

Original Research

**COMPARISON OF SATELLITE DATA WITH LOW-COST SENSOR NETWORK
PM DATA IN FOUR CITIES IN SERBIA**

DOI: 10.5937/EcoTER26210C

**Željko Ćirović^{1,a*}, Duška Kleut^{1,b}, Danka B. Stojanović^{1,c},
Milena Jovašević-Stojanović^{1,d}**

¹ Vinča Institute of Nuclear Sciences-National Institute of the Republic of Serbia,
University of Belgrade, Mike Petrovića Alasa 12-14, Belgrade, SERBIA

^a <https://orcid.org/0000-0002-8069-0982> ^b <https://orcid.org/0000-0002-5881-6497>

^c <https://orcid.org/0000-0002-7178-1970> ^d <https://orcid.org/0000-0003-0765-6603>

Abstract

Conventional monitoring stations provide reliable measurements of particulate matter (PM) mass concentrations and are valuable for giving information about air quality but due to high costs and operational challenges, only a limited number of such monitors can be deployed. On the other hand, a network of low-cost PM sensors is also a promising approach to increase the measurement density. In addition to the ground-based measurement techniques described earlier, remote sensing can also be used to derive PM concentrations. In this study, we propose an integration of information from multiple measurement approaches. Here, we demonstrate correlations between results obtained by the data from the monitoring stations, low-cost sensors and satellite S-MESH data in western Serbia in the period from April to December 2022. The results indicate good correlations of all low-cost sensor data with S-MESH data with mean correlations around (>70%). Within this study, it is confirmed that the low-cost sensor network and satellite data provide valuable information on PM concentrations at locations which do not have monitoring stations.

Keywords: particulate matter, pollution mapping, satellite data, low-cost sensor network.

INTRODUCTION

Air pollution, especially from fine particulate matter smaller than 2.5 micrometers (PM_{2.5}), is a significant public health threat and contributes to millions of premature deaths each year. A large number of epidemiological research has consistently linked high PM_{2.5} levels to serious health issues such as cardiovascular and respiratory diseases [1]. In Europe, air quality management standards are in place to combat elevated PM_{2.5} concentrations. Nevertheless, 94% of the EU's urban population remains exposed to harmful PM_{2.5} levels [2], underscoring the urgent need for accurate, high-resolution monitoring to support effective interventions and evidence-based policymaking.

Traditionally, ground-level PM_{2.5} has been measured using air quality monitoring stations with reference-grade instruments, which yield high-quality local data. However, due to their high installation and maintenance costs, these stations are sparse and unevenly distributed,

* zeljko.cirovic@vin.bg.ac.rs

resulting in limited spatial coverage [3]. They often fail to capture inter-urban pollution gradients [4,5]. To address these gaps, various modelling and satellite-based estimation methods have been developed. Yet, each approach has inherent limitations, particularly regarding temporal resolution, spatial representativeness, and accuracy under changing atmospheric conditions.

Satellite-based instruments present an alternative approach by delivering global observations of Aerosol Optical Depth (AOD), which can act as a stand-in for ground-level $\text{PM}_{2.5}$ concentrations [6]. Sensors like the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Visible Infrared Imaging Radiometer Suite (VIIRS) have been extensively employed to estimate surface $\text{PM}_{2.5}$ using various methods [7]. The S-MESH framework [8] produces high-resolution, bias-corrected daily $\text{PM}_{2.5}$ hindcast data for 2021–2022 by downscaling CAMS regional forecasts. It integrates satellite AOD (from VIIRS), CAMS-based $\text{PM}_{2.5}$ predictions, meteorological variables, and ground station data (used solely for training) within a stacked machine learning model. In recent years, low-cost sensors have gained traction due to their ability to fill spatial gaps in existing monitoring networks. These devices are small, affordable, and capable of generating highly resolved temporal data, making them especially appealing for urban-scale studies and citizen science initiatives [9–11]. Remote sensing can also help evaluation of the reliability of low-cost sensors [12–14].

In this study, we use the results of a low-cost sensor network that included 12 locations in 4 municipalities in western Serbia in the period from May to December 2022 [15]. We evaluate the correlations between results obtained by the data from the monitoring stations, low-cost sensors and satellite S-MESH data in order to provide more information about air pollution levels. The methods used here are especially important in municipalities in which monitoring of PM within local networks does not exist.

MATERIALS AND METHODS

Low-Cost Sensor Network and Monitoring stations

The low-cost sensor network data was obtained as part of the UNICEF Serbia pilot initiative "Schools for Better Air Quality". Each low-cost sensor was equipped with a PMS7003 air dust sensor to measure particulate matter fractions. The total low-cost sensor network consisted of 129 devices at 43 locations in 15 municipalities in central and western Serbia [15]. Within the study presented here, research was done for four cities including, Čačak, Kosjerić, Novi Pazar and Užice (12 locations in total). These cities are equipped with Serbian Environmental agency monitoring stations [16] and results obtained with low-cost sensors and monitoring stations are compared in our previous publication [15]. Here, we analysed the data set acquired in the period from May to December 2022.

S-MESH approach

The S-MESH framework [8] produces high-resolution, bias-corrected daily hindcasts of $\text{PM}_{2.5}$ for 2021–2022 by downscaling CAMS regional forecasts. It employs a stacked machine learning model that integrates satellite AOD (from VIIRS), CAMS $\text{PM}_{2.5}$ forecasts, meteorological data, and station observations. The CAMS forecast serves as a predictor to enhance accuracy and fill gaps in AOD data. Rather than aiming to generate real-time

forecasts, S-MESH offers a timely alternative to conventional CAMS reanalysis data. The S-MESH framework used in this study follows a similar structure: a stacked ensemble based on the XGBoost algorithm [17]. It consists of two first-stage beta models, each with slightly different input feature sets, and a second-stage meta model that combines the predictions from the beta models with selected additional predictors.

Data analysis methods

The correlations, corresponding scatter plots and monthly graphs of the $PM_{2.5}$ data obtained from the campaign, from low-cost sensors and monitoring stations, as well as from the S-MESH framework were calculated using *openair* package in R-studio and the plots were created using this tool built for the purpose of analysing air quality data. Land use data shown in Table 1. was derived from CORINE Land Cover 2018 [18] (vector/raster 100 m, Europe, 6-yearly) data which provides pan-European CORINE Land Cover inventory for 44 thematic classes for the 2018 reference year. The dataset has a Minimum Mapping Unit (MMU) of 25 hectares (ha) for areal phenomena and a Minimum Mapping Width (MMW) of 100 m for linear phenomena and is available as vector and as 100 m raster data.

RESULTS AND DISCUSSION

The results presented in Table 1 show a comprehensive overview of the data obtained during the campaign at four locations in each municipality, and their comparison with the data gathered from the S-MESH model database. Figure 1 presents scatter plots and correlation equations for the daily campaign measurements and the S-MESH framework data in the city of Novi Pazar (results for other cities can be seen in the table 1). These were calculated at the GPS coordinates of each of the four locations (one monitoring station and three schools) in the four municipalities, covering the entire campaign period (May–December 2022).

The low-cost sensor network data and the reference stations set in the same municipality show high correlation of $>93\%$, except for one school set in a remote village 13 km away from the reference station location in the town centre (Figure 2). The diversity of the distances and the urban and geographical landscape of the four municipalities indicated the need to examine the relationship between all $PM_{2.5}$ data and the distances of the schools and the reference stations, hence the inclusion of that column in Table 1. As a measurement of the S-MESH estimation for the municipality the correlations of all 4 locations were estimated comparing to the S-MESH data at the reference station location as well as at their own GPS coordinates. In some instances, a few locations were part of the same S-MESH datapoint since the resolution of the data is 1 km (Figure 2). In all municipalities the campaign data shows high correlation with the S-MESH data ($>70\%$) around the central locations with larger concentration of pollution sources. As the distance of the school from the reference station increases the S-MESH data at that location correlate less with the campaign data by 5–10%. Two cases of interest are the ones with unique land use characteristics. First, the remote school in Kosjerić municipality shows an overall low correlation of 25–30% with both the reference data and the S-MESH data. Second, in the school in Novi Pazar municipality, the correlation with its own S-MESH location is 20% lower compared to the correlation between the reference station and its corresponding S-MESH data. This suggests a strong influence

from factors likely not included in the S-MESH framework, probably due to industrial facilities in the vicinity of the latter. Highest overall correlation of 84% of the S-MESH framework data with the campaign measurements was measured at the Novi Pazar reference station location, while the remote school in Kosjerić shows overall lowest correlation to all data.

Table 1 The correlations of the campaign measurements with the reference station and the distance from it, the correlation with the reference station location S-MESH data, the school's location own S-MESH data and the land use data for that particular (own) location

Location	SEPA correlation	SEPA Distance	SEPA-SMESH correlation	Own SMESH correlation	Coordinates	Land use
Cacak-SEPA	1	0.00	0.718	0.73	43.89333, 20.34553	Sport and leisure facilities
Cacak-CHSch	0.945	1.54 km	0.735	0.70	43.88243, 20.33253	Discontinuous urban traffic
Cacak-ElSchSS	0.962	1.50 km	0.748	0.75	43.86786, 20.36911	Discontinuous urban traffic
Cacak-ElSchTR	0.959	3.31 km	0.725	0.65	43.90631, 20.35658	Discontinuous urban traffic
Kosjeric-SEPA	1	0.00	0.737	0.74	44.00294, 19.90518	Discontinuous urban traffic
Kosjeric-ElSchJD	0.262	12.9km	0.301	0.25	44.04410, 19.75163	Land principally occupied by vegetation, with significant areas of natural vegetation
Kosjeric-ElSchMI	0.977	0.87km	0.742	0.71	43.99370, 19.90793	Discontinuous urban traffic
Kosjeric-TechSch	0.976	1.08km	0.745	0.72	43.99262, 19.91031	Discontinuous urban traffic
NoviPazar-SEPA	1	0.00	0.841	0.84	43.13928, 20.51617	Discontinuous urban traffic
NoviPazar-ElSchHZ	0.936	4.75km	0.843	0.76	43.14105, 20.45814	Discontinuous urban traffic
NoviPazar-ElSchMS	0.977	2.47 km	0.843	0.66	43.15987, 20.52680	Complex cultivation patterns
NoviPazar-ElSchRB	0.982	0.87km	0.826	0.83	43.13397, 20.51742	Discontinuous urban traffic
Uzice-SEPA	1	0.00	0.764	0.76	43.85427, 19.84321	Discontinuous urban traffic
Uzice-ElSchNI	0.973	1.05 km	0.755	0.65	43.85404, 19.85919	Discontinuous urban traffic
Uzice-MedSch	0.974	1.00 km	0.753	0.65	43.86037, 19.85493	Discontinuous urban traffic
Uzice-TechSch	0.972	0.175km	0.731	0.73	43.85581, 19.84585	Discontinuous urban traffic

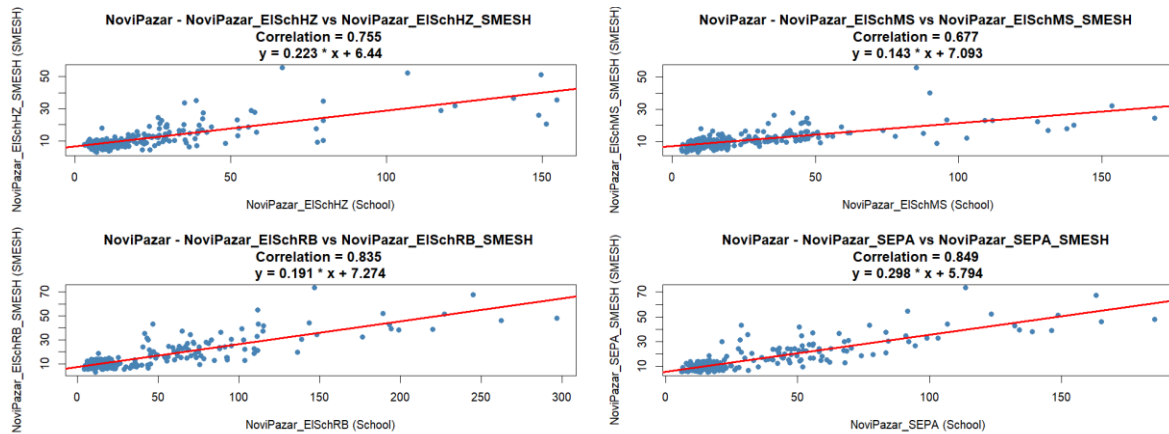


Figure 1 Scatter plots of daily campaign measurements compared to S-MESH data at locations in Novi Pazar

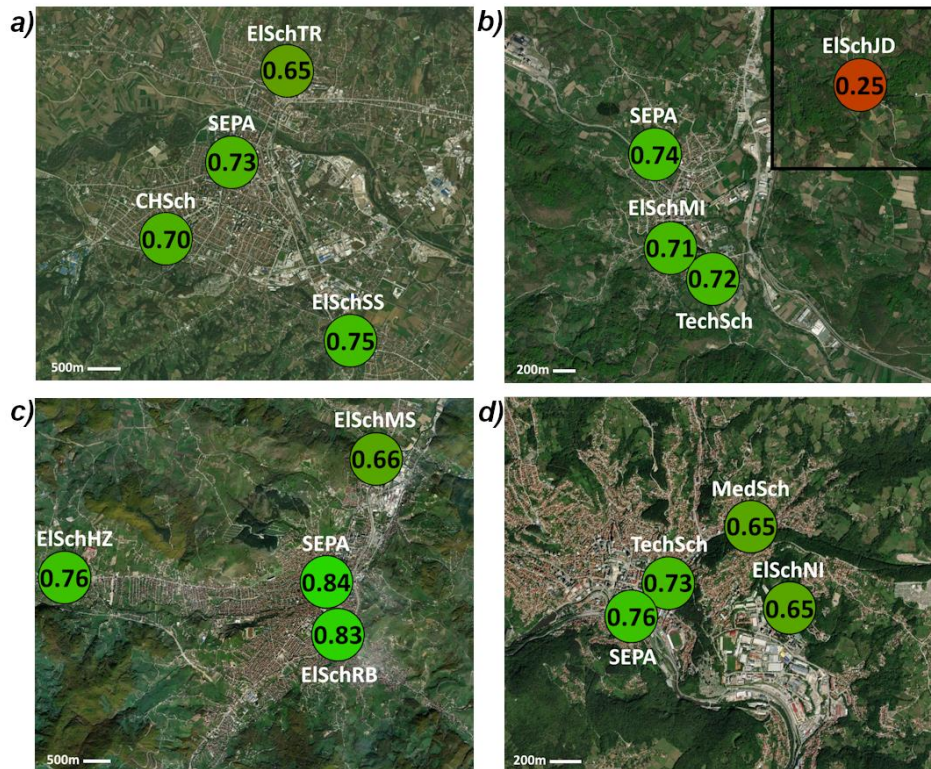


Figure 2 Maps of the campaign locations with the S-MESH data correlation values

The correlations of the campaign data with the S-MESH framework data on a monthly level presented in Figure 3 show much higher variance in correlations on a smaller sample of 30–31 days. Monthly correlation values vary mostly between 25–75%. In two schools which are the most remote from their closest monitoring stations the monthly correlations were negative in one summer month in this period. The best monthly correlations overall in all 4 municipalities are measured in July and October of 2022, which is presumably related to the remarkably clear and dry weather, expected in July but surprisingly recorded in weather data history for October 2022 in Serbia. This can be attributed to the better predictions of the S-MESH framework data.

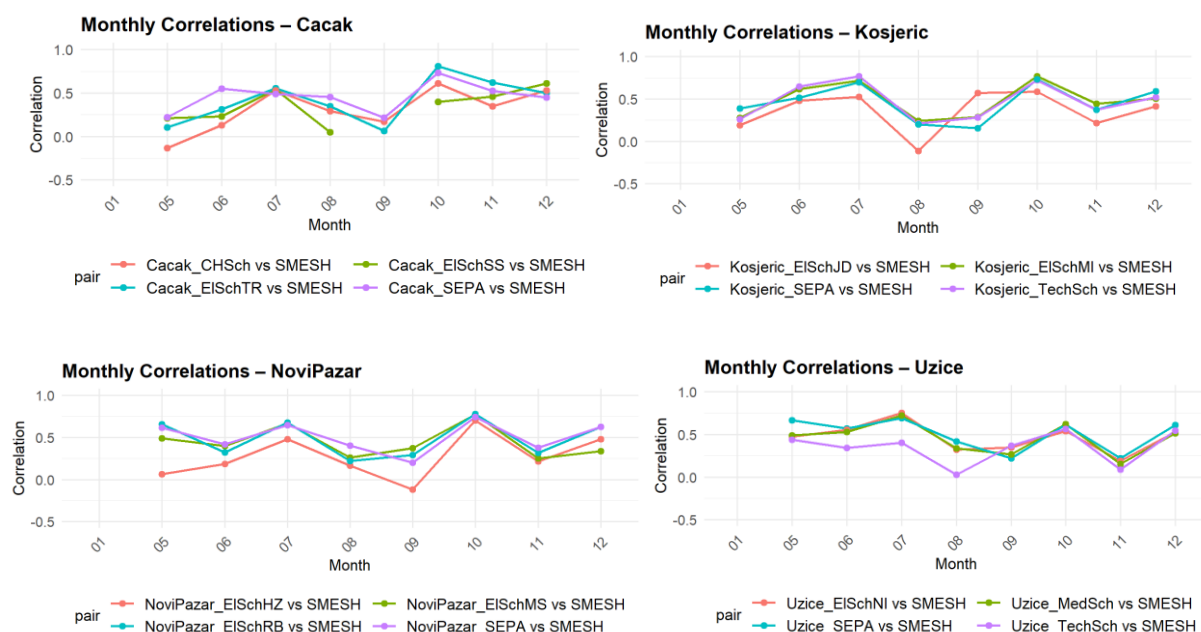


Figure 3 Changes in the monthly values of correlation between daily data gathered during the campaign and the S-MESH data at the measurement locations

RMSE (avg. $30.31 \mu\text{g m}^{-3}$), MAE (avg. $18.17 \mu\text{g m}^{-3}$) and bias (avg. $-17.49 \mu\text{g m}^{-3}$) values were calculated for all measurement sites. All biases were negative meaning that S-MESH consistently underestimates $\text{PM}_{2.5}$ compared to ground sensors. In cities with complex pollution (Novi Pazar, Cacak), SMESH misses peaks which could underestimate exposure risks. Because the model is trained and validated using monitoring stations, it will usually perform best in regions with dense (usually urban), representative monitoring coverage. That is also true with our dataset with the EISchJD site being a clear outlier. Additionally, while 1 km can be considered high resolution, it still is not street scale and can miss sharp gradients near roads, intersections, street canyons, small heating sources etc.

CONCLUSION

In this study, we show that data from monitoring stations, low-cost sensors, and satellite S-MESH data in western Serbia between April and December 2022 are well correlated. Findings reveal that almost all low-cost sensor data align well with the S-MESH data, with average correlation values of approximately 70%. This study shows that the low-cost sensor networks in combination with satellite derived S-MESH data can provide information on long-term PM concentration trends. This is of particular importance for locations which do not have monitoring stations, but also confirms the usefulness of satellite-derived ground based PM concentrations in general.

ACKNOWLEDGEMENT

This research was funded by the European Union's Horizon Europe Research and Innovation Program under GA 952433, H2020 (VIDIS project) and GA 101060170 (WeBaSOOP project), the Government of Norway and implemented by UNICEF initiative "Schools for Better Air Quality: Citizens-Based Monitoring, STEM Education, and Youth Activism in Serbia" and the Ministry

Science, Technological Development and Innovation of the Republic of Serbia under GA 451-03-33/2026-03/ 20001.

REFERENCES

- [1] Pryor, J.T., Lachlan O.C., Stephanie E.S. *Front. Public Health.* 2022, 10, 882569.
- [2] EEA, Air quality status report 2025.
- [3] Liu, Y. *Atmos. Environ.* 2013, 68, 52–53.
- [4] Cheeseman, M., Ford, B., Rosen, Z. *et al.* *Atmos. Meas. Tech. Discuss.*, 2021, 1–30.
- [5] Just, Allan C., *et al.* *Environ. Sci. Technol.* 2015, 49.14, 8576–8584.
- [6] Handschuh, J., Erbertseder, T., Baier, F. *Remote Sens.* 2023, 15 (8).
- [7] Xu, X., Zhang, C., Liang, Y. *Atmos. Environ.* 2021, 256, 118302.
- [8] Shetty, S., Hamer, P.D., Stebel, K., *et al.* *Environ. Res.* 2025, 264 (P2), 120363.
- [9] Castell, N., Dauge, F.R., Schneider, P., *et al.* *Environ. Int.* 2017, 99, 293–302.
- [10] Hassani, A., Bykuć, S., Schneider, P. *et al.* *Atmos. Environ.* 2023, 314, 120108.
- [11] Lepioufle, J.-M., Schneider, P., Hamer, P.D., *et al.* *Environ. Data Sci.* 2024, 3, e19.
- [12] Li, J., *et al.* *Atmos. Environ.* 2020, 223, 117293.
- [13] Huang, K., Bi, J., Meng, X., *Sci. Total Environ.* 2019, 697, 134094.
- [14] Gupta, P., Doraiswamy, P., Levy, R., *et al.* *GeoHealth* 2018, 2.6, 172–181.
- [15] Stojanović, D.B., *et al.* *Sensors* 2024, 24.13, 4052.
- [16] Serbian Environmental Protection Agency. Serbian Environmental Protection Agency. Belgrade: Ministry of Environmental Protection; 2026. Available from: <https://sepa.gov.rs/>
- [17] Chen, T., Guestrin, C., 2016. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 785–794.
- [18] Copernicus Land Monitoring Service. CORINE Land Cover 2018 (vector/raster 100 m), Europe, 6-yearly. 2026. Available from: <https://land.copernicus.eu/en/products/corine-land-cover/clc2018>