THE ASIAN BULLETIN OF BIG DATA MANAGEMENT
Vol. 5. Issue 3 (2025)
https://doi.org/10.62019/hcqtrs67

ASIAN BULLETIN OF BIG DATA MANAGEMENT
http://abbdm.com/
ISSN (Print): 2959-0795
ISSN (online): 2959-0809

A Transformer-Based Framework for Accurate and Interpretable PM2.5 Forecasting through Integration of Country-Level Embeddings and Explainable AI for Enhanced Environmental Decision-Making

Syed Azeem Inam*, Hassan Hashim, Asif Mehmood Awan, Syeda Wajiha Naim, Haider Rajput, Saddam Umer

Chronicle

Article history
Received: July 14, 2025
Received in the revised format: July 30, 2025
Accepted: Aug 24, 2025
Available online: Sept 18, 2025

Syed Azeem Inam*, Hassan Hashim, Asif Mehmood Awan, Haider Rajput, and Saddam Umer are currently affiliated with the Department of Artificial Intelligence and Mathematical Sciences, Sindh Madressatul Islam University, Karachi, Pakistan.
Email: syed.azeem@smiu.edu.pk
Email: hassan.hashim@smiu.edu.pk
Email: asif.awan@smiu.edu.pk
Email: haiderrajput.ai@gmail.com
Email: sadamumer566@gmail.com

Syeda Wajiha Naim is affiliated with the Department of Software Engineering, Sindh Madressatul Islam University, Karachi, Pakistan.
Email: swnaim@smiu.edu.pk

Corresponding author*

Abstract

Air quality prediction is increasingly vital to environmental health, particularly for urban and rural regions, where dangerous particulate matter hazards often exist. This study presents a novel methodology for estimating PM2.5 concentrations using a transformer model with country-level embedding and explainable AI (XAI). The proposed approach is superior to conventional machine learning and deep learning techniques as it provides high accuracy and is interpretable and applicable to various geographic regions. Given the country-specific embeddings, the transformer architecture models the replacements caused by time and location variations in pollutants' concentrations, consequently allowing accurate predictions even for regions having sparse data. Furthermore, SHAP and LIME elucidate the model's tendency to predict, providing policymakers with valuable insights. Overall, the proposed architecture presents a stronger predictive power than other forecasting models, with an R-squared value of 0.98 and a mean absolute error of 0.011. Also, using country embedding has helped improve accuracy and the ability to apply to different regions. Hence, this research offers a plausible framework to forecast air pollution and evidence-based government policymaking and planning about air pollution and its health and environmental effects.

Keywords: Transformer Architecture, Explainable AI, Environmental Management, Pollution Forecasting, PM2.5 Concentration

© 2025 The Asian Academy of Business and Social Science Research Ltd, Pakistan.

INTRODUCTION

Forecasting the concentration of fine Particulate Matter (PM) of diameter less than 2.5 μm commonly produced from combustion, industrial emissions, and secondary atmospheric reactions has become indispensable for environmental management, connecting scientific advances with policy interventions to protect ecosystems and human health (Inam, Iqbal, et al. 2024; Inam, Khan, et al. 2025; Rahujo et al. 2025). Once airborne, these particles travel long distances and deposit heavy metals and organic contaminants onto soil and water bodies. Accumulation of PM2.5-bound toxins in soil alters pH and nutrient availability, disrupting microbial communities and inhibiting plant growth. Similarly, in aquatic systems, these particles impair water quality, bioaccumulate in organisms, and compromise the food network. Because air pollution transcends regional boundaries, a forecasting framework must be highly accurate and interpretable to guide emission regulations, land-use planning, and water-quality safeguards, thereby ensuring ecosystem resilience (Fang et al., 2019; J. He et al., 2023; Inam, Zaidi, et al., 2025; Martinez et al., 2024; Onwudiegwu et al., 2025; Zhakypbek et al., 2024). Early PM2.5 forecasting efforts relied on statistical methods of autoregressive integrated moving average (ARIMA), moving-average techniques, and linear regression that assumed stationarity and linear pollutant interactions. Although these models provided baseline predictions, they oscillate when data is incomplete or pollutant dynamics are non-linear. With the advent of machine learning, algorithms such as support vector regression (SVR), decision-tree regressors, random forest, and gradient-boosting methods like XGBoost captured non-linear relationships and achieved R^2 scores typically between 0.75 and 0.85 in urban environments, but struggled with missing values, shifting emission patterns, and generalizing across heterogeneous regions (Mathew et al., 2023; Vignesh et al., 2023). The emergence of deep learning marked a paradigm shift.

Convolutional neural networks (CNNs) initially extracted spatial features from gridded pollutant data, while recurrent neural networks (RNNs), notably Long Short-Term Memory (LSTM) and bidirectional LSTM (BiLSTM), modeled temporal dependencies. Hybrid CNN-LSTM architectures achieved R^2 values exceeding 0.90, and BiLSTM variants further captured diurnal cycles and seasonal trends. However, RNNs require sequential processing, which limits parallelization, and suffer from vanishing gradient issues, constraining their ability to learn long-range dependencies. Moreover, these deep models operate as black boxes, offering little insight into how meteorological drivers, precursor pollutant concentrations, and land-use characteristics influence forecasts (Esager & Ünlü, 2023; Z. He & Guo, 2024). Initially designed for natural language processing, the transformer architecture revolutionized time-series forecasting by replacing recurrence with self-attention. This mechanism processes entire input sequences simultaneously,

capturing long-range dependencies without recursive loops and enabling parallel computation. In PM2.5 forecasting, sparse-attention transformer networks achieved an R^2 score of 0.937 with an RMSE value of 19 μg/m³ on Beijing data, outperforming LSTM baselines. Temporal Fusion Transformers (TFTs) achieved an R^2 score of 0.97 and an RMSE of 4.2 μg/m³ in multi-day air-quality projections. Graph-enhanced transformers weave spatial adjacency directly into their attention scores, which sharpens forecasts in areas shaped by rugged terrain and clustered emission sources. The drawback is that these models still behave like black boxes, a problem in fields where decisions demand visible reasoning (Rai et al., 2023; Rath & P, 2025; Zhang & Zhang, 2023). To really understand what these models are doing, researchers are leaning on explainable AI tools like SHAP and LIME. SHAP uses ideas from game theory to break down how much each feature contributes overall. Hence, you get a global view of what drives predictions, whereas LIME works differently; it builds a simple, temporary model around a single case to show which factors matter most for that specific forecast.

Combined, SHAP and LIME render transformer forecasts far less of a riddle. They can demonstrate, for example, how temperature, humidity, wind, or precursor pollutants can initiate the increase or decrease in the concentration of PM2.5 (Aldughayfiq et al., 2023; Gaspar et al., 2024; Roshinta & Gábor, 2024; Salih et al., 2024). Such clarity ensures that people are more disposed to believe the results and can satisfy regulatory requirements where explainability is not a luxury. The comparison of different state-of-the-art PM2.5 concentration models is provided in Table 1. The present study proposes a framework integrating transformer architecture with country-level embedding vectors and XAI tools. Country embedding encodes regional emission inventories, meteorological norms, and land-use typologies into dense feature representations, augmenting time-series inputs to capture cross-regional variability. It further addresses data scarcity in regions lacking dense monitoring networks. By learning shared pollutant and meteorological patterns across countries with similar environmental profiles, embedding vectors transfers knowledge from data-rich to data-sparse areas, enhancing forecast reliability and informing transboundary cooperation on air-quality management.

Table 1.
Characteristics of Various State-of-the-art Models for PM2.5 Forecasting

Model	R^2 Score	Model Type	Key Characteristic
Sparse Attention Transformer (STN) (Zhang & Zhang, 2023)	0.9370	Transformer-based	Employs sparse self-attention
ResInformer (Al-qaness et al., 2023)	Not Specified	Transformer-based	Incorporates residual Informer blocks
Temporal Fusion Transformer (TFT) (Rath & P, 2025)	0.9700	Transformer-based	Applies multi-horizon attention layers
Graph Transformer (TDGTN) (Zhang et al., 2022)	Not Specified	Graph Transformer	Embeds spatial adjacency in attention
BiConvLSTM + STA (Lakshmi & Krishnamoorthy, 2024)	0.9686	Hybrid	Combines convolutional and recurrent layers

Moreover, the transformer's parallel self-attention offers computational efficiency suitable for real-time forecasting across multiple monitoring stations, equipping decision-makers with timely predictions essential for early warnings, air-quality alerts, and adaptive emission regulations. Also, SHAP's global explanations highlight the dominant drivers of PM2.5 variability. At the same time, LIME's localized surrogate models reveal the factors precipitating specific high-pollution episodes, which allow regulatory agencies to target interventions precisely when and where they will be most effective. Table 2 presents the challenges and limitations in the present state-of-the-art models, along with the solutions provided by the proposed architecture.

Table 2.
Challenges/Limitations of the State-of-the-art Models and their Solution with the Proposed Architecture

Challenge / Limitation	Traditional Weakness	Transformer + XAI Solution
Interpretability	Black-box models hinder trust	SHAP and LIME furnish global and local explanations
Long-term Dependencies	RNNs face vanishing gradients	Self-attention captures dependencies in parallel
Cross-regional Generalization	Models trained on specific regions	Country embedding enables adaptation across geographies
Feature Importance Attribution	Lacks transparent attribution mechanisms	SHAP values quantify each feature's contribution
Complex Temporal-Spatial Patterns	Linear or sequential models fall short	Multi-head attention discerns multivariate patterns
Policy Decision Support	Opaque rationale undermines regulatory use	XAI integration provides evidence-based transparency

Extreme Event Prediction	Poor handling of rare pollution events	Embedding and attention improve anomaly forecasting
Computational Efficiency	Sequential RNN bottlenecks	Parallel attention reduces training and inference time
Multivariate Input Integration	Limited variable scope	Attention processes multiple inputs simultaneously
Uncertainty Quantification	No inherent confidence estimates	Attention weights and SHAP dispersion indicate uncertainty

The current framework proposes an integrated forecasting framework combining substantial precision, strict explainability, and local flexibility. The proposed framework closes the gap between advanced machine learning and actionable environmental policy by combining a high-performance transformer model with country embeddings and robust XAI tools. The resulting transparent, high-fidelity PM2.5 forecasts empower policymakers to develop evidence-based strategies for minimizing soil and water contamination, preserving biodiversity, and protecting public health.

METHODOLOGY

Dataset Description

The dataset driving this study captures detailed environmental readings from various countries and regions (see Figure 1 and Figure 2). Each row represents a unique exposure event, identified by an exposure ID, and annotated with the type of exposure (i.e., ambient air, water source). Geographic context is provided through the country name, its corresponding three-letter ISO3 code, and the broader region to which it belongs. Pollutant information is split into two complementary columns: a concise code in pollutant and a descriptive label in pollutant name. At the same time, units specify how concentrations are measured, for instance, micrograms per cubic meter for air pollutants or parts per million for water contaminants.

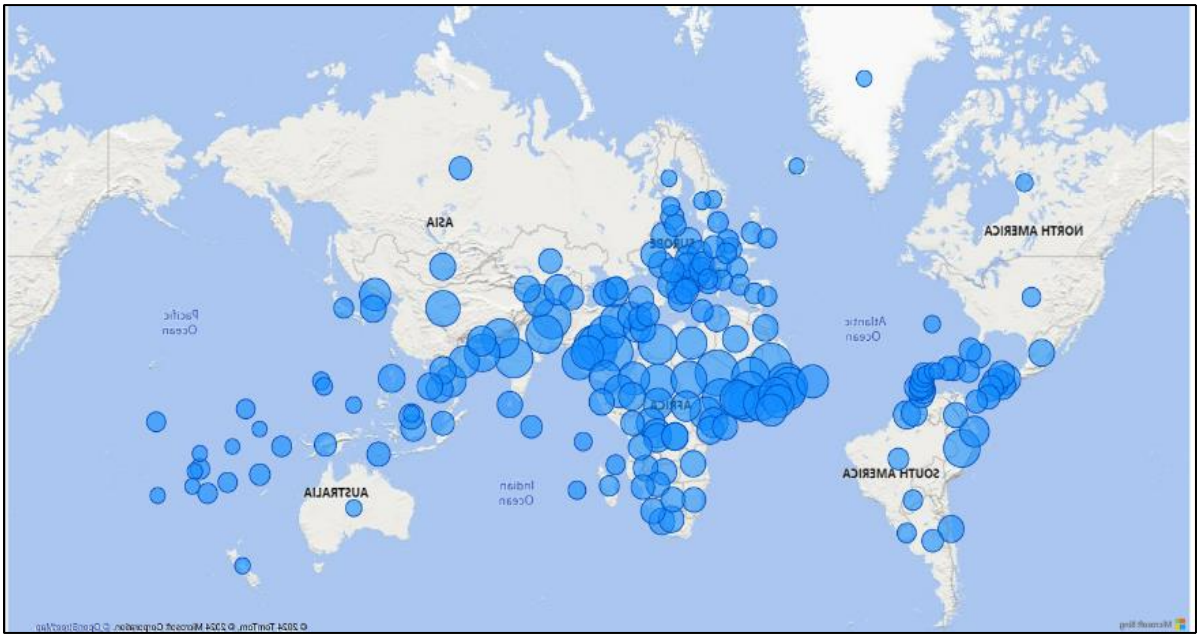

Figure 1.
PM2.5 by country in the SOGA Dataset (Inam, Khan, et al., 2024)

Figure 2.
Countries Under Consideration in the SOGA Dataset (Inam, Khan, et al., 2024)

Spanning multiple years, this rich dataset allows the study to trace pollutant trends over time and compare environmental burdens across continents. The study first loaded the CSV via pandas to prepare the data, explicitly naming columns to prevent misalignment. The study then addressed missing entries by interpolating temporal gaps for continuous pollutants and imputing with region-specific medians when values were sparse. Country identifiers were standardized to ensure that "United States," "USA," and "US" all map to the same ISO3 code. Finally, categorical features such as Pollutant and Region were label-encoded, and numeric features were scaled to a standard range to prevent any single variable from dominating the model's learning process. This careful pre-processing lays the groundwork for predictive modeling: it ensures that the spatial distribution of pollutants is faithfully represented and that the model can generalize patterns across time and geography.

Model Architecture (Transformer + Country Embedding)

At the center of the proposed approach is a transformer network enhanced by a dedicated Country Embedding layer, designed to integrate global self-attention capabilities with localized environmental context. The rationale is straightforward: pollutant behavior follows temporal trajectories (seasonal cycles, long-term trends) and spatial idiosyncrasies (industrial concentration, regulatory frameworks). The study allows the model to internalize these broader socio-environmental factors alongside raw pollutant readings by embedding each country into a continuous vector space.

Country Embedding Layer: Each ISO3 code is mapped to a learned 128-dimensional vector. These embeddings evolve during training to reflect similarities and differences between countries, so, for example, two neighboring nations with comparable air quality profiles will acquire closely aligned vectors.

Transformer Encoder Stack: The study employs six layers of the classic transformer encoder, where, within each layer:

- A multi-head self-attention mechanism (8 heads, 512-dimensional hidden states) dynamically weights parts of the input sequence, enabling the model to focus on critical time points or related pollutants.

- A position-wise feedforward network refines these attention outputs, capturing non-linear interactions.

- Residual connections and layer normalization ensure stable gradients and robust learning over 100 training epochs.

Feature Fusion and Output: The resulting pollutant sequence of the last transformer layer is concatenated with the country embedding. Such integrated feature representation is then fed into fully connected layers, which either regress (to predict concentration levels) or classify (to categorize risk levels), depending on the task. Moreover, the training takes advantage of (see Table 3) the batch size of 64, and a dropout rate of 0.1 is employed over the entire training to prevent overfitting. The researchers use the Adam optimizer and a learning rate of 0.001, a compromise between stability and speed of convergence. In practice, this hybrid structure is most effective at noticing the nuances of pollutant behavior, learning, for instance, that spikes of PM in one country can predict the same behavior in an adjacent area, while still reflecting the individual environmental fingerprint of that country.

Table 3.
Hyperparameter Configuration of the Proposed Model Architecture

Hyperparameter	Value
Number of Transformer Layers	6
Hidden Dimension	512
Number of Attention Heads	8
Dropout Rate	0.1
Learning Rate	0.001
Batch Size	64
Epochs	100
Embedding Dimension	128

RESULTS

Many models were tested in the experiments (see Table 4), and it is evident that the transformer using country embedding is the best-performing model. It had the lowest mean squared error of 0.000469, the lowest root mean squared error of 0.021663, and a mean absolute error of 0.11236 with an R-squared of 0.982513. These values show that the model fits most of the variance in the target and leaves the residuals small. The second most powerful baseline is the plain random forest regressor with slightly worse MSE and RMSE, reporting R-squared of 0.980301. Attempts to stack or hybridize the random forest with fully connected neural networks or recurrent units did not yield gains and, in fact, led to small degradations. XGBoost-based variants performed competitively but stayed behind the top two. At the same time, linear and polynomial regressor families, even when combined with recurrent layers, remained clustered around an R-squared near 0.977 with larger errors. Classical deep learning models such as LSTM, BiLSTM, and feedforward networks performed poorly, producing RMSE around 0.16 and R-squared close to zero. PINN and mismatched polynomial multilayer perceptron combinations diverged with extreme errors, confirming they are unsuitable for this forecasting.

Table 4.
Performance Evaluation of Experiment Models for PM2.5 Forecasting

S. No.	Algorithm	MSE	RMSE	MAE	R^2
1.	Transformer + Country Embedding	0.000469	0.021663	0.011236	0.982513
2.	Random Forest Regressor	0.000529	0.022992	0.009216	0.980301
3.	Random Forest Regressor + FCNN	0.000544	0.023332	0.010441	0.979715
4.	Random Forest Regressor + LSTM	0.000555	0.023552	0.010277	0.979330
5.	XGB Regressor	0.000575	0.023978	0.015691	0.978575
6.	Linear regression + BiLSTM	0.000616	0.024820	0.016345	0.977045
7.	XGBoost Regressor + FCNN	0.000616	0.024820	0.016345	0.977045
8.	Polynomial regression + BiLSTM	0.000616	0.024820	0.016345	0.977045
9.	Random Forest Regressor + LSTM + BiLSTM	0.000664	0.025767	0.016091	0.975258
10.	Decision Tree Regressor	0.000824	0.028703	0.011105	0.969301
11.	Transformer	0.002313	0.048094	0.042620	0.913807
12.	Extra Trees Regressor + LSTM + BiLSTM	0.003509	0.059233	0.035735	0.869258
13.	Extra Trees Regressor	0.003678	0.060649	0.035796	0.862931
14.	Extra Trees Regressor + FCNN	0.003682	0.060680	0.035834	0.862792
15.	LSTM + BiLSTM (Tuned)	0.024809	0.157510	0.114866	0.075507
16.	LSTM + BiLSTM	0.025538	0.159807	0.119532	0.048346
17.	LSTM	0.025844	0.160760	0.120017	0.036959
18.	FCNN	0.025929	0.161025	0.116842	0.033786
19.	BiLSTM	0.026370	0.162389	0.121608	0.017345
20.	Linear regression + MLP	0.026606	0.163113	0.126412	0.008562
21.	PINN	952.6554	30.86511	22.69124	0.002721
22.	Polynomial Regression	0.026820	0.163767	0.120827	0.000604
23.	Linear Regression	0.026878	0.163946	0.120913	-0.001590
24.	SVR	0.027550	0.165981	0.116921	-0.026599
25.	Poly Regression + MLP	8.832946	2.972027	2.534358	-328.1485

The cross-validation curves (see Figure 3) confirm the stability of the proposed architecture. Across five folds, the mean absolute error and root mean squared error fluctuate slightly around their fold means with no sign of overfitting. The mean squared error follows the same pattern, and the R-squared remains consistently high across folds, hovering near 0.98. The fifth fold shows the lowest error and the highest R-squared, but the spread is small, supporting the average metrics' reliability.

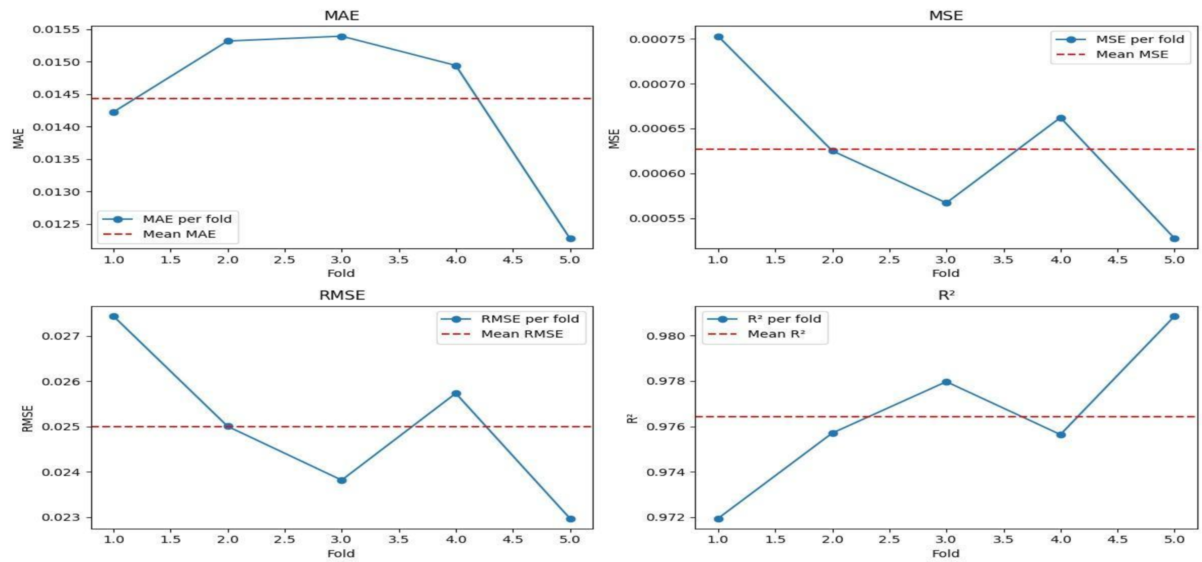

Figure 3.
Cross-Validation Scores of the Transformer + Country Embedded Architecture

Local interpretability with LIME (see Figure 4) highlights two drivers. “Year” greater than 0.77 contributes a negative adjustment to the prediction, while the condition that the encoded country lies between 99 and 151 pushes the prediction upward. The magnitudes show that the country-related signal slightly outweighs the year effect for this case, resulting in a modest positive net contribution.

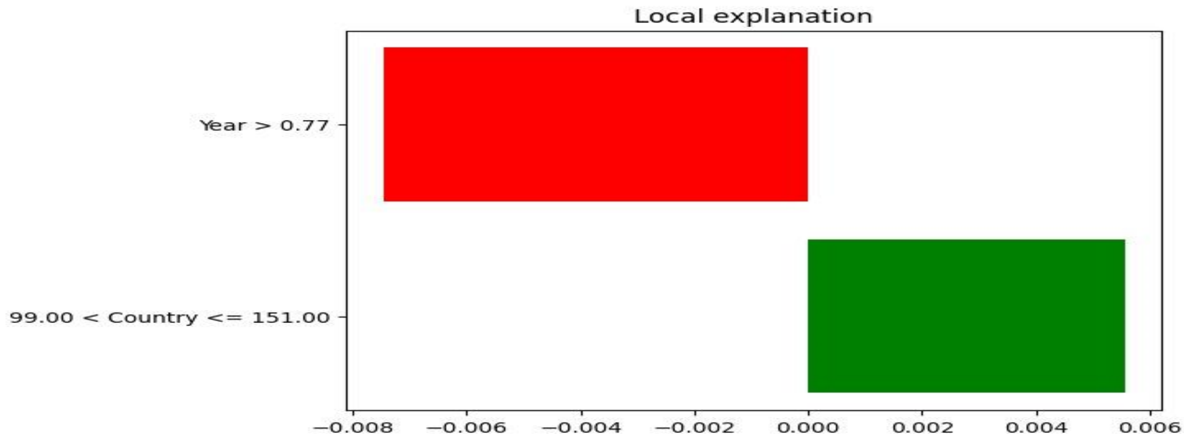

Figure 4.
LIME Interpretation of the Proposed Transformer + Country Embedding Architecture

SHAP values (see Figure 5) have the widest country distribution, which suggests that the country is the most influential feature in the dataset. Both low and high-country embeddings can shift predictions in either direction, and the shifts can be moderate in both directions with many observations. “Year” shows a SHAP spread much more concentrated around zero, indicating a more consistent but weaker effect. A lower value of the “Year” is more likely to decrease the output slightly, and a higher value is more likely to increase the output; however, the effect is negligible compared to that of the country. Introducing a learned country representation enables the transformer architecture to gain a country-specific structure not found in traditional models or plain transformers. This explains the improvement of the error metrics and the robustness of the cross-validation.

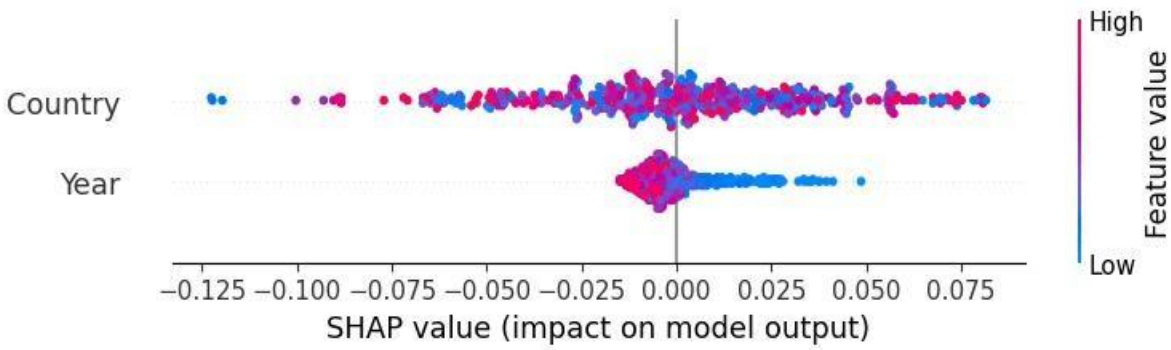

Figure 5.
SHAP Summary of the Proposed Transformer + Country Embedding Architecture

CONCLUSION

Transformer networks with country-specific embedding vectors and developing explainable AI approaches are valuable additions to PM2.5 concentration forecasting. The study demonstrates that the quality of the model offered is relatively high in predicting the air quality, as the R-squared is close to 1, and the mean absolute error is low. Moreover, SHAP and LIME explications address a major bottleneck of applying machine learning to a regulatory framework, which is necessary to make the decision-making procedure transparent. The results demonstrate the importance of using advanced machine learning and explainable AI to produce credibility and policy recommendations. Furthermore, this method can, in principle, be applied to a broad range of other environmental prediction applications, which were previously challenging to control through the prediction and mitigation of complex eco-problems at runtime.

DECLARATIONS

Acknowledgement: We appreciate the generous support from all contributors to research and their various affiliations.
Funding: No public, private, or nonprofit funding body provided a particular grant for this research.
Availability of data and material: The data sources for the variables are stated in the approach.
Authors' contributions: Each author participated equally in creating this work.
Conflicts of Interest: The authors declare no conflict of interest.
Consent to Participate: Yes
Consent for publication and Ethical approval: Because this study does not include human or animal data, ethical approval is not required for publication. All authors have given their consent.

REFERENCES

Abaidullah, A., & Basheer, M. F. (2024). Nexus among Entrepreneurial Activities, Human Capital, and Economic Growth to achieve Sustainable Development Goals (SDGs): Moderating Role of Financial Development. Journal of Finance and Accounting Research, 6(1), 1-27.

Aldughayfiq, B., Ashfaq, F., Jhanjhi, N. Z., & Humayun, M. (2023). Explainable AI for Retinoblastoma Diagnosis: Interpreting Deep Learning Models with LIME and SHAP. *Diagnostics*, *13*(11), 1932. https://doi.org/10.3390/diagnostics13111932

Al-qaness, M. A. A., Dahou, A., Ewees, A. A., Abualigah, L., Huai, J., Abd Elaziz, M., & Helmi, A. M. (2023). ResInformer: Residual Transformer-Based Artificial Time-Series Forecasting Model for PM2.5 Concentration in Three Major Chinese Cities. *Mathematics*, *11*(2), 476. https://doi.org/10.3390/math11020476

Basheer, M. F., Sabir, S. A., & Hassan, S. G. (2024). Financial development, globalization, energy consumption, and environmental quality: Does control of corruption matter in South Asian countries?. Economic Change and Restructuring, 57(3), 112.

Esager, M. W. M., & Ünlü, K. D. (2023). Forecasting Air Quality in Tripoli: An Evaluation of Deep Learning Models for Hourly PM2.5 Surface Mass Concentrations. *Atmosphere*, *14*(3), 478. https://doi.org/10.3390/atmos14030478

Fang, D., Chen, B., Hubacek, K., Ni, R., Chen, L., Feng, K., & Lin, J. (2019). Clean air for some: Unintended spillover effects of regional air pollution policies. *Science Advances*, *5*(4). https://doi.org/10.1126/sciadv.aav4707

Gaspar, D., Silva, P., & Silva, C. (2024). Explainable AI for Intrusion Detection Systems: LIME and SHAP Applicability on Multilayer Perceptron. *IEEE Access*, *12*, 30164–30175. https://doi.org/10.1109/ACCESS.2024.3368377

He, J., Chen, J., Xiao, J., Zhao, T., & Cao, P. (2023). Defining Important Areas for Ecosystem Conservation in Qinghai Province under the Policy of Ecological Red Line. *Sustainability*, *15*(6), 5524. https://doi.org/10.3390/su15065524

He, Z., & Guo, Q. (2024). Comparative Analysis of Multiple Deep Learning Models for Forecasting Monthly Ambient PM2.5 Concentrations: A Case Study in Dezhou City, China. *Atmosphere*, *15*(12), 1432. https://doi.org/10.3390/atmos15121432

Huo, S., Ni, L., Basheer, M. F., Al-Aiban, K. M., & Hassan, S. G. (2024). The role of fintech, mineral resource abundance, green energy and financial inclusion on ecological footprint in

E7 countries: New insight from panel nonlinear ARDL cointegration approach. Resources Policy, 94, 105083.

Inam, S. A., Iqbal, D., Hashim, H., & Khuhro, M. A. (2024). An Empirical Approach Towards Detection of Tuberculosis using Deep Convolutional Neural Network. International Journal of Data Mining, Modelling and Management, 16(1), 101–112. https://doi.org/10.1504/IJDMMM.2024.136232

Inam, S. A., Khan, A. A., Ahmed, N., Mazhar, T., Shahzad, T., Khan, S., Saeed, M. M., & Hamam, H. (2025). A Novel Deep Learning Approach for Investigating Liquid Fuel Injection in Combustion System. Discover Artificial Intelligence, 5(1), 32. https://doi.org/10.1007/s44163-025-00248-2

Inam, S. A., Khan, A. A., Mazhar, T., Ahmed, N., Shahzad, T., Khan, M. A., Saeed, M. M., & Hamam, H. (2024). PR-FCNN: A Data-Driven Hybrid Approach for Predicting PM2.5 Concentration. *Discover Artificial Intelligence*, *4*(1), 75. https://doi.org/10.1007/s44163-024-00184-7

Inam, S. A., Zaidi, S. M. H., Khan, A. A., & Ullah, S. (2025). A Neural Network Approach to Carbon Emission Prediction in Industrial and Power Sectors. *Discover Applied Sciences*, *7*(6), 640. https://doi.org/10.1007/s42452-025-07257-x

Lakshmi, S., & Krishnamoorthy, A. (2024). Effective Multi-Step PM2.5 and PM10 Air Quality Forecasting Using Bidirectional ConvLSTM Encoder-Decoder With STA Mechanism. *IEEE Access*, *12*, 179628–179647. https://doi.org/10.1109/ACCESS.2024.3509142

Li, J., Hu, L., & Basheer, M. F. (2024). Linking green perceived value and green brand loyalty: a mediated moderation analysis of green brand attachment, green self-image congruity, and green conspicuous consumption. Environment, Development and Sustainability, 26(10), 25569-25587.

Martinez, N. E., Canoba, A., Donaher, S. E., Garnier-Laplace, J., Kinase, S., Mayall, A., Stark, K., & Whicker, J. (2024). An introduction to ecosystem services for radiological protection. *Annals of the ICRP*, *53*(1_suppl), 246–254. https://doi.org/10.1177/01466453241283931ah

Mathew, A., Gokul, P. R., Raja Shekar, P., Arunab, K. S., Ghassan Abdo, H., Almohamad, H., & Abdullah Al Dughairi, A. (2023). Air quality analysis and PM 2.5 modelling using machine learning techniques: A study of Hyderabad city in India. *Cogent Engineering*, *10*(1). https://doi.org/10.1080/23311916.2023.2243743

Onwudiegwu, C., Nabebe, G., & Izah, S. C. (2025). Environmental and Public Health Implications of Pesticide Residues: From Soil Contamination to Policy Interventions. *Greener Journal of Biological Sciences*, *15*(1), 1–12. https://doi.org/10.15580/gjbs.2025.1.120424187

Qadir, F., Basheer, M. F., & Chaudhry, S. (2024). Transgender Entrepreneurs are Paving the Path of Social Entrepreneurship: Exploring Motivators of Entrepreneurial Intent. Journal of Business and Management Research, 3(3), 785-812.

Rahujo, A., Atif, D., Inam, S. A., Khan, A. A., & Ullah, S. (2025). A Survey on the Applications of Transfer Learning to Enhance the Performance of Large Language Models in Healthcare Systems. Discover Artificial Intelligence, 5(1), 90. https://doi.org/10.1007/s44163-025-00339-0

Rai, V., Kumar, S., Singh, T., & Kapoor, R. (2023). PM2.5 Level Forecasting using Transformer-Based Model. *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 216–221. https://doi.org/10.1109/ICACITE57410.2023.10182822

Rath, S., & P, M. (2025). Air Pollution Forecasting Using Machine Learning with Temporal Fusion Transformer and Graph Neural Networks. *2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 1798–1803. https://doi.org/10.1109/IDCIOT64235.2025.10914714

Roshinta, T. A., & Gábor, S. (2024). A Comparative Study of LIME and SHAP for Enhancing Trustworthiness and Efficiency in Explainable AI Systems. *2024 IEEE International Conference on Computing (ICOCO)*, 134–139. https://doi.org/10.1109/ICOCO62848.2024.10928183

Salih, A., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Menegaz, G., & Lekadir, K. (2024). *A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME*. https://doi.org/10.1002/aisy.202400304

Vignesh, P. P., Jiang, J. H., & Kishore, P. (2023). Predicting PM 2.5 Concentrations Across USA Using Machine Learning. *Earth and Space Science*, *10*(10). https://doi.org/10.1029/2023EA002911

Yin, X., Khan, A. J., Basheer, M. F., Iqbal, J., & Hameed, W. U. (2025). Green human resource management: a need of time and a sustainable solution for organizations and environment. Environment, Development and Sustainability, 27(1), 1379-1400.

Zhakypbek, Y., Kossalbayev, B. D., Belkozhayev, A. M., Murat, T., Tursbekov, S., Abdalimov, E., Pashkovskiy, P., Kreslavski, V., Kuznetsov, V., & Allakhverdiev, S. I. (2024). Reducing Heavy Metal Contamination in Soil and Water Using Phytoremediation. *Plants*, *13*(11), 1534. https://doi.org/10.3390/plants13111534

Zhang, Z., & Zhang, S. (2023). Modeling air quality PM2.5 forecasting using deep sparse attention-based transformer networks. *International Journal of Environmental Science and Technology*, *20*(12), 13535–13550. https://doi.org/10.1007/s13762-023-04900-1

Zhang, Z., Zhang, S., Zhao, X., Chen, L., & Yao, J. (2022). Temporal Difference-Based Graph Transformer Networks For Air Quality PM2.5 Prediction: A Case Study in China. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.924986

2025 by the authors; The Asian Academy of Business and social science research Ltd Pakistan. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).