

PREDICTION OF AIR POLLUTANT CONCENTRATIONS BY SELECTED NEURAL NETWORK ARCHITECTURES

Terézia Kurucová, Alžbeta Andrejková, Viktor Pristaš, Ľubomír Antoni, Gabriel Semanišin
Institute of Computer Science, Faculty of Science, Pavol Jozef Šafárik University in Košice,
Jesenná 5, 040 01 Košice, Slovakia
terezia.kurucova@student.upjs.sk, alzbeta.andrejkova@student.upjs.sk

Abstract

Air pollution currently represents a significant environmental challenge with a direct impact on both public health and ecosystems. One of the aims of this paper is to predict the concentration of particulate matter PM₁₀ and CO₂ in the air using machine learning methods, specifically by employing various types of neural networks, including feedforward, convolutional, and recurrent models. The aim is also to compare the performance of these architectures in processing univariate time series with different temporal resolutions (monthly, hourly). The study uses real data from the Slovak Hydrometeorological Institute (SHMÚ) and modelled emissions from the EDGAR database. The models were trained, optimized, and evaluated using standard error metrics. The results showed that Conv1D models achieved the highest accuracy on hourly data of PM₁₀, while feedforward architectures performed best on monthly national data of CO₂. The findings confirm the potential of artificial intelligence methods as a tool for environmental monitoring and decision support in air quality protection. Opportunities for extending the models with meteorological variables and applying them in real-time warning systems are also discussed.

Keywords: air pollution, PM₁₀, CO₂ emissions, neural networks, forecasting

Napovedovanje koncentracij onesnaževal zraka z uporabo izbranih nevronske mrežne arhitekture

Izveček

Onesnaženost zraka trenutno predstavlja pomemben okoljski izziv z neposrednim vplivom na javno zdravje in ekosisteme. Eden od ciljev tega prispevka je napovedovanje koncentracije trdnih delcev PM₁₀ in CO₂ v zraku z uporabo metod strojnega učenja, zlasti z uporabo različnih vrst nevronske mreže, vključno s povratnimi, konvolucijskimi in rekurentnimi modeli. Cilj je tudi primerjava učinkovitosti teh arhitektur pri obdelavi enodimenzionalnih časovnih vrst z različno časovno ločljivostjo (mesečno, urno). Študija uporablja realne podatke Slovaškega hidrometeorološkega inštituta (SHMÚ) ter modelirane emisije iz baze podatkov EDGAR. Modeli so bili učeni, optimizirani in ocenjeni z uporabo standardnih metrik napak. Rezultati so pokazali, da so modeli Conv1D dosegli najvišjo natančnost pri urnih podatkih PM₁₀, medtem ko so povratne arhitekture dosegle najboljše rezultate pri mesečnih nacionalnih podatkih CO₂. Ugotovitve potrjujejo potencial metod umetne inteligence kot orodja za okoljsko spremljanje in podporo odločanju pri varovanju kakovosti zraka. Obravnavane so tudi možnosti razširitve modelov z meteorološkimi spremenljivkami ter njihova uporaba v sistemih zgodnjega opozarjanja v realnem času.

Ključne besede: onesnaženje zraka, PM₁₀, emisije CO₂, nevronske mreže, napovedovanje

1 INTRODUCTION

Air pollution remains one of the most pressing environmental and public-health challenges, with particulate matter (PM_{10} and $PM_{2.5}$) among the key drivers of disease burden worldwide [1]. In parallel, the accurate forecasting of greenhouse-gas emissions, especially CO_2 , is essential for designing effective prevention strategies and monitoring progress toward climate targets [2]. Data-driven methods, and neural networks in particular, have become prominent tools for modeling nonlinear dynamics in environmental time series [3], complementing inventories and monitoring systems such as EDGAR [4], [5], [6].

Recent studies show that deep learning can capture short-term fluctuations of urban air quality more effectively than conventional statistical baselines. Convolutional and recurrent architectures (e.g., CNNs, LSTMs, and ConvLSTMs) have achieved strong performance in forecasting PM concentrations at hourly to daily horizons, often outperforming autoregressive models when sufficient data are available [7], [8], [9], [10]. For CO_2 forecasting, both machine learning and deep-learning approaches, ranging from support vector regression (SVR) and random forests to feedforward and recurrent networks, have reported competitive accuracy on annual, monthly, and even daily series, particularly when combined with careful preprocessing and hyperparameter tuning [2].

This paper unifies two case studies focused on univariate forecasting with different temporal resolutions. The first case study targets PM_{10} using hourly observations from the Slovak Hydrometeorological Institute (SHMÚ); the second targets national CO_2 emissions using monthly aggregates. Across both, we evaluate selected neural architectures: feedforward (Dense/FFNN), one-dimensional convolutional (Conv1D), and recurrent (LSTM). For the CO_2 study, we also benchmark classical ML baselines (linear regression, SVR, random forest). Our design emphasizes consistent sliding-window inputs, identical train/validation/test splits per frequency, and error measured mainly by mean absolute error (MAE), along with RMSE, R^2 , and MAPE where appropriate [3].

PM_{10} concentrations and CO_2 emissions were selected as prediction targets due to their significant environmental and societal impact. PM_{10} is directly associated with respiratory and cardiovascular diseases and is widely used as an indicator of air quality in urban and industrial regions. CO_2 emissions represent an important indicator of anthropogenic activity and contribute to climate change processes. Although both variables are routinely measured by environmental monitoring stations, predictive models provide additional practical value by enabling short-term forecasting of hazardous pollution levels, supporting early-warning systems, and assisting decision-making processes in environmental management. Furthermore, predictive approaches are particularly useful in situations with missing sensor data, limited spatial coverage of monitoring stations, delayed measurements, or rapidly changing environmental conditions where real-time forecasting is required.

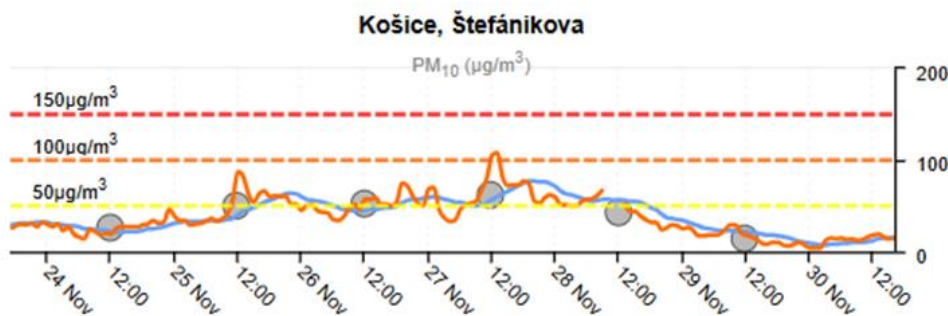


Figure 1: Example of hourly PM_{10} concentrations ($\mu g/m^3$) recorded at the monitoring station in Košice, Štefánikova, between 24 and 30 November 2024. The solid orange line represents hourly measured concentrations. The blue line denotes a 12-hour moving average, while the grey circles correspond to daily averages. Dashed horizontal lines indicate regulatory threshold values: $50 \mu g/m^3$ (yellow, information threshold), $100 \mu g/m^3$ (orange, warning threshold), and $150 \mu g/m^3$ (red, alert threshold). Such temporal series formed the basis for the model training, validation, and testing datasets used in this study.

The contributions are the following: (1) results of the joint experiment that enable the comparison of neural architectures across hourly, and monthly horizons; (2) an empirical assessment that confirms the suitability of Conv1D for short-horizon PM_{10} (hourly/monthly) and Dense/FFNN for longer horizons (annual), in line with trends reported in previous work [7], [8], [10]; and (3) a discussion of practical extensions, adding meteorological covariates and deploying the models within operational warning systems. The remainder of the paper details methodology and results of the PM_{10} case study (Subsection 2.1 and 3.1) and the CO_2 case study (Subsection 2.2 and 3.2), and concludes with implications and future directions (Section 4).

2 METHODOLOGY

In this section, we present the methodology for the two case studies mentioned.

2.1 Case Study 1: PM_{10} Forecasting

The first case study addresses the prediction of particulate matter (PM_{10}) concentrations, which are among the most harmful air pollutants affecting public health and urban ecosystems. For this purpose, we employed hourly air quality measurements obtained from the Slovak Hydrometeorological Institute (SHMÚ), which maintains an extensive network of automatic monitoring stations across Slovakia. The SHMÚ air quality database is widely used in environmental and epidemiological research, and provides validated measurements of key pollutants including PM_{10} , $PM_{2.5}$, NO_2 , and O_3 [11], [12].

The dataset used in this study consisted of univariate time series of PM_{10} concentrations collected at an urban background station in Košice. Hourly averages were derived according to European Air Quality Directive standards to ensure comparability and consistency [12]. The data span multiple years (from 1st January 2020 to 2nd January 2025) and capture both seasonal variability and extreme episodes of air pollution, which are important characteristics for time series forecasting tasks. The dataset consists of hourly averaged PM_{10} concentrations expressed in micrograms per cubic meter ($\mu g/m^3$), provided by the Slovak Hydrometeorological Institute (SHMÚ). Figure 1 illustrates an example of such a time series measured at the Košice–Štefánikova monitoring station during the period 24–30 November 2024, including regulatory threshold levels used for air quality assessment.

The dataset was transformed into a supervised learning problem using sliding windows of fixed length. Particularly, the input data were structured as univariate sliding windows of length w , where the past w hourly concentrations were used to predict the next hour. Different window sizes (3, 6, 7, and 8 hours) were tested to evaluate the influence of temporal context on predictive performance.

Following the approach implemented in `keras.utils.timeseries_dataset_from_array`, each input sample consisted of a sequence of eight consecutive hourly observations (`sequence_length = 8`), and the target variable was defined as the PM_{10} concentration in the subsequent time step (`delay = 1`). This setup allowed the models to learn short-term temporal dependencies while predicting the next-hour concentration.

To ensure a realistic forecasting scenario, the time series was split chronologically into non-overlapping training, validation, and test sets. Using the sliding-window approach, the effective number of available samples was determined not only by the prediction delay but also by the input window length. More precisely, for a time series of length N , an input window of size w , and prediction delay d , the total number of generated samples was equal to $N - w - d + 1$. This formulation more accurately reflects the number of valid input-output pairs that can be constructed without exceeding the temporal boundaries of the dataset.

The samples were allocated proportionally: 50% for training, 25% for validation, and 25% for testing. In practice, the partitioning was implemented as:

- **Training set:** first 50% of samples,
- **Validation set:** next 25% of samples,
- **Test set:** final 25% of samples.

Furthermore, although samples within the training and validation subsets were shuffled during optimization, the chronological order between the training, validation, and testing partitions was strictly preserved in order to avoid temporal leakage. The use of shuffling inside the training subset together with a batch size of one was intended to improve stochastic optimization stability and reduce sensitivity to local temporal correlations between adjacent samples. Since each training sample already consisted of a complete temporal window, the internal temporal structure of individual sequences remained preserved during model training.

The sliding-window sampling procedure generated supervised pairs by shifting the input window step-by-step through each subset. Samples were shuffled within the training and validation sets to improve generalization, while batch size was set to one (`batch_size = 1`) to preserve temporal granularity. This design ensured that the validation and test sets always represented later, unseen portions of the time series, thereby avoiding temporal leakage and enabling fair evaluation of the forecasting models.

To model PM_{10} dynamics, we experimented with three types of neural network architectures:

- **Dense feedforward networks (FFNN):** employed as a baseline deep learning approach for univariate regression tasks. In our implementation, the network was initialized with the HeNormal initializer to ensure stable weight distribution for ReLU activations. The architecture consisted of two hidden layers with 32 and 64 neurons respectively, followed by a linear output layer for regression. The network was trained with the Adam optimizer and mean squared error (MSE) loss over 300 epochs. Model performance was monitored with mean absolute error (MAE), while checkpoints were applied to save the best-performing model. Although FFNNs can capture non-linear dependencies between input features, they do not explicitly account for temporal dependencies, which limits their suitability for sequential air quality data.
- **One-dimensional convolutional neural networks (Conv1D):** used to extract local temporal features such as short-term spikes or diurnal cycles. The implemented Conv1D model applied two convolutional blocks: the first with 64 filters of kernel size 5, followed by max pooling, and the second with 128 filters of kernel size 5 and another pooling layer. After convolutional layers, the network employed a fully connected layer of 256 units with ReLU activation and HeNormal initialization, combined with dropout ($p = 0.3$) for regularization. The output layer was a single regression neuron. The training was performed with Adam and MSE loss, using early stopping with patience 100 and model checkpoints. Conv1D architectures have proven effective for high-

frequency air quality time series, as they capture repeating patterns and localized dependencies in the data.

- **Long short-term memory networks (LSTM):** applied to capture long-range temporal dependencies and seasonal cycles in pollutant concentrations. The implemented model stacked three LSTM layers (32, 64, and 128 units) with dropout regularization (0.3–0.4) to reduce overfitting. The LSTM stack was followed by two dense layers (64 and 32 units) with tanh activation and combined L_1 - L_2 regularization ($l1_l2(0.001, 0.001)$). This setup allowed the network to balance memorization of long sequences with robust generalization. Training used the Adam optimizer, MSE loss, a batch size of 128, and adaptive learning rate reduction (`ReduceLROnPlateau`) when validation performance plateaued. LSTMs are widely recognized for their superior ability to model sequential data in air quality applications, especially when long-term dependencies and seasonality must be considered.

The models were trained using the Adam optimizer with an early stopping criterion to avoid overfitting. Performance was evaluated on a held-out test set using two standard regression metrics: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). MAE provides a direct measure of average prediction deviation, while RMSE penalizes larger errors more strongly and is therefore sensitive to extreme pollution episodes.

2.2 Case Study 2: CO₂ Emissions Forecasting

The second case study employed CO₂ emission data from the Emissions Database for Global Atmospheric Research EDGAR [6], jointly developed by the European Commission, Joint Research Centre (JRC), and the International Energy Agency (IEA). The dataset consists of monthly records from 1970 to 2023, available for all countries. Although emissions are reported separately by sector, we considered total national emissions. Each record contains the country code, country group, country name, and monthly emission values. Unlike the air quality dataset used in Case Study 1, the emissions are expressed in terms of CO₂-equivalents (CO₂eq), representing a unit based on the global warming potential (GWP) of different greenhouse gases. For this case study, we focused on Slovakia and Poland.

2.2.1.1 Data preprocessing.

The dataset was partitioned chronologically, with data from 1970–2010 used for training and data from 2011–2023 reserved for testing. To formulate the time series as a supervised learning task, the monthly data were reshaped into sliding windows of sizes 3, 5, and 7. For example, a window size of 5 corresponds to using five consecutive months as input features and predicting the emission value of the subsequent month as the target. This transformation was applied separately to each country.

2.2.1.2 Models and hyperparameter optimization.

We implemented a set of classical machine learning models (support vector regression, and random forest regression), along with a feedforward neural network (FFNN). Hyperparameter tuning was performed using `GridSearchCV` (and in selected cases `RandomizedSearchCV`) combined with cross-validation.

- **Feedforward neural network (FFNN):** Grid search explored combinations of 3–6 hidden layers with initial layer sizes of 128, 256, or 512 neurons. The number of neurons in subsequent layers followed a halving rule, with a minimum size of 8 neurons to ensure sufficient model capacity. Training used the Adam optimizer with mean squared error (MSE) loss and a batch size of 32. During hyperparameter search, the number of epochs was fixed at 10 to reduce computational cost. The best architecture was then retrained with 100 epochs to determine the optimal training duration. Five-

fold cross-validation ($k = 5$) was employed, and MAE served as the selection metric during grid search.

- **Support vector regression (SVR):** Experiments were conducted with both radial basis function (RBF) and polynomial kernels. For RBF, the parameter grid included $C \in \{1, 10, 20, 50\}$, $\epsilon \in \{0.01, 0.1, 0.2\}$, and $\gamma \in \{\text{scale}, \text{auto}\}$. For the polynomial kernel, degrees 2–5 were tested, with additional parameter $\text{coef0} \in \{0, 1, 2\}$. Due to the high computational cost of the degree-5 polynomial kernel, RandomizedSearchCV was used with 20 iterations. All SVR models required data standardization, and parameter tuning was guided by 5-fold cross-validation.
- **Random forest regression (RFR):** we performed hyperparameter tuning $\text{max_depth} \in \{3, 4, 5\}$, $\text{min_samples_split} \in \{2, 3, 5\}$, and $\text{criterion} \in \{\text{squared_error}, \text{absolute_error}\}$. Cross-validation was carried out with $k = 2$, balancing reliability with computational feasibility. Unlike SVR, no feature standardization was necessary. Once the best parameters were identified, the model was retrained on the training set and evaluated on the test set.

2.2.1.3 Evaluation metrics.

Initially, models were compared using multiple error metrics: mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R^2), and mean absolute percentage error (MAPE). In the final stage of analysis, however, we report performance exclusively using R^2 , as it provides a normalized measure of explained variance and allows direct comparability between models and datasets.

Different evaluation metrics were used for the PM_{10} and CO_2 forecasting experiments due to the different temporal characteristics and practical objectives of the analyzed tasks. For hourly PM_{10} prediction, MAE was selected because it directly reflects the average magnitude of prediction errors in the original concentration units and is well suited for short-term air quality forecasting, where accurate estimation of absolute deviations is important for practical environmental monitoring.

In contrast, the monthly CO_2 emission forecasting task focused primarily on capturing long-term temporal trends and overall variability rather than short-term fluctuations. Therefore, R^2 was used as the primary evaluation metric, since it quantifies how well the model explains the variance in the observed data and is commonly applied in regression-based trend forecasting tasks.

As a consequence, the reported PM_{10} and CO_2 results are not directly numerically comparable because the metrics measure different aspects of predictive performance. Instead, each metric should be interpreted within the context of its respective forecasting objective and temporal resolution.

3 RESULTS

3.1 PM_{10} forecasting results

The models were evaluated on the hourly PM_{10} concentration data using mean absolute error (MAE) as the primary metric. Table 1 summarizes the performance of the baseline naive method and the three deep learning architectures.

Table 1: Evaluation of MAE on hourly PM_{10} data (for windows size of 8).

Model	Train MAE	Val MAE	Test MAE
Naive	7.44	4.44	6.22

Model	Train MAE	Val MAE	Test MAE
Dense	9.63	6.14	8.23
Conv1D	4.76	3.12	4.03
LSTM	7.85	5.22	7.19

The naive baseline (computed as the average of values within the time window) achieved a validation MAE of 4.44, which served as a reference point for the neural models. The dense feedforward network underperformed compared to the baseline, with higher errors across all data splits. In contrast, the one-dimensional convolutional network (Conv1D) outperformed all other models, achieving the lowest validation (3.12) and test (4.03) MAE, demonstrating its effectiveness in capturing short-term temporal patterns in air quality data. The LSTM network showed moderate improvements over the dense network but did not surpass the performance of Conv1D, indicating that for this dataset the convolutional approach was more suitable than recurrent modeling of long-range dependencies.

The observation that the training MAE was slightly higher than the validation MAE can be explained by the training configuration and evaluation protocol used in the experiments. During the training phase, regularization mechanisms such as dropout and stochastic optimization introduced additional variability, making the prediction task effectively more difficult for the model. In contrast, validation evaluation was performed in inference mode, where dropout was disabled and predictions were therefore more stable. Additionally, the chronological splitting strategy may have resulted in slightly different statistical properties between the training and validation subsets, with the validation portion containing smoother temporal patterns and lower short-term variability. Consequently, lower validation MAE values do not necessarily indicate data leakage or overfitting, but rather reflect the combined effect of regularization and temporal distribution differences between the subsets.

3.2 CO₂ forecasting results

The performance of the models in forecasting monthly CO₂ emissions was evaluated using the coefficient of determination (R^2). Table 2 summarizes the results for Slovakia and Poland.

Table 2: Evaluation of R^2 for monthly CO₂ emissions forecasting (2011–2023).

Method	Slovakia	Poland
FFNN	0.92	0.88
RFR	0.21	0.64
SVR-Poly2	0.89	0.86
SVR-Poly3	0.66	0.88
SVR-Poly4	0.71	0.86
SVR-Poly5	0.69	0.83
SVR-RBF	0.72	0.86

For Slovakia, the feedforward neural network (FFNN) achieved the best performance with $R^2 = 0.92$, followed closely by support vector regression (SVR) with polynomial kernel of degree 2 ($R^2 = 0.89$). Figure 2 illustrates the comparison between the observed and predicted CO₂ emissions for Slovakia in the test period. The real values (solid blue line) and predicted values (dashed red line) demonstrate a close alignment, indicating that the FFNN effectively captured both seasonal fluctuations and longer-term emission dynamics.

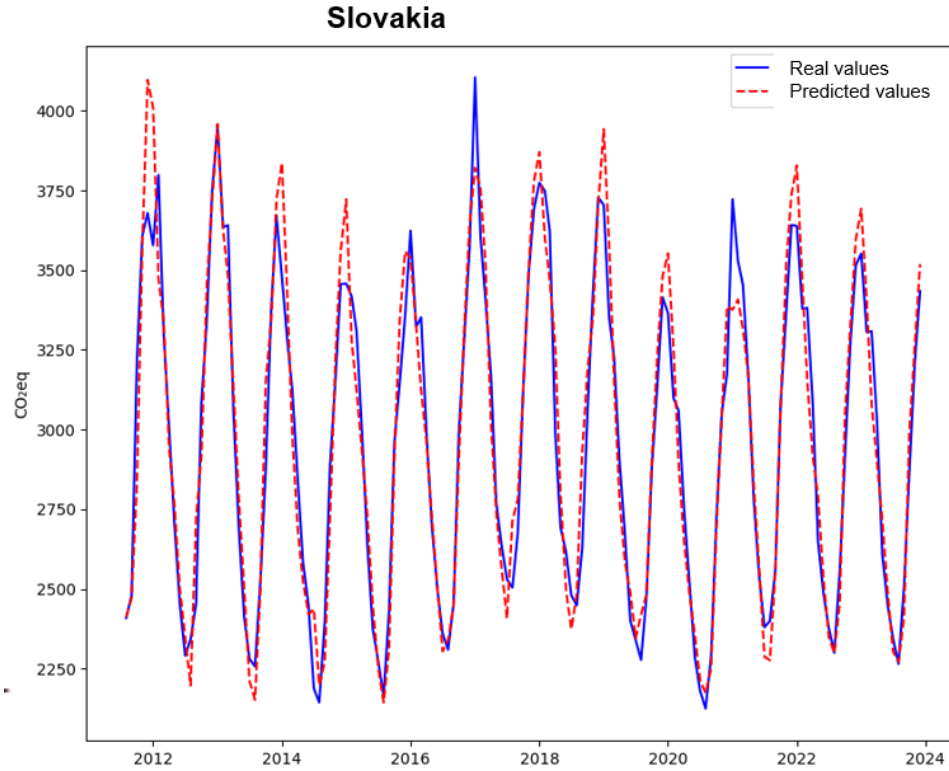


Figure 2: Comparison of observed (blue) and predicted (red, dashed) monthly CO₂-equivalent emissions in Slovakia for the period 2011–2023, obtained using the FFNN model. The horizontal axis represents time, and the vertical axis denotes CO₂eq emissions.

For Poland, the results were more balanced across models. FFNN again achieved strong performance ($R^2 = 0.88$), comparable to SVR with polynomial kernel of degree 3 ($R^2 = 0.88$). Other SVR configurations and the RBF kernel also performed consistently well (R^2 around 0.83–0.86). RFR showed moderately good performance ($R^2 = 0.64$), which was significantly better than in the Slovak case.

Overall, FFNN and SVR models outperformed RFR, demonstrating that models capable of capturing non-linear temporal dependencies provided the most accurate forecasts of monthly CO₂ emissions.

4 DISCUSSION

The forecasting models were trained and evaluated using data collected from a single monitoring location in Košice, Slovakia. Hourly PM₁₀ measurements were obtained from the Slovak Hydrometeorological Institute (SHMÚ), while CO₂ emission-related variables were derived from the EDGAR database. Since the study focused on one geographical area, spatial variability between multiple monitoring stations was not explicitly modeled. Consequently, the proposed models primarily capture temporal dynamics specific to the analyzed urban environment. This design allows for a controlled evaluation of prediction performance while reducing the influence of inter-station heterogeneity. Nevertheless, the proposed methodology can be extended in future work to incorporate multiple monitoring stations and spatial features in order to improve model generalizability across different geographical regions and environmental conditions.

The results of the two case studies illustrate how different machine learning approaches perform under distinct environmental forecasting contexts. In the first case study, focusing on short-term PM₁₀ forecasting using hourly air quality data, the one-dimensional convolutional

neural network (Conv1D) outperformed both dense feedforward networks (FFNN) and recurrent LSTM models. This finding highlights the effectiveness of convolutional filters in detecting short-term patterns such as diurnal cycles and episodic pollution spikes. Although LSTMs are widely used for sequential modeling, in this setting the limited dataset size and the dominance of local temporal structures favored convolutional approaches over recurrent ones.

In the second case study, which examined long-term CO₂ emission forecasting using monthly records from EDGAR and IEA sources, FFNN and support vector regression (SVR) provided the best results. These outcomes show that non-linear regression methods are well-suited to datasets where both long-term emission pathways and seasonal cycles interact. Taken together, the two studies demonstrate that the optimal modeling strategy depends strongly on the temporal granularity and structural properties of the data.

An additional opportunity for extending this work lies in the integration of meteorological variables such as temperature, wind speed, humidity, and atmospheric pressure. These factors are known to strongly influence both air quality dynamics and emission dispersion, and their inclusion could improve model robustness and generalization. Moreover, embedding the trained models into real-time early warning systems would allow authorities to issue short-term alerts for high-pollution events or monitor deviations from expected emission trajectories.

5 CONCLUSION

This work applied selected machine learning models to two environmental forecasting problems: prediction of PM₁₀ concentrations and forecasting of CO₂ emissions. The findings show that Conv1D networks are highly effective for high-frequency air quality data, where short-term temporal dependencies dominate. FFNNs and SVR models are well-suited to aggregated datasets such as monthly CO₂ emissions, where smooth non-linear patterns must be captured.

From a practical perspective, the results have several implications:

- **Air quality management:** Accurate PM₁₀ forecasts can support the design of early warning systems and short-term prevention policies, such as traffic restrictions or targeted public health alerts.
- **Climate policy:** Reliable forecasts of national CO₂ emissions provide policymakers with an additional tool for monitoring emission reduction progress and assessing the effectiveness of implemented measures.

Future research should explore models that combine convolutional and recurrent layers, as well as the integration of explanatory variables such as meteorological factors for air quality and macroeconomic indicators for emissions. Incorporating explainable AI methods may also increase trust in these models, enhancing their practical uptake by environmental agencies and policy institutions.

6 ACKNOWLEDGMENTS

This student paper (T. Kurucová, A. Andrejková) was co-authored with university teachers (V. Pristaš, Ľ. Antoni, G. Semanišin) as part of the Erasmus+ project Developing Talents in Artificial Intelligence to Solve Disruptive Environmental Problems (Project No. 2023-1-PL01-KA220-HED-000166765). The project aims to enhance the quality and impact of AI education while strengthening the digital readiness of higher education institutions and students.

We gratefully acknowledge the Slovak Hydrometeorological Institute (SHMÚ) for providing access to the air quality data used in this study.

References

- [1] World Health Organization, *WHO global air quality guidelines: Particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. Geneva: World Health Organization, 2021.
- [2] A. A. Ajala, O. L. Adeoye, O. M. Salami, and A. Y. Jimoh, "An examination of daily CO₂ emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models," *Environmental Science and Pollution Research International*, vol. 32, no. 5, pp. 2510–2535, 2025.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, MA: MIT Press, 2016. Available: <http://www.deeplearningbook.org>
- [4] G. Janssens-Maenhout *et al.*, "EDGAR v4.3.2 global atlas of the three major greenhouse gas emissions for the period 1970–2012," *Earth System Science Data*, vol. 11, no. 3, pp. 959–1002, 2019.
- [5] M. Crippa *et al.*, "High resolution temporal profiles in the emissions database for global atmospheric research (EDGAR)," *Scientific Data*, vol. 7, p. 121, 2020.
- [6] M. Crippa *et al.*, "GHG emissions of all world countries," Publications Office of the European Union, 2024. doi: [10.2760/4002897](https://doi.org/10.2760/4002897).
- [7] H. Se-Chae, Y.-J. Ahn, and H.-S. Kim, "PM₁₀ and PM_{2.5} real-time prediction models using an interpolated convolutional neural network," *Scientific Reports*, vol. 11, p. 91253, 2021.
- [8] İ. Aksangür, S. Seker, and C. Yilmaz, "Evaluation of data preprocessing and feature selection process for prediction of hourly PM₁₀ concentration using long short-term memory models," *Environmental Pollution*, vol. 311, p. 119973, 2022.
- [9] K. Zhang, X. Yang, H. Cao, J. Thé, Z. Tan, and H. Yu, "Multi-step forecast of PM_{2.5} and PM₁₀ concentrations using convolutional neural network integrated with spatial–temporal attention and residual learning," *Environment International*, vol. 171, p. 107691, 2023.
- [10] M. Kostadinov *et al.*, "Forecasting air pollution with deep learning with a focus on impact of urban traffic on PM₁₀ and noise pollution," *PLOS ONE*, vol. 19, no. 12, p. e0313356, 2024.
- [11] Slovak Hydrometeorological Institute, "Air quality data portal." 2025. Available: <http://www.shmu.sk>
- [12] European Environment Agency, "Air quality in europe – 2020 report," Publications Office of the European Union, Luxembourg, 09/2020, 2020. Available: <https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report>

Terézia Kurucová is a master's student in Data Analysis and Artificial Intelligence at Faculty of Science at Pavol Jozef Šafárik University in Košice. Her research interests focus on machine learning, time series analysis, and environmental data modelling. She is actively involved in projects related to air quality monitoring and neural network applications.

Alžbeta Andrejková is a master's student of Applied Informatics at Faculty of Science at Pavol Jozef Šafárik University in Košice. She specializes in data analysis, predictive modelling, and artificial intelligence techniques. Her academic work concentrates on processing environmental data and evaluating neural network architectures for forecasting tasks. She participates in student research activities and applied AI projects.

Viktor Pristaš is a PhD student in the field of Computer Science at Faculty of Science at Pavol Jozef Šafárik University in Košice. His research focuses on advanced machine learning methods, and time series forecasting, particularly in environmental and ecological domains. He has contributed to several scientific publications and research projects addressing energetics modelling and intelligent data analysis.

Lubomír Antoni is an associate professor at the Institute of Computer Science, Faculty of Science, Pavol Jozef Šafárik University in Košice. His research interests include artificial intelligence, fuzzy systems, data mining, and applied machine learning.

Gabriel Semanišin is a professor of Computer Science at Faculty of Science, Pavol Jozef Šafárik University in Košice. As part of his research activities, he focuses mainly on algorithmic graph theory and its application in various areas of theoretical and applied informatics. He is a co-guarantor of the study programs Applied Informatics, Data Analysis and Artificial Intelligence and Computer Science. He was a supervisor of 6 PhD students in study programs Computer Science, Discrete Mathematics and Theory of Teaching Informatics.