An Explainable Hybrid Convolutional Neural Network for NO_2 Concentration Forecasting in Milan

G. Piccoli[1], S. Raccagni[2], C. Carnevale[3], S. De Nardi[4]

***Abstract*— Nitrogen dioxide (NO_2) is a major air pollutant with well-documented adverse effects on human health and ecosystems, particularly in urban areas. Accurate short-term forecasting of NO_2 concentrations is therefore essential to support air quality monitoring and mitigation strategies. In this work, we propose a hybrid convolutional neural network (HCNN) model to predict daily NO_2 concentrations in the city of Milan by jointly exploiting spatially distributed nitrogen oxides emission maps over the Lombardy region and the previous day's NO_2 concentration as an autoregressive input. The model combines a convolutional module to process emission maps with a feedforward network to integrate spatial and pointwise information. The HCNN achieved good predictive performance and demonstrated the ability to identify exceedances of health-related daily exposure thresholds. Model behavior was further investigated using explainable artificial intelligence techniques based on SHAP, enabling the assessment of feature importance and the identification of emission sectors and spatial regions most relevant to the forecasts. Overall, the results demonstrate the potential of hybrid deep learning models coupled with explainable methods to provide accurate and interpretable air quality forecasts, while also enabling the analysis of the learned relationships between emissions and predicted concentrations, a key aspect for future control-oriented applications aimed at identifying optimal emission reduction pathways.**

I. INTRODUCTION

Air pollution remains a major global concern, with a significant portion of the world's population exposed to pollutant concentrations exceeding levels considered harmful to human health [1]. Prolonged exposure is associated with increased risks of respiratory and cardiovascular diseases, reduced quality of life, and lower life expectancy [2], [3]. In addition to human health impacts, air pollution adversely affects ecosystems, agriculture, and biodiversity, contributing to environmental degradation [4].

Air pollutants can be classified as primary, directly emitted from natural or anthropogenic sources (e.g., SO_2, NO_x, CO, VOCs, particulate matter), or secondary, formed through atmospheric chemical reactions (e.g., O_3, NO_2) [1]. Among these, nitrogen oxides (NO_x) play a key role in atmospheric chemistry and air quality degradation. In particular, nitrogen dioxide (NO_2) is a highly reactive gas that contributes to ozone formation, acid deposition, and reduced visibility, with significant impacts on both human health and ecosystems.

NO_2 originates from both natural processes, such as volcanic activity, and predominantly from anthropogenic sources, including combustion processes related to traffic, industry, and domestic heating [5]. Exposure to NO_2 has been linked to several adverse health outcomes, including respiratory and cardiovascular diseases. To mitigate these risks, the World Health Organization (WHO) defined guideline values in 2021 [1], while the European Union introduced limits through Directive (EU) 2024/2881 [6], setting exposure thresholds to be achieved by 2030 (Table I).

TABLE I

EXPOSURE THRESHOLDS FOR NO_2 AS DEFINED BY THE WHO AQG (2021) AND EU DIRECTIVE 2024/2881.

Exposure duration	**1 hour**	**1 day**	**1 year**
WHO AQG (2021)	200 $\mu g/m^3$	25 $\mu g/m^3$	10 $\mu g/m^3$
EU 2024/2881	200 $\mu g/m^3$	50 $\mu g/m^3$	20 $\mu g/m^3$

Understanding and forecasting NO_2 concentrations is therefore essential for supporting effective air quality management policies. In this context, Artificial Neural Networks (ANNs) have been widely used as an alternative to deterministic models due to their flexibility, low input requirements, and ability to capture complex nonlinear relationships [7], [8], [9], [10], [11], [12], [13], [14]. However, their "black box" nature limits interpretability, raising concerns in terms of trust, accountability, and regulatory compliance. To address this issue, Explainable Artificial Intelligence (XAI) techniques have been developed [15], [16], [17], and are increasingly supported by regulatory frameworks such as the DARPA XAI program [18], the European AI Act [19], and UNESCO guidelines [20]. Building on these considerations, this work addresses not only the forecasting of daily NO_2 concentrations in the city of Milan, but also the limited interpretability of current data-driven air quality models. In fact, while neural approaches can achieve high predictive performance, they often provide little insight into how different emission sources contribute to the predicted pollution levels. To overcome this limitation, we propose a neural network model that combines previous-day NO_2 measurements with spatial NO_x emission maps for the Lombardy region. XAI techniques are then applied to open the black-box structure of the model and to quantify the influence of different emission

[1]G. Piccoli is with Department of Mechanical and Industrial Engineering, University of Brescia, Via Branze 38, I-25123, Brescia (IT) gabriele.piccoli@unibs.it

[2]S. Raccagni is with Department of Mechanical and Industrial Engineering, University of Brescia, Via Branze 38, I-25123, Brescia (IT) sara.raccagni@unibs.it

[3]C. Carnevale is with Department of Mechanical and Industrial Engineering, University of Brescia, Via Branze 38, I-25123, Brescia (IT) claudio.carnevale@unibs.it

[4]S. De Nardi is with Department of Mechanical and Industrial Engineering, University of Brescia, Via Branze 38, I-25123, Brescia (IT) sabrina.denardi@unibs.it

TABLE II
GNFR SOURCE CATEGORIES

Code	Sector name	Brief description
A	Public Power	Emissions from public electricity and heat production plants.
B	Industry	Emissions from industrial combustion and manufacturing processes.
C	Other Stationary Combustion	Fuel combustion in residential, commercial, and institutional sectors.
D	Fugitives	Non-combustion emissions from fuel extraction, processing, and distribution.
E	Solvents	Emissions from the use of solvents in industrial and domestic applications.
F	Road Transport	Exhaust and non-exhaust emissions from on-road vehicles.
G	Shipping	Emissions from maritime transport, including inland and coastal shipping.
H	Aviation	Emissions from aircraft operations during landing and take-off cycles.
I	Off–Road	Emissions from mobile machinery not used on public roads.
J	Waste	Emissions from waste treatment, disposal, and incineration.
K	Agriculture Livestock	Emissions from livestock management and manure handling.
L	Agriculture Other	Emissions from crop production, fertilizer use, and agricultural soils.

sectors on NO_2 levels. In this context, the integration of a convolutional neural network based on spatial emission maps with XAI methods for the interpretation of sector-specific contributions represents a relatively unexplored research direction in air quality forecasting, supporting more transparent and informed environmental decision-making. The objective of this work is the development of a neural network model able to provide accurate predictions of NO_2 concentrations while explicitly linking them to the spatial distribution of sector-specific NO_x emissions. This aspect is particularly relevant for the future integration of the model into a control-oriented framework aimed at identifying optimal spatially distributed emission reduction strategies to mitigate pollution levels.

II. METHODS

The time series representing the NO_2 concentration in the city of Milan was collected from the European Environment Agency (EEA) Air Quality Download Service [21] and is available on a daily basis from 2010 to 2022, for a total amount of 4721 measurements. The NO_x emission maps, covering the same period with the same daily temporal resolution, were obtained from the Copernicus Atmosphere Monitoring Service (CAMS) [22], [23] and consist of matrices representing the Lombardy region, discretized into 45×45 cells with a spatial resolution of 6 km $\times$ 6 km. NO_x emissions are separated into 12 source categories following the GNFR (Gridded Nomenclature For Reporting) classification, as reported in Table II. To simplify the dataset, emissions from sectors K and L were aggregated into a single "Agriculture" sector, while sectors D, E, and G were grouped under a general "Other" sector due to their negligible contribution, resulting in 9 effective sectors and emission maps with dimensions $(4721, 45, 45, 9)$.

The NO_2 pointwise concentrations and the NO_x emission maps were combined into a single hybrid dataset, where lagged NO_2 observations were used together with NO_x emission data as additional input features.

$$\hat{NO}_2(t) = f(NO_2(t-1), NO_x(t)) \quad (1)$$

Both NO_2 and NO_x data were subsequently normalized to the interval $[0, 1]$ using min–max scaling. Afterwards, the dataset was split into a training and test set. The former comprised the data from the 2010-2019 period and was used to train a neural network model to learn the relation described in (1), the latter covered the 2020-2022 period and was used to evaluate the model's performance and its ability to generalize to new, previously unseen data.

The model used is a hybrid convolutional neural network (HCNN). HCNN is a broad term used in the literature to refer to models that combine convolutional neural networks (CNNs) with another neural network architecture or modeling approach in order to leverage the strengths of both [24], [25], [26], [27], [28], [29], [30]. In this work CNNs were combined with feedforward neural networks (FNNs), to account for both the spatial distribution of NO_x emission maps and the pointwise NO_2 concentrations. In the following, a brief introduction to CNNs and FNNs is provided.

CNNs are typically used to process two-dimensional data. They extract patterns from their input by applying a set of learnable filters (kernels) using the 2D convolution operation [31]. This operation involves sliding each filter across the input and computing the element-wise product between the filter weights and the corresponding region of the input, followed by a summation. Each filter is applied across all input channels and the results are then summed element-wise. After the convolution, a nonlinear activation function is applied to introduce nonlinearity. The number of filters determines the number of output channels, as the 2D convolution is performed in parallel for each filter. It is common practice to apply pooling operations after convolutional layers to reduce computational complexity while preserving the most relevant information extracted from the input. Pooling involves sliding a window over the input and computing a single representative value from each region. Several pooling techniques exist, with max pooling being one of the most widely used. In this method, the maximum value within each region is selected as the representative value.

FNNs are one of the simplest and most fundamental types of artificial neural networks used for supervised learning tasks [32]. An FNN consists of a sequence of fully connected layers of neurons, where each neuron is connected to all neurons in both the preceding and the succeeding layers. This architecture ensures that information flows in a unidirectional manner, from input to output, without forming cycles. Each neuron computes a weighted sum of the outputs of the preceding layer and, then, applies an activation function to introduce non-linearity. The input to the network is provided to the first layer, and the output of the final layer represents the overall output of the network.

The presented HCNN consists of a convolutional module that processes the current NO_x emission maps, and a feedforward module that combines the output of the convolutional module with the NO_2 concentration value from the previous day to predict the current NO_2 concentration. The final model architecture was determined through hyperparameter tuning performed using the Keras Tuner framework [33]. During this process, 100 architectures were trained and tested using various combinations of hyperparameters. The hyperparameters, their search ranges, and the selected best values for the final model are reported in Table III.

TABLE III

HYPERPARAMETER RANGES AND SELECTED BEST VALUES

Hyperparameter	Min	Max	Step	Best
number of conv. layers	2	6	1	3
number of filters	16	256	16	16
filter size	3	7	2	3
number of FNN layers	1	4	1	1
number of nodes	8	256	4	12
learning rate	10^{-5}	10^{-1}	log-uniform	10^{-2}

The HCNN model was subsequently trained for 100 epochs using the Adam optimizer [34] with the mean squared error (MSE) as the loss function. Its predictive performance was evaluated together with its ability to identify exceedances of the exposure thresholds reported in Table I.

Eventually, XAI techniques were applied using the SHAP library [15] to enhance the interpretability of the model and provide insights into the relationship linking the NO_x emissions and past NO_2 concentrations to the current concentration values. SHAP explanations are based on Shapley values, a concept from cooperative game theory that provides a fair way to distribute the total gain generated by a coalition of players according to their individual contribution. In the context of XAI, the gain of a forecast is defined as the deviation from the average model output, the players are the input features (i.e. the past NO_2 concentration and the NO_x emissions sectors) and the Shapley values quantify each feature's contribution to shifting a forecast from this baseline. Formally, the SHAP explanation can be expressed in its simplest additive form as:

$$\hat{y}_i = E[\hat{y}] + \sum_{j=1}^{N} \phi_{i,j} \tag{2}$$

where $\hat{y}_i$ is the model forecast for sample i, $E[\hat{y}]$ is the expected model output (i.e. the average forecast over the dataset), N is the number of input features and $\phi_{i,j}$ is the Shapley value of the feature j for the sample i. The SHAP explanations were used to identify the emission sources contributing most significantly to the predicted NO_2 concentrations, together with the spatial regions of the emission maps having the largest influence on the model output. This analysis allowed the assessment of the model's ability to represent the underlying pollution dynamics in a physically coherent and interpretable way, which is a key requirement for its future integration into a control-oriented framework for emission reduction strategies.

III. RESULTS

In this section, we present the performance of the model when asked to predict the NO_2 concentration in Milan based on the previous day concentration and the emission maps of NO_x from various source sectors. Figure 1 shows the model forecasts for the 2020–2022 test period compared to the observed NO_2 concentrations. A clear seasonal pattern is visible, likely reflecting changes in human activity between summer and winter, as well as physical phenomena such as temperature inversions, which result in higher NO_2 concentrations during winter months. The predictive performance of the model was quantitatively assessed using mean absolute error (MAE), normalized MAE (NMAE), correlation (CORR) and R^2, as reported in Table IV.

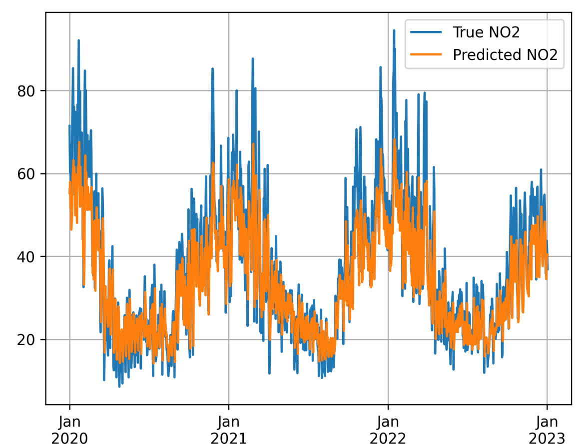

Fig. 1. Predicted vs. observed NO_2 concentrations in Milan for the 2020–2022 test period.

TABLE IV

MODEL PERFORMANCE ACROSS DIFFERENT TIME PERIODS

Metric	2020-2022	2020	2021	2022
MAE	6.997	7.183	7.001	6.774
NMAE	0.187	0.195	0.185	0.181
CORR	0.866	0.894	0.857	0.839
R^2	0.695	0.735	0.682	0.660

The model performance does not show significant variation across the different years; the MAE and NMAE values indicate that the model is capable of accurately predicting NO_2 concentrations, the correlation values suggest that the model successfully captures seasonal trends, while some daily fluctuations may not be perfectly reproduced, and the

R^2 values demonstrate that the model provides a substantial improvement over a simple mean predictor. Overall, these results indicate that the HCNN can reliably predict both the magnitude and seasonal variability of NO_2 concentrations in Milan.

Furthermore, the ability of the model to identify exceedances of the daily exposure thresholds–25μg/m^3 and 50μg/m^3 from Table I–was evaluated using two confusion matrices. Each matrix compares the true NO_2 measurements above and below the threshold with the corresponding model forecasts over the 2020-2022 period. The accuracy

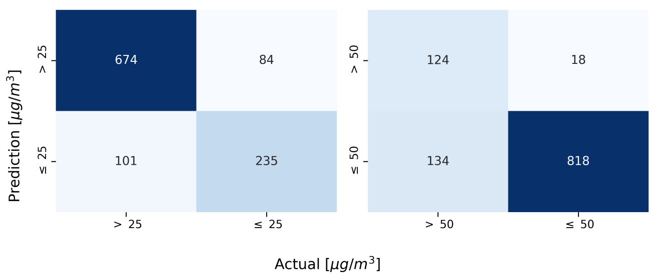

Fig. 2. Confusion matrices showing the model's ability to predict exceedances of daily NO_2 thresholds of 25μg/m^3 (left) and 50μg/m^3 (right) over the 2020–2022 period. Rows correspond to predicted classes, while columns represent observed (ground-truth) classes.

of threshold exceedance detection was evaluated using the hit ratio (HR) and false alarm ratio (FAR). HR represents the proportion of actual exceedances correctly identified by the model, while FAR indicates the proportion of predicted exceedances that did not actually occur. Table V shows that

TABLE V

HIT RATIO AND FALSE ALARM RATIO FOR NO_2 DAILY THRESHOLD EXCEEDANCES OVER THE 2020-2022 PERIOD.

Metric	25 μg/m^3	50 μg/m^3
HR	87.0%	48.1%
FAR	11.1%	12.7%

the model performs well in detecting exceedances of the lower 25μg/m^3 threshold, while performance decreases for the higher 50μg/m^3 threshold, where only about half of the actual exceedances are correctly identified. The false alarm ratio remains comparable in both cases, indicating a similar tendency toward over-forecasting of exceedance events. This behavior is consistent with the distribution of observed NO_2 concentrations, as illustrated in Figure 2, where most samples fall between the two thresholds (i.e., above 25μg/m^3 and below 50μg/m^3). This class imbalance is a well-known factor affecting the performance of statistical and machine learning models in rare-event detection problems, often leading to reduced sensitivity at higher thresholds where exceedances are less frequent. As a result, the model shows better performance at the lower threshold, where exceedance events are more common and therefore more effectively learned during training. In this context, the higher sensitivity achieved at the lower threshold is particularly relevant, since it corresponds to the more stringent WHO daily guideline value, designed to provide a health-oriented reference for air quality assessment, thus improving the detection of exceedances that are more critical in terms of potential health impacts.

Consequently, the model behavior was further investigated using XAI techniques based on SHAP. The analysis focused on the overall importance of the input features in the model forecasts, their relationship with the output, and the spatial regions of the NO_x emission maps that the model considered most relevant. Feature importance was evaluated as the average of the absolute Shapley values associated with each feature. This metric quantifies the extent to which each feature contributes to deviations of the model output from its baseline value, thereby reflecting its overall influence on the forecasts. Figure 3 shows the importance of the input

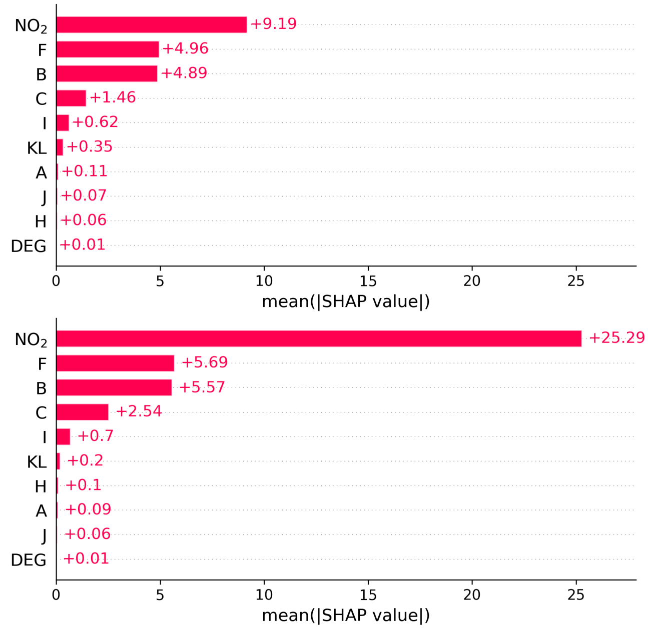

Fig. 3. Feature importance evaluated by SHAP across the entire test set (top) and for the 10% of the samples with the highest NO_2 concentration (bottom). Higher values indicate greater contribution of each feature to the model output.

features, highlighting a predominance in the autoregressive NO_2 term in driving the model output, followed by the sectors F (road transport), B (industry) and C (other stationary combustion). The importance of the emission-related features remains comparable between the entire set and the 10% portion having the highest NO_2 concentration whereas the autoregressive feature more than doubles. This increase in importance may be attributed to the strong persistence of the phenomenon, whereby the autoregressive component increasingly dominates the forecast relative to the emission inputs. However, this result should not be interpreted as evidence that controlling NO_x emissions is ineffective. Rather, it underscores the importance of limiting emissions from the outset to prevent NO_2 concentrations from becoming elevated in the first place. Regarding the emission, the model identified vehicle emissions as well as industrial and other stationary combustion sources as the most relevant features, which is consistent with known sources of NO_2

in the atmosphere. This suggests that the model successfully captured the emission components most strongly related to the predicted output.

Then, regions of the input maps where the model focuses the most were analyzed. Such measure was obtained by summing the Shapley values of each cell across the sectors and considering the average across the test set. In this way it is possible to assess the importance associated with each cell of the input maps, and thereby identify the most important regions for the model. Figure 4 shows the importance assigned by the model to each grid cell (right), together with the variability of the cells over the test set, quantified as the standard deviation of their values (left). Both panels are displayed using a color scale where lighter shades correspond to lower values, while darker shades indicate higher values. The main result is that the model correctly identifies the region surrounding Milan as the most relevant for its forecasts. In particular, the model focuses on densely populated and touristic areas near the city, such as the Como, Lecco, Iseo, and Maggiore lakes, as well as the cities of Novara and Bergamo. Conversely, areas characterized by low emission variability, such as Sondrio, Pavia, Lodi, and Cremona, or located farther from Milan, such as Mantova, are assigned little importance by the model.

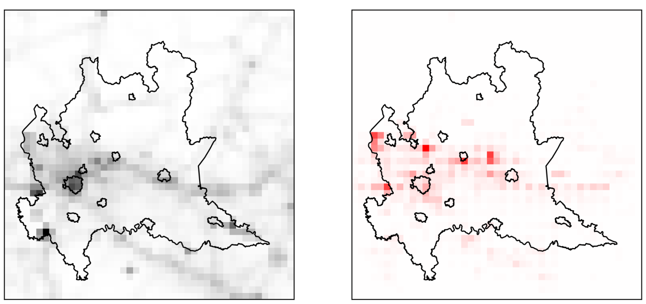

Fig. 4. Spatial distribution of SHAP-based feature importance across the NO_x emission grid (right) and corresponding spatial variability of emissions in the test set, measured as standard deviation (left). In both panels, darker colors indicate higher values, while lighter colors indicate lower values.

IV. CONCLUSIONS

In this paper, we presented a data-driven framework to forecast daily NO_2 concentrations in the city of Milan. A hybrid convolutional neural network was employed, combining two-dimensional maps of NO_x emissions across Lombardy with pointwise measurements of the previous day's NO_2 concentrations. Although NO_x includes NO_2 as one of its components, emission maps were used as a proxy for spatially distributed precursor sources, enabling the model to capture the contribution of different emission sectors and their geographical distribution.

The model was trained on data from 2010 to 2019 and evaluated over the 2020–2022 period, showing good predictive performance and a reliable ability to detect exceedances of regulatory thresholds, particularly the $25\mu g/m^3$ daily limit defined by the WHO.

Furthermore, model behavior was investigated using the explainable AI technique SHAP, providing insights into both feature relevance and spatial patterns driving the forecasts. The past NO_2 concentration emerged as the most influential input, reflecting the strong temporal persistence of the system. Emission-related features, particularly those associated with road transport and industrial activities, also showed a significant contribution. While these results are consistent with established knowledge of NO_x emission sources, the XAI analysis provides a quantitative and model-specific validation of these relationships, demonstrating that the model has learned physically meaningful patterns rather than relying on spurious correlations. In addition, the spatial analysis highlighted that the model assigns greater importance to densely populated and touristic areas surrounding the target location, suggesting that it effectively captures the spatial structure of emission-driven pollution dynamics.

The XAI assessments suggest that the proposed model could be effectively integrated into a control-oriented framework aimed at identifying optimal emission reduction pathways. Building on these promising results, future work may extend the proposed modeling approach by exploring more challenging forecasting settings, such as multi-step prediction or fully lagged input configurations, as well as by refining the spatial and sectoral resolution of emission data. In addition, dedicated training strategies aimed at improving the detection of rare high-concentration events could be investigated. Moreover, the insights provided by SHAP could be leveraged to design more efficient input representations, for instance by focusing on the most influential regions and emission sectors, thereby improving model interpretability and computational efficiency without compromising predictive performance. The applicability of the proposed methodology to different geographical locations could also be investigated; however, such an extension would require the availability of sufficiently long and reliable datasets for both model retraining and independent validation.

REFERENCES

[1] Weltgesundheitsorganisation and W. H. Organization, *WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. World Health Organization, 2021.

[2] European Environment Agency, *Harm to Human Health from Air Pollution in Europe : Burden of Disease 2023*, ser. EEA Briefing. Publications Office, 2023. [Online]. Available: https://data.europa.eu/doi/10.2800/721439

[3] ——, *Europe's Air Quality Status 2023*, ser. EEA Briefing. Publications Office, 2023. [Online]. Available: https://data.europa.eu/doi/10.2800/59526

[4] ——, *Impacts of Air Pollution on Ecosystems in Europe*, ser. Web Report. Publications Office, 2024. [Online]. Available: https://data.europa.eu/doi/10.2800/9670592

[5] ——, *Air Pollution in Europe: 2023 Reporting Status under the National Emission Reduction Commitments Directive.* Publications Office, 2023. [Online]. Available: https://data.europa.eu/doi/10.2800/339055

[6] European Parliament, Council of the European Union, "Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on ambient air quality and cleaner air for Europe (recast)," 2024. [Online]. Available: http://data.europa.eu/eli/dir/2024/2881/oj

[7] A. Bozdağ, Y. Dokuz, and Ö. B. Gökçek, "Spatial prediction of PM10 concentration using machine learning algorithms in Ankara, Turkey," *Environmental Pollution*, vol. 263, 2020.

[8] Y. Park, B. Kwon, J. Heo, X. Hu, Y. Liu, and T. Moon, "Estimating PM2.5 concentration of the conterminous United States via interpretable convolutional neural networks," *Environmental Pollution*, vol. 256, 2020.

[9] Q. Di, H. Amini, L. Shi, I. Kloog, R. Silvern, J. Kelly, M. B. Sabath, C. Choirat, P. Koutrakis, A. Lyapustin, Y. Wang, L. J. Mickley, and J. Schwartz, "Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging," *Environmental Science & Technology*, vol. 54, no. 3, pp. 1372–1384, 2020. [Online]. Available: https://doi.org/10.1021/acs.est.9b03358

[10] A. Gilik, A. S. Ogrenci, and A. Ozmen, "Air quality prediction using CNN+LSTM-based hybrid deep learning architecture," *Environmental Science and Pollution Research*, vol. 29, no. 8, pp. 11 920–11 938, 2022.

[11] R. Navares and J. L. Aznarte, "Predicting air quality with deep learning LSTM: Towards comprehensive models," *Ecological Informatics*, vol. 55, p. 101019, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574954119303309

[12] S. R. Shams, A. Jahani, S. Kalantary, M. Moeinaddini, and N. Khorasani, "The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO2 concentration," *Urban Climate*, vol. 37, 2021.

[13] J. Wei, S. Liu, Z. Li, C. Liu, K. Qin, X. Liu, R. T. Pinker, R. R. Dickerson, J. Lin, K. F. Boersma, L. Sun, R. Li, W. Xue, Y. Cui, C. Zhang, and J. Wang, "Ground-Level NO2 Surveillance from Space Across China for High Resolution Using Interpretable Spatiotemporally Weighted Artificial Intelligence," *Environmental Science & Technology*, vol. 56, no. 14, pp. 9988–9998, 2022. [Online]. Available: https://doi.org/10.1021/acs.est.2c03834

[14] B. Zhang, H. Zhang, G. Zhao, and J. Lian, "Constructing a PM2.5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks," *Environmental Modelling & Software*, vol. 124, p. 104600, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364815219300192

[15] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[16] J. Klaise, A. V. Looveren, G. Vacanti, and A. Coca, "Alibi explain: Algorithms for explaining machine learning models," *Journal of Machine Learning Research*, vol. 22, no. 181, pp. 1–7, 2021. [Online]. Available: http://jmlr.org/papers/v22/21-0017.html

[17] M. T. Ribeiro, S. Singh, and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 1135–1144. [Online]. Available: https://dl.acm.org/doi/10.1145/2939672.2939778

[18] D. Gunning and D. W. Aha, "DARPA's Explainable Artificial Intelligence Program," *AI Magazine*, vol. 40, no. 2, pp. 44–58, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1609/aimag.v40i2.2850

[19] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)," 2024. [Online]. Available: http://data.europa.eu/eli/reg/2024/1689/oj

[20] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," 2022. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137.locale=en

[21] European Environment Agency. Download Service. European Envirnomental Agency. [Online]. Available: https://eeadmz1-downloads-webapp.azurewebsites.net/

[22] M. Guevara, O. Jorba, C. Tena, H. Denier Van Der Gon, J. Kuenen, N. Elguindi, S. Darras, C. Granier, and C. Pérez García-Pando, "Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-TEMPO): Global and European emission temporal profile maps for atmospheric chemistry modelling," *Earth System Science Data*, vol. 13, no. 2, pp. 367–404, 2021. [Online]. Available: https://essd.copernicus.org/articles/13/367/2021/

[23] J. Kuenen, S. Dellaert, A. Visschedijk, J.-P. Jalkanen, I. Super, and H. Denier Van Der Gon, "CAMS-REG-v4: A state-of-the-art high-resolution European emission inventory for air quality modelling," *Earth System Science Data*, vol. 14, no. 2, pp. 491–515, 2022. [Online]. Available: https://essd.copernicus.org/articles/14/491/2022/

[24] X. Wang, S. Yang, J. Zhang, M. Wang, J. Zhang, W. Yang, J. Huang, and X. Han, "Transformer-based unsupervised contrastive learning for histopathological image classification," *Medical Image Analysis*, vol. 81, p. 102559, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841522002043

[25] M. Sajjad, Z. A. Khan, A. Ullah, T. Hussain, W. Ullah, M. Y. Lee, and S. W. Baik, "A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting," *IEEE Access*, vol. 8, pp. 143 759–143 768, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9141253

[26] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667096820300070

[27] L. Ding, W. Fang, H. Luo, P. E. D. Love, B. Zhong, and X. Ouyang, "A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory," *Automation in Construction*, vol. 86, pp. 118–124, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580517302650

[28] K. Choi, G. Fazekas, M. Sandler, and K. Cho, "Convolutional recurrent neural networks for music classification," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 2392–2396. [Online]. Available: https://ieeexplore.ieee.org/document/7952585

[29] R. Barzegar, M. T. Aalami, and J. Adamowski, "Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model," *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 2, pp. 415–433, 2020. [Online]. Available: https://doi.org/10.1007/s00477-020-01776-2

[30] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting," *IEEE Access*, vol. 8, pp. 180 544–180 557, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9210478

[31] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999–7019, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9451544/

[32] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometrics and intelligent laboratory systems*, vol. 39, no. 1, pp. 43–62, 1997.

[33] T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi *et al.*, "Kerastuner," https://github.com/keras-team/keras-tuner, 2019.

[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017. [Online]. Available: https://arxiv.org/abs/1412.6980