existing Pole-pole data in 3D calculated (measured) data set will be used to calculate square array data sets. In order to calculate square array apparent resistivity value (ρ_α) it is possible to use four Pole-pole apparent resistivity measurements (Figure 17a - yellow and white lines) by summing the four values and finding the average value of that sum. The same procedure is applied for calculating (ρ_β) in (Figure 17b). In order to calculate square array resistance values (R_α) and (R_β), (ρ_α) and (ρ_β) are divided by appropriate square array geometric factor, $K = 2\pi a / (2\sqrt{2})$. The same procedure is applied for calculating (R'_α) and (R'_β) for square array rotated by 45° (Figure 16b) and thus one gets the complete input data for crossed square anisotropy analysis proposed by Habberjam [9-10]. Apparent electrical strike (θ), apparent coefficient of electrical anisotropy (n) and mean geometric apparent resistivity (ρ_m) are calculated from electrical

resistance ($R_\alpha, R_\beta, R'_\alpha, R'_\beta$) in two square orientations, under the assumption of homogeneous anisotropic half-space. Accuracy of determination of axes of anisotropy is no longer dependent on number of array directions, but on disturbances caused by heterogeneities which lead to deviation from homogeneous and anisotropic subsurface model. This inevitable geological noise may be quantitatively accounted for, by using parameter (ξ) which defines the deviation from homogeneous anisotropic subsurface model. Since three unknown parameters (n, θ, ρ_m) have to be determined, it is enough to consider values of resistance measured in three directions, and the fourth measurement enables determination of parameter (ξ). Orientation of the system is based on the direction defined by the electrodes 1-4, and the input values are electrical resistances $R_\alpha, R_\beta, R'_\alpha, R'_\beta$. There is, however, a slight difference of about 5% in square array side (a) in two different square array orientation (Figure 16b) so the depth of investigation, could be slightly erroneous. Therefore the procedure for anisotropy analysis based on crossed square resistivity measurements will not be used in this paper.

Figure 18 shows the part of 3D (*.dat) file with coordinates of current electrodes (X_c, Y_c) and potential electrodes (X_p, Y_p) of different Pole-pole configurations which are involved in calculated 3D data set.

The very simple 1D two-layer anisotropic model was used (Figure 11), with the thin covering isotropic layer (0.7 m, 50 Ohm) and anisotropic half-space consisting of thin (0.5m) vertical layers with alternating resistivity values (10 and 300 Ohm). This vertical layering enables determination of anisotropy coefficient (λ) with square array resistivity measurements from the surface. Square array was favourable oriented to the electrical strike, so that (ρ_α) value represents apparent resistivity measured perpendicular to electrical strike and (ρ_β) apparent resistivity measured parallel to the strike of thin vertical layers which cause electrical anisotropy (Figure 17). The four Pole-pole arrays and corresponding apparent resistivities which can be used for calculating the square array apparent resistivity (ρ_α) values are marked in Figure 18 (red rectangles).

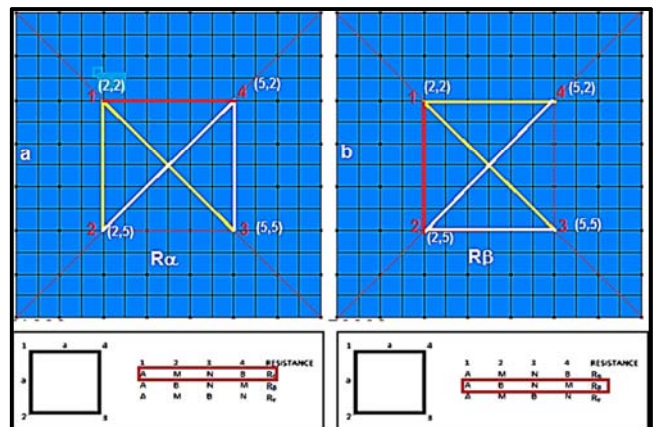


Figure 17. Square array measurements and Pole-pole configurations (a) perpendicular square array orientation (b) parallel square array orientation [9, 10].

Xc	Yc	Xp	Yp	ρ_a	Xc	Yc	Xp	Yp	ρ_a
2.000	2.000	1.000	4.000	40.554	5.000	2.000	2.000	4.000	23.437
2.000	2.000	2.000	4.000	39.239	5.000	2.000	3.000	4.000	24.635
2.000	2.000	3.000	4.000	35.025	5.000	2.000	4.000	4.000	28.028
2.000	2.000	4.000	4.000	28.486	5.000	2.000	5.000	4.000	30.525
2.000	2.000	5.000	4.000	23.294	5.000	2.000	6.000	4.000	23.747
2.000	2.000	6.000	4.000	17.404	5.000	2.000	7.000	4.000	15.312
2.000	2.000	7.000	4.000	12.337	5.000	2.000	0.000	5.000	24.752
2.000	2.000	0.000	5.000	46.893	5.000	2.000	1.000	5.000	24.074
2.000	2.000	1.000	5.000	43.708	5.000	2.000	2.000	5.000	24.070
2.000	2.000	2.000	5.000	42.862	5.000	2.000	3.000	5.000	26.171
2.000	2.000	3.000	5.000	36.986	5.000	2.000	4.000	5.000	27.819
2.000	2.000	4.000	5.000	30.640	5.000	2.000	5.000	5.000	28.531
2.000	2.000	5.000	5.000	23.724	5.000	2.000	5.000	5.000	28.531

Figure 18. The four Pole-pole apparent resistivity values extracted from 3D (*.dat) file later used for (ρ_α) calculation.

Xc	Yc	Xp	Yp	ρ_a	Xc	Yc	Xp	Yp	ρ_a
2.000	2.000	4.000	2.000	29.125	1.000	5.000	5.000	7.000	22.811
2.000	2.000	3.000	2.000	22.521	1.000	5.000	6.000	7.000	17.833
2.000	2.000	6.000	2.000	16.376	1.000	5.000	7.000	7.000	12.135
2.000	2.000	7.000	2.000	11.917	2.000	5.000	3.000	5.000	39.918
2.000	2.000	0.000	3.000	42.031	2.000	5.000	4.000	5.000	28.562
2.000	2.000	1.000	3.000	38.092	2.000	5.000	5.000	5.000	24.074
2.000	2.000	2.000	3.000	43.000	2.000	5.000	5.000	5.000	24.070
2.000	2.000	3.000	3.000	34.451	2.000	5.000	6.000	5.000	16.546
2.000	2.000	4.000	3.000	27.542	2.000	5.000	7.000	5.000	11.826
2.000	2.000	5.000	3.000	22.217	2.000	5.000	0.000	6.000	42.912
2.000	2.000	6.000	3.000	17.565	2.000	5.000	1.000	6.000	37.870
2.000	2.000	7.000	3.000	11.952	2.000	5.000	2.000	6.000	41.121
2.000	2.000	0.000	4.000	45.261	2.000	5.000	3.000	6.000	33.927
2.000	2.000	1.000	4.000	40.554	2.000	5.000	4.000	6.000	25.907
2.000	2.000	2.000	4.000	39.239	2.000	5.000	5.000	6.000	25.907
2.000	2.000	3.000	4.000	35.025	2.000	5.000	6.000	6.000	25.907
2.000	2.000	4.000	4.000	28.486	2.000	5.000	7.000	6.000	25.907
2.000	2.000	5.000	4.000	23.294	2.000	5.000	8.000	6.000	25.907
2.000	2.000	6.000	4.000	17.404	2.000	5.000	9.000	6.000	25.907
2.000	2.000	7.000	4.000	12.337	2.000	5.000	0.000	7.000	23.232
2.000	2.000	0.000	5.000	46.893	2.000	5.000	1.000	7.000	23.232
2.000	2.000	1.000	5.000	43.708	2.000	5.000	2.000	7.000	15.674
2.000	2.000	2.000	5.000	42.862	2.000	5.000	3.000	7.000	15.674
2.000	2.000	3.000	5.000	36.986	2.000	5.000	4.000	7.000	15.674
2.000	2.000	4.000	5.000	30.640	2.000	5.000	5.000	7.000	15.674
2.000	2.000	5.000	5.000	23.724	2.000	5.000	6.000	7.000	15.674
2.000	2.000	6.000	5.000	18.514	2.000	5.000	7.000	7.000	15.674

Figure 19. The four Pole-pole apparent resistivity values extracted from 3D (*.dat) file later used for (ρ_β) calculation..

Table 1. Coefficient of anisotropy (λ) and mean geometric resistivity ρ_m for anisotropic model (Figure 11).

a(m)	ρ_α (Ohmm)	ρ_β (Ohmm)	$\rho_m = \sqrt{\rho_\alpha \times \rho_\beta}$	$\lambda = \sqrt{\rho_\alpha / \rho_\beta}$ (%)
1	35.50	34.83	35.163	1.01
3	29.80	23.00	26.180	1.14
5	28.91	18.23	22.96	1.26
7	28.19	14.24	20.04	1.41

Furthermore it is possible to calculate the coefficient of anisotropy (λ) and mean geometric resistivity (ρ_m). Since the 1D anisotropic model includes the thin covering isotropic layer it affects square array resistivity measurements. With square array side a=3 meters the anisotropy coefficient calculated in this way is $\lambda = \sqrt{\rho_\alpha / \rho_\beta} = 1.138$ (Table 1.). For

By averaging the sum of the four Pole-pole apparent resistivity data, one gets square array apparent resistivity value $\rho_\alpha = 29.8$ Ohmm, with the array orientation perpendicular to the strike of 1D anisotropic model. On the other hand, by averaging the sum of four Pole-pole apparent resistivity data in Figure 19, one gets square array apparent resistivity value $\rho_\beta = 23.0$ Ohmm, with the array orientation parallel to the strike of anisotropic model. It is obvious that there is no paradox of anisotropy for square array.

smaller size of the square array side (a=1m) the anisotropy coefficient is only 1.01, but its value increases with increasing value of (a), taking values of 1.26 (a=5m) and 1.41 (a=7m).

3.2. 3D Forward and Inverse Modelling of the Three Layer 1D Anisotropic Model

The three layer 1D anisotropic model is shown in Figure 20, with 6.4 meters thick middle anisotropic layer (consisting of 0.5 meters thin vertical layers with varying resistivity of 10 and 300 Ohmm, respectively), while covering layer (1.1 meter thick, 50 Ohmm) and the third layer (3000 Ohmm) are homogenous and isotropic.

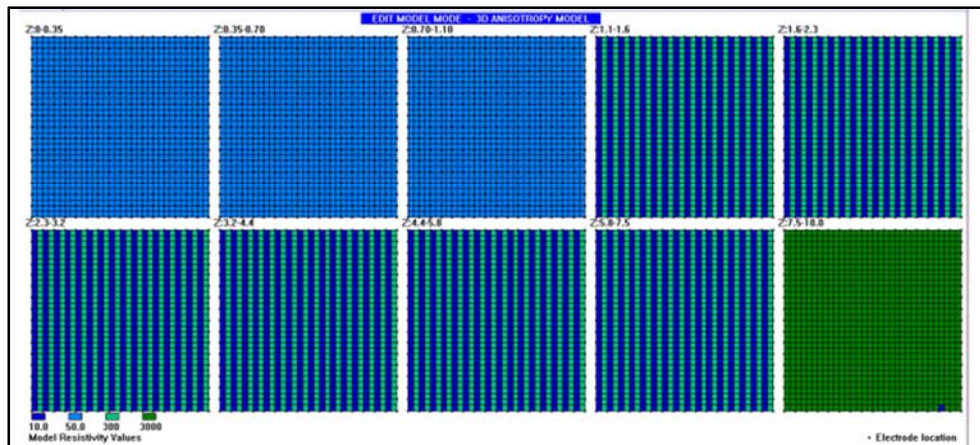


Figure 20. The three layer 1D anisotropic model (17x17 electrode grid).

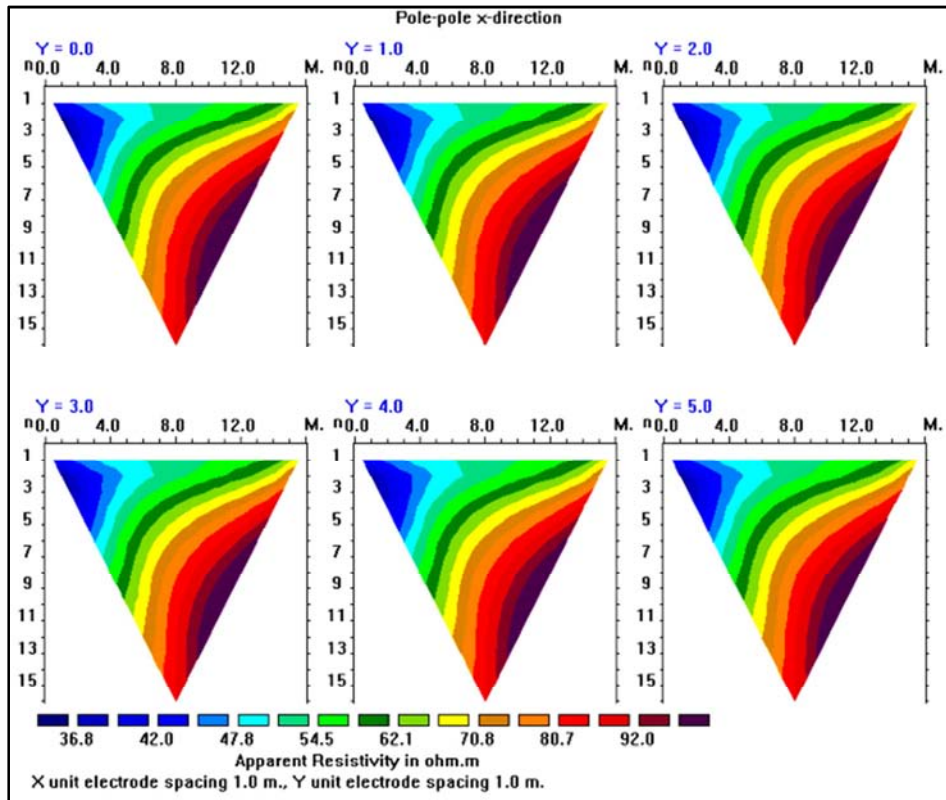


Figure 21. Pole-pole array apparent resistivity cross-sections - X direction.

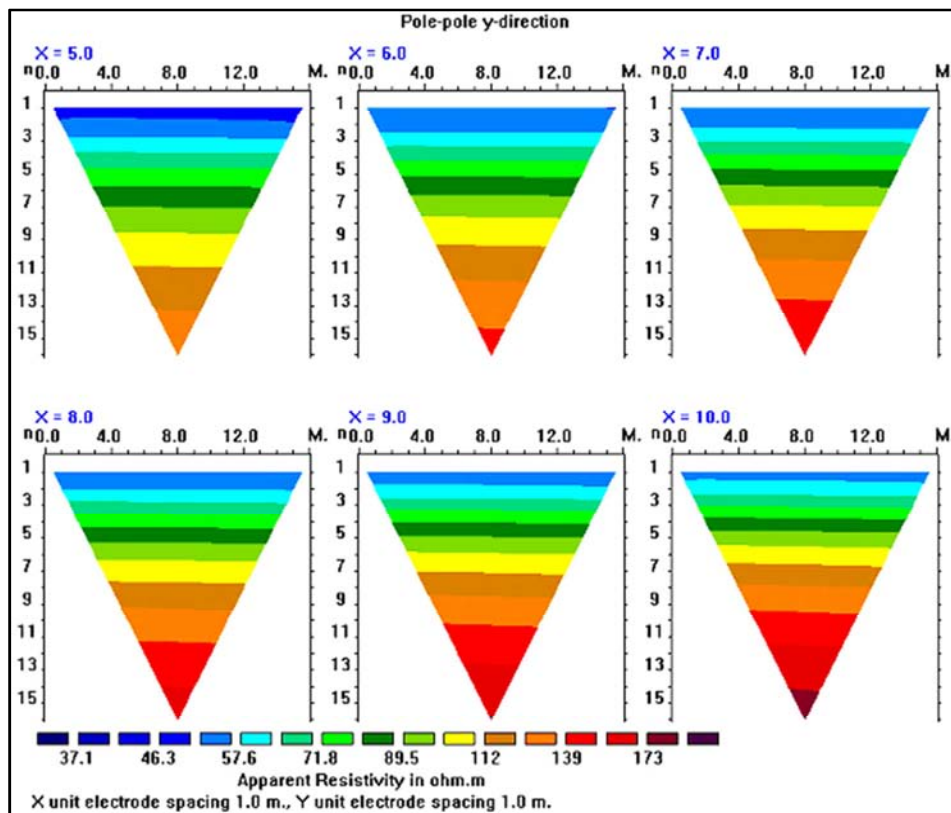


Figure 22. Pole-pole array apparent resistivity cross-sections - Y direction.

It is obvious that apparent resistivity cross-sections are both misleading (Figures 21 and 22), so that the result of 3D inversion is totally erroneous (Figure 23).

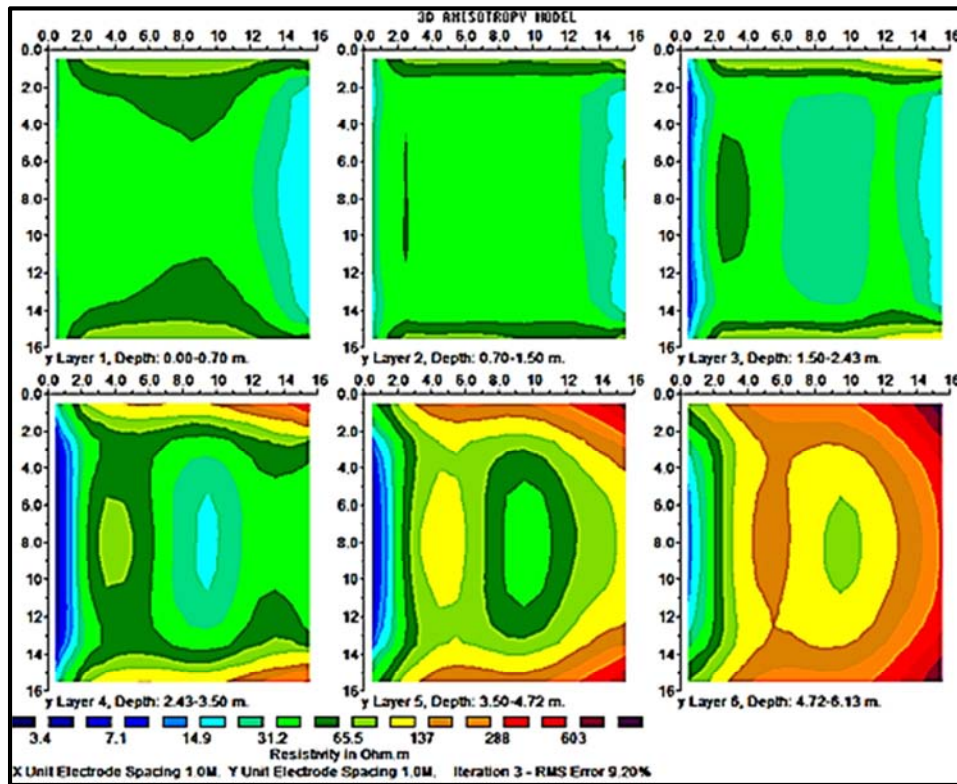


Figure 23. 3D inversion model.-horizontal sections

Square array was favourable oriented to the electrical strike, so that (ρ_{α}) value represents apparent resistivity measured perpendicular and (ρ_{β}) apparent resistivity measured parallel to the strike of thin vertical layers (inside the middle layer in 1D model) which cause electrical

anisotropy. The vertical layering inside the middle layer enables determination of true anisotropy coefficient (λ) values for expanding square array dimensions and thus increasing depth of investigation (Table 2).

Table 2. Coefficient of anisotropy (λ) and mean geometric resistivity ρ_m for anisotropic model (Figure 20).

a(m)	ρ_{α} (Ohmm)	ρ_{β} (Ohmm)	$\rho_m = \sqrt{\rho_{\alpha} \times \rho_{\beta}}$ (Ohmm)	$\lambda = \sqrt{\rho_{\alpha} / \rho_{\beta}}$ (%)
2	56.9	55.4	56.1	1.01
4	76.4	69.4	72.8	1.05
6	95.1	83.0	88.8	1.07
8	108.8	90.4	99.2	1.10
10	117.6	95.9	106.2	1.11
12	127.5	97.6	111.6	1.14
14	133.6	98.9	114.9	1.16
16	139.7	98.2	117.1	1.19

The covering isotropic layer as well as the isotropic substratum (Figure 20) affected measured apparent resistivity (ρ_{α}) and (ρ_{β}) so that (λ) values are reduced compared to the results given in Table 1. The mean geometric apparent resistivity curve (ρ_m), which is orientation- independent, was interpreted by using 1D software with fixing depth to the boundaries in the three layer 1D anisotropic model (Figure 24). This square array resistivity curve was first shifted in accordance with the adequate percentage (33%) to give an equivalent Wenner array (ρ_m) curve [12-13]. Resistivity of the relatively thin covering homogenous and isotropic layer is underestimated (40 Ohmm) due to lack of square array measurements at spacing lesser than 2 meters. On the other hand, maximal spacing of square array for 17x17 electrode grid was 16 meters so that there were not enough square

array data for better defining the resistivity of half-space. Interpreted resistivity value of the middle anisotropic layer (76.5 Ohmm) is much higher than the resistivity of this layer obtained by 2D inversion (16.5-27.5 Ohmm in Figure 3) and by 1D inversion (25-28.5 Ohmm in Figures 4 and 5). This is a consequence of 1D inversion of orientation-independent ρ_m square array sounding curve.

Two-dimensional forward and inverse modelling was performed for Wenner array oriented perpendicularly to the electrical strike in the 1D anisotropic model in Figure 2 and as a result apparent resistivity values were minimal. Therefore interpreted resistivity value of the middle anisotropic layer is also minimal and variable with array orientation. The same is the case with 1D inversion of the Wenner sounding curve which was reconstructed from 2D Wenner cross-section.

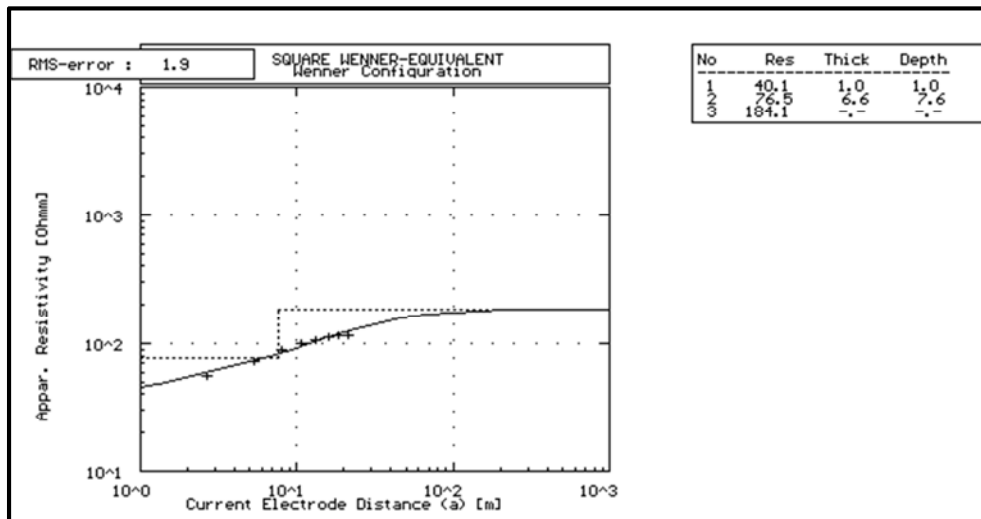


Figure 24. One dimensional interpretation of the mean geometric (ρ_m) curve.

The square array mean geometric apparent resistivity curve (ρ_m), which is orientation-independent, is therefore the best choice for getting reliable results of 1D inversion.

4. Conclusions

This study is based on a number of synthetic 1D and 2D anisotropic models which were used for computing the corresponding Wenner and Pole-pole array apparent resistivity cross-sections by using forward modelling. Random noise of $\pm 2\%$ was added to the each computed apparent resistivity data set in order to simulate measured apparent resistivity data. Apparent resistivity data which were obtained in this way served as an input for 1D, 2D and 3D inversions in order to get the idea about the effects of electrical anisotropy on the results of these inversions.

The three layer 1D anisotropic model was first considered. The relatively thick middle layer characterized by an internal oblique stratification with high resistivity contrast of 30:1 between alternating thin oblique layers was anisotropic while covering layer and substratum were homogeneous and isotropic. The result of 2D inversion of the Wenner apparent resistivity data corresponding to the 1D three layer anisotropic model showed that the interpreted depth to the substratum was underestimated by at least 15-20%, and that the middle layer appears to be isotropic with relatively low “true” resistivity (16.5-27.5 Ohmm). This leads to completely wrong geophysical as well as geological interpretation of 2D inverse model. Further problem in interpreting apparent resistivity data acquired with any collinear array arises from the fact that these data are orientation-dependent in the presence of electrical anisotropy.

Since 2D interpretation is somewhat limited by the fact that there are no sharp boundaries between layers it was necessary to use 1D resistivity sounding data inversion. The 1D anisotropic model was further examined by 1D interpretation of Wenner apparent resistivity sounding curve reconstructed from Wenner array 2D cross-section. In order to better estimate the “true” resistivity of the middle

anisotropic layer this 1D inversion was performed by fixing all known parameters of the model, namely thicknesses of layers and resistivities of covering layer and half-space. The “true” resistivity of the middle layer determined by this 1D interpretation was 28.5 Ohmm. The range of equivalence was also examined by keeping the RMS error value around 2% and by allowing changing of the middle layer resistivity, the first layer parameters and the resistivity of half-space being fixed. Further lowering of the resistivity of the middle anisotropic layer was possible to the value of 25 Ohmm. As a consequence the thickness of the middle anisotropic layer was decreased from 32.5 to about 27 meters and the depth to substratum from 35 to 29.4 meters, which is closer to the results of 2D inversion.

The 2D anisotropic model was composed of the same three layers as in 1D anisotropic model, the middle layer had the same internal structure consisting of thin alternating oblique layers which cause electrical anisotropy. In addition, this model is characterized by the trench structure in the basement which can cause structural anisotropy as well. The results of 2D inversion showed underestimated depth to substratum (of about 15-20%) as well as a relatively low resistivity of the middle anisotropic layer (14-23 Ohmm) which appears to be a homogenous and isotropic layer. This can lead to wrong geophysical and geological interpretation. On the other hand the geometry of the trench-like structure was pretty well defined.

In order to investigate 3D forward and inverse modelling ten parallel 2D Wenner array profiles were created with the same middle anisotropic layer present in 2D models and with the same covering layer and basement parameters. The depth to the basement and to the bottom of the trench is getting shallower from profile 1 towards profile 10, while the trench structure is getting narrower in the same direction. Three-dimensional inversion based on 2D inversion of apparent resistivity data sets along ten parallel profiles (5 meters apart) revealed that the anisotropic layer filling up the 2D trench structure shows very low “true” resistivity values (7.5 to 15 Ohmm), which is much lower than its values based on 1D

and 2D inversions. The covering layer (5 meters thick and with resistivity of 50 Ohmm) is generally well defined as well as the trench structure which gets wider and deeper from profile 10 towards profile 1.

For 1D anisotropic two-layer model, consisting of the thin (0.7m) homogenous and isotropic (50 Ohmm) covering layer and anisotropic half-space with vertical thin layers (0.5m) with alternating resistivity values (10 and 300 Ohmm), the Pole-pole apparent resistivity 3D data set was calculated and random noise of $\pm 2\%$ was added to simulate Pole-pole measured data set. The results of 3D inversion of the complete Pole-pole data sets in the case of this simple 1D anisotropic model showed totally erroneous results regarding geometry of the 3D inverse model and resistivity distribution in the 3D inverse model.

The same was the case regarding the three layer 1D anisotropic model which consists of 1.1 meter thick homogenous and isotropic covering layer (50 Ohmm), 6.4 meters thick anisotropic middle layer consisting of vertical thin (0.5 m) layers with alternating resistivity values (10 and 300 Ohmm) and homogenous and isotropic half-space (3000 Ohmm). In order to get some information on electrical anisotropy which caused totally erroneous 3D inverse models even for the simple 1D anisotropic models it is possible to use the complete 3D measured Pole-pole array data sets. This kind of 3D measurements enable the reconstruction of square array data sets which can further be used for electrical anisotropy analysis and for getting orientation-independent apparent resistivity sounding curves (ρ_m). Square array (ρ_α) and (ρ_β) curves were calculated by using the four adequate Pole-pole array resistivity values. Square array was favorable oriented to the strike of vertical thin layers with alternating resistivity values (10 and 300 Ohmm), so that (ρ_α) value represents apparent resistivity measured perpendicular to the electrical strike and (ρ_β) apparent resistivity measured parallel to the electrical strike. For increasing square array spacing the values of true anisotropy coefficients (λ) were determined ranging from 1.01 to 1.19 (three layer 1D anisotropic model) and from 1.01 to 1.41 (two layer 1D anisotropic model).

Finally, the square array mean geometric apparent resistivity curve ($\rho_m = \sqrt{\rho_\alpha \times \rho_\beta}$), which is orientation-independent, was interpreted by using 1D inversion with fixing depth to the boundaries in the three layer 1D anisotropic model. In order to interpret this square array mean geometric resistivity curve by using 1D standard software (for Wenner array) it was first necessary to shift it for the adequate percentage (33%) to give an equivalent Wenner array mean geometric resistivity curve. The result of 1D inversion shows the the middle anisotropic layer is characterized by much higher “true” resistivity (76.5 Ohmm) compared to the resistivity of this layer obtained by 2D inversion of Wenner apparent resistivity cross-section (16.5-27.5 Ohmm) and 1D inversion of corresponding Wenner sounding curve (25-28.5 Ohmm). Two-dimensional forward and inverse modelling in this case was performed by using Wenner array oriented perpendicularly to the electrical strike,

apparent resistivity values being minimal and therefore interpreted “true” resistivity value of the middle anisotropic layer is also minimal and variable with array orientation. The square array mean geometric resistivity curve (ρ_m), which is orientation-independent, is therefore the best choice for getting reliable results of 1D inversion but also for qualitative interpretation by using mean geometric resistivity cross-sections (ρ_m), as well as (ρ_m) maps and 3D blocks of the mean geometric resistivity distribution.

References

- [1] N. Inyang Udosen. N. Jimmy George, 2018, Characterization of electrical anisotropy in North Yorkshire, England using square arrays and electrical resistivity tomography *Geomech. Geophys. Geo-energ. Geo-resour.* July 2018 DOI: 10.1007/s40948-018-0087-5.
- [2] Habberjam, G. M., Apparent resistivity, anisotropy and strike measurements. *Geophysical Prospecting*, 23, p. 211-247, 1975.
- [3] Bhattacharya B. B., Patra H. P. Direct methods in geoelectric sounding: Principle and interpretation. Elsevier Science Publ. Co. Inc., 1968, 131 p.
- [4] Loke, M. H., Tutorial: 2-D and 3-D electrical imaging surveys 1996-2018, 2018.
- [5] Loke, M. H., Rapid 2-D Resistivity & IP inversion using the least-squares method, 2008.
- [6] Loke, M. H., Rapid 3D Resistivity & IP inversion using the least-squares method, 2008.
- [7] Bart Vander Velpen, RESIST. A computer program to process resistivity sounding data on PC Compatibles, *Computers & Geosciences*, 19(5):691-703 · May 1993, DOI: 10.1016/0098-3004(93)90102-B.
- [8] Koefoed. O., *Geosounding Principles: Resistivity Sounding Measurements (METHODS IN GEOCHEMISTRY AND GEOPHYSICS)*, Elsevier Scientific Publishing Company, 1979.
- [9] Habberjam G. M., Apparent resistivity observations and the use of square array techniques. (*Geoexploration Monographs*, Number 9), 1979. ISBN 978-3-443-13013-8.
- [10] Habberjam, G. M., The effects of anisotropy on square array resistivity measurements. *Geophysical Prospecting*, 20, p. 249-266, 1972.
- [11] Habberjam, G. M., and Watkins, G. E. The use of a square configuration in resistivity prospecting. *Geophysical Prospecting*, 15, p. 445-467, 1967.
- [12] Lane, J. W., Jr., Haeni, F. P., and Watson, W. M., 1995, Use of square-array direct-current resistivity methods to detect fractures in crystalline bedrock in New Hampshire: *Groundwater*, v. 33, no. 3, p. 476-485.
- [13] Geological Survey (USGS), Nye County Nuclear Waste Repository Project, 2005, Square-Array Direct-Current Resistivity Measurements Conducted at Nye County near Borehole NC-EWDP-29P.