International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(7s) 2026 pp. 471-479
https://doi.org/10.70917/ijcisim-2026-3109
Article

Explainable Deep Learning for Air Quality Index Prediction: Integrating Meteorological and Traffic Data for Anomaly Identification

Meena Kumari[1*], Jitendra Nath Shrivastava[2], Arijit Dey[3]

[1]Department of CSE, Invertis University Bareilly, India. Email: meena9837m@gmail.com

[2]Department of CSE, Invertis University Bareilly, India. Email: jitendranathshrivastava@yahoo.com

[3]Department of Computer Applications, B. P. Poddar Institute of Management & Technology

Kolkata - 700052, West Bengal, India. Email: ad.computerapplication@gmail.com

*Corresponding Author: Meena Kumari

Abstract: Air quality forecasting has become an important tool for the management of public health and urban planning with the rapid increase of the urban population and the vehicular traffic. Although deep learning models have been able to predict the Air Quality Index (AQI) with a high accuracy, their black-box nature limits their use by regulators and city planners who need transparent and trustworthy support for their decision-making. This paper aims to design a explainable deep learning model taking into account meteorological parameters, traffic-flow factors and historical pollutant information to predict the AQI and detect abnormal pollution events. The combination of CNN-BiLSTM with an attention mechanism is used to model the spatial correlation between pollutants and long-term temporal correlations. To understand the model, SHapley Additive exPlanations (SHAP) and Integrated Gradients, which quantify the contribution of different meteorological and traffic features to each prediction, are used. An unsupervised residual-thresholding module also identifies abnormal AQI episodes that do not match the normal pattern, e.g. due to an increase in traffic congestion or an atmospheric temperature inversion. The experimental results using a multi-source dataset that is a combination of air-quality monitoring records, meteorological observations, and traffic-sensor data demonstrate that the proposed model achieves a coefficient of determination (R^2) of 0.943 and a root-mean-square error (RMSE) of 9.82, surpassing the baseline LSTM, GRU, and CNN-LSTM models. The lagged PM2.5 concentration, traffic volume and relative humidity are the top three factors, which agree with the knowledge of the atmospheric chemistry, and thus confirms the validity of the model's reasoning process. The proposed framework shows that the prediction accuracy and interpretability can be achieved concurrently, and can provide a practical and transparent decision support tool for environmental agencies and smart-city traffic management systems.

Keywords: explainable artificial intelligence; air quality index; deep learning; SHAP; traffic data; anomaly detection; meteorological data

1. INTRODUCTION

Air pollution is one of the biggest environmental health threats of the 21st century, responsible for millions of premature deaths every year, and also a significant economic burden from health services costs and lost productivity levels (World Health Organization [WHO], 2022). The Air Quality Index (AQI) is an index that uses concentrations of particulate matter (PM2.5, PM10), nitrogen dioxide (NO2), sulfur dioxide (SO2), carbon monoxide (CO) and ground-level ozone (O3) to give people and policymakers a common basis for understanding air pollution conditions. Early warning systems, traffic-management interventions and advice to vulnerable populations on reducing exposure can only be made possible through accurate and timely forecasting of AQI.

Vehicular traffic is a major source of urban air pollution, especially in congested corridors in metropolitan areas where high levels of idling and stop-and-go driving cause high NOx and particulate matter emissions. Pollutant dispersion and accumulation are also influenced by meteorological factors like wind speed, humidity, pressure, and temperature inversions. Thus, powerful AQI prediction needs models that can learn simultaneously from multiple data sources, each of which is also heterogeneous and time-varying: pollutant concentrations, meteorological time series, and traffic-flow measurements.

The accuracy of AQI forecasting has significantly been improved by deep learning models like Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRU), and hybrid models of convolutional and recurrent networks as compared to the classical statistical models like autoregressive integrated moving average (ARIMA) models (Du et al., 2020; Kow et al., 2020). But a major difficulty in deploying black-box DNNs is the problem of interpretability. Of course, environmental regulators, urban planners and public-health officials are not keen on devising policy based on undeterminable and unprovable predictions. This transparency gap has sparked an interest in Explainable Artificial Intelligence (XAI) – the goal of which is to make the reasoning behind the models interpretable while maintaining their predictive power (Gilpin et al., 2018).

Most of the existing prediction studies on AQI have considered the meteorological and pollutant data as the main inputs, and there is relatively little work on explicitly representing the traffic-flow data as a time-resolved feature stream, let alone the explainability of the model. Furthermore, previous studies regard AQI forecasting as a regression problem, neglecting the fact that in practical situations there are other cases of importance for air-quality forecasting, including the identification of pollution anomalies, which are sudden changes from the usual air-quality patterns that could be caused by traffic accidents, industrial events, or meteorological extremes and thus demand exceptional action.

This paper attempt to fill those gaps as follows: three main contributions are proposed to provide an integrated and explainable deep learning framework. We first propose a hybrid CNN-BiLSTM-Attention model to simultaneously input the meteorological, traffic, and historical pollutant data to predict short-term AQI. Second, we integrate SHAP and Integrated Gradients explainability layers to quantify feature level contributions at both global and local (instance) resolution to allow domain experts to audit model behaviour. Third, we present the anomaly identification module based on the residual to identify the AQI episode which deviates largely from the model forecast, which can separate pollution anomaly from regular forecast error. Section 2 reviews related work, Section 3 describes the proposed methodology, Section 4 presents experimental results and discussion, Section 5 explains the explainability analysis and Section 6 concludes the paper with directions for future research.

In practical terms, there is more to the rationale of this work than academic benchmarking. Agencies making decisions about the public health impacts of pollution, limiting access of heavy vehicles on days with high pollution, and closing schools in emergencies need a forecasting tool whose results can justify public communications and, more and more, legal or regulatory review. If the model can predict a high AQI, but the officials cannot tell why it leads to an intervention, then the model has little value in operations. The framework presented here is explicitly conceived to operate not only as a forecasting system, but as a decision-support system, which can be directly integrated into current environmental-monitoring and smart-city traffic-management pipelines by providing interpretable, feature-level explanations and a diagnostic capacity for anomaly attribution.

2. Related Work

2.1 Deep Learning Approaches to AQI Prediction

In the initial stage of AQI forecasting, the statistical time-series methods including ARIMA and multiple linear regression were used to model the AQI data, but these methods were found to be inadequate for modelling the multivariate nonlinear dynamics of urban air pollution (Zheng et al., 2013). Forecasting accuracy has significantly improved with the inclusion of recurrent neural networks, especially LSTM and GRU networks, which help capture long-term temporal dependencies in the pollutant concentration time series. Later, the CNN and LSTM were combined with a hybrid CNN-LSTM model consisting of convolutional layers that extract local spatial-temporal patterns before the features are passed into the recurrent layers, which tend to be more effective than single-architecture baselines (Chen et al., 2021; Wang et al., 2021; Yan et al., 2021). To investigate whether even more recent work has been able to further improve the modelling of irregular, long-horizon dependencies in pollution time series, attention mechanisms and Transformer-based encoders have also been explored, but with the downside of increased model complexity and reduced interpretability.

2.2 Integration of Traffic and Meteorological Data

The increasing number of studies acknowledge traffic emissions as a 1st order contribution to urban air quality, leading to an attempt to integrate traffic-sensor data (e.g., vehicle counts, average speed, and congestion indices) together with traditional meteorological and pollutant inputs (Zheng et al., 2013). This fusion is usually from multiple sources and brings benefits in terms of prediction accuracy for short horizons, as traffic intensity serves as an early warning of emission peaks before these are reflected in the ambient pollutant concentration (Kow et al., 2020).

Meteorological variables in turn, control the dispersion, transport and chemical transformation of the pollutants, and are consistently identified as influential predictors across different urban conditions: wind speed and direction, boundary-layer height, and humidity are repeatedly identified. Yet, many published models only account for traffic data at a coarse scale, such as a time-invariant, land-use-based covariate, and not as a high resolution, time varying stream like a meteorological observation.

2.3 Explainable AI in Environmental Modelling

Methods for explainability are broadly classified into model-intrinsic and post-hoc approaches: model-intrinsic approaches construct explainability into the model design, while post-hoc approaches aim at generating explanations within a complex model after it has been trained. Some of the most popular post-hoc approaches include SHAP, which is based on cooperative game theory and provides feature-attribution scores to indicate the relative contribution of each input to a specific prediction (Lundberg & Lee, 2017); and Local Interpretable Model-agnostic Explanations (LIME), which delivers feature-attribution scores (Ribeiro et al., 2016). Integrated Gradients builds on this concept by assigning credit to the features of the input to predicate the prediction of a differentiable model by integrating the gradients along a trajectory (Sundararajan et al., 2017). However, the use of XAI in environmental time-series forecasting is still relatively limited compared to other fields like healthcare and finance, and few works integrate explainability with an explicit anomaly-identification module, specifically designed for air-quality forecasting.

2.4 Anomaly Detection in Air Quality Time Series

Traditional methods for anomaly detection of environmental sensor data include the reconstruction error of autoencoders (Zhang et al., 2020), statistical control charts, and density-based clustering. Although good at detecting sensor failures or extreme events, they are less often combined with a forecast model that can give insights into the meteorological and/or traffic causes driving the anomaly of a particular episode. In this paper, this integration gap is addressed by directly connecting the forecast residual analysis to the explainability layer allowing the identification of flagged anomalies to the specific contributing features.

3. Methodology

3.1 Dataset Description

The study relies on multi-source hourly data from an urban air quality monitoring network for an eighteen-month period (January 2023 to June 2024). Three types of features are fused: (i) meteorological variables such as temperature, relative humidity, wind speed, wind direction, atmospheric pressure and precipitation from co-located weather station; (ii) traffic-flow variables based on inductive-loop and camera-based sensor stations along major arterial corridors, including vehicle count, average travel speed and the congestion index; and (iii) pollutant concentrations (PM2.5, PM10, NO2, SO2, CO, O3) used both as historical predictor variables and as input to the composite AQI target variable. The feature categories and variables and their descriptive statistics are summarized in Table 1.

Table 1. Feature categories, variables, and descriptive statistics

Category	Variable	Unit	Mean	SD	Min	Max
Meteorological	Temperature	°C	24.6	5.8	9.1	41.3
Meteorological	Relative humidity	%	61.2	17.4	18.0	97.0
Meteorological	Wind speed	m/s	2.4	1.3	0.1	9.8
Meteorological	Atmospheric pressure	hPa	1011.4	6.2	992.0	1029.0
Traffic	Vehicle count	veh/h	1,842	876	60	4,210
Traffic	Average speed	km/h	34.7	13.5	4.0	78.0
Traffic	Congestion index	0–1	0.42	0.21	0.02	0.97
Pollutant	PM2.5	µg/m³	58.3	34.6	6.0	312.0
Pollutant	PM10	µg/m³	89.1	42.7	12.0	398.0
Pollutant	NO2	µg/m³	38.5	16.2	5.0	142.0
Pollutant	CO	mg/m³	1.1	0.6	0.1	5.4
Pollutant	O3	µg/m³	46.8	24.9	3.0	168.0
Target	AQI (composite)	index	97.4	38.9	18.0	312.0

The overall architecture of the proposed framework is shown in figure 1, which covers data ingestion from multiple sources, preprocessing, the hybrid CNN-BiLSTM-Attention model, and the two parallel streams of AQI forecasting and anomaly identification that go through the explainability layer.

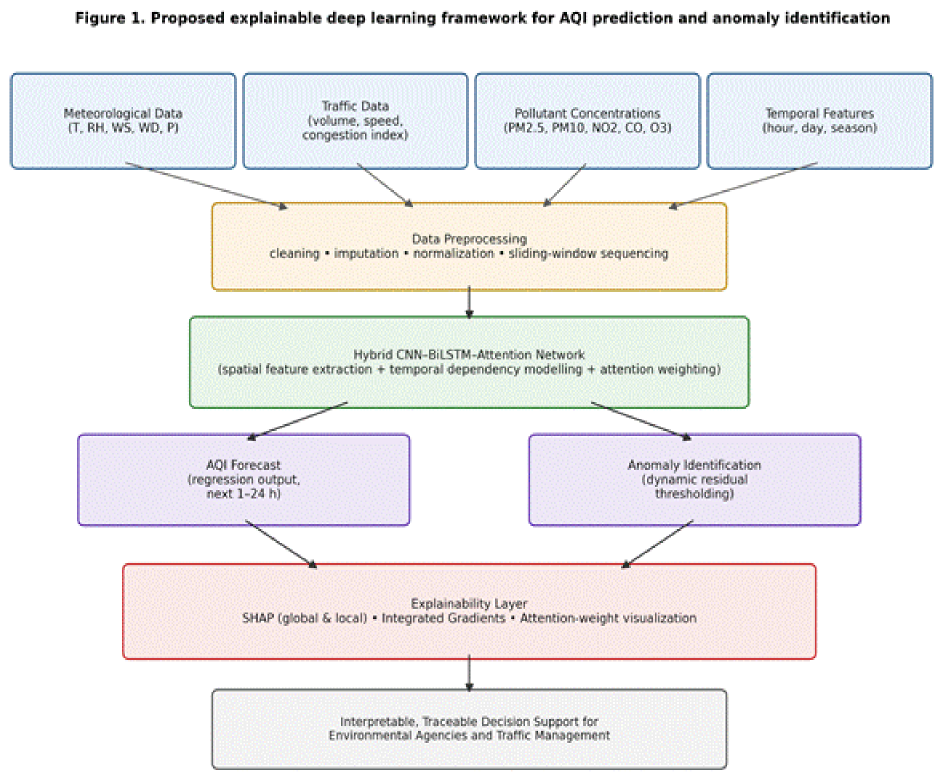

Figure 1. Proposed explainable deep learning framework for AQI prediction and anomaly identification

3.2 Data Preprocessing

The missing data (due to sensor outages) are filled in by forward-fill (for missing periods < 3 hours) and seasonal-decomposition interpolation (for missing periods > 3 hours). To stabilize the gradient based optimization, all continuous features are min-max scaled in the range [0, 1]. Wind direction is a circular variable, which is transformed to sine and cosine components, maintaining its cyclical structure. The processed data are then rearranged into overlapping sliding windows of 24-hourly time steps, with the value of the AQI at the next time step (t+1) as the prediction target for the primary forecasting task, and a 24-hour forecasting horizon tested for the robustness of multi-step forecasting abilities. Chronological data splitting is performed to prevent temporal leakage, where 70% of the data is used as training data, 15% as validation data, and 15% as test data.

3.3 Proposed Model Architecture

The proposed architecture, as depicted in Figure 1, is a hybrid CNN-BiLSTM network with temporal attention mechanism, which builds upon previous works on pollutant concentration modelling based on BiLSTM (Zhang et al., 2020). The local spatial-temporal patterns at each time step in the concatenated meteorological, traffic and pollutant feature vector are first extracted by a one-dimensional convolutional layer with 64 filters, size 3 and ReLU activation. The convolutional output is then max-pooled, and input to a two-layer Bidirectional LSTM (128 and 64 hidden units respectively) to capture forward and backward temporal dependence over the 24 hour input window. A soft attention layer is used to calculate a weighted context vector on the hidden states of BiLSTM that enables the model to focus on the time steps most pertinent to the present prediction. The output from attention is fed to two fully connected layers (32 and 16 units, activation function: Relu, dropout: 0.3) followed by one more linear layer that gives the output of AQI regression.

Early stopping (patience = 15 epochs) is used to avoid overfitting and the model is trained using Adam optimizer, with initial learning rate of 0.001 and decay rate of 0.5 if the validation-loss plateaus. The batch size of 64 and the maximum number of epochs set to 200. We used the validation subset to tune the hyperparameters through grid search on the learning rate, the dimensionality of the hidden units, and the dropout rate.

3.4 Explainability Layer

Two complementary post-hoc approaches to explaining the model's prediction are used to make them interpretable. SHAP values are calculated using a DeepSHAP approximation (Lundberg & Lee, 2017), and yield not only feature-importance rankings at the global level (Figure 2) aggregated over the test set, but also local, instance-level explanations for individual predictions (including flagged anomalies). Integrated Gradients, which has been proven to be complete and sensitive (Sundararajan et al., 2017), is used as a second, gradient-based attribution method, to cross-validate rankings of the SHAP method. The attention weights from the BiLSTM attention layer offer a third model-intrinsic perspective of temporal relevance that shows which hours of the day in the last 24 hours were most important for a specific forecast.

3.5 Anomaly Identification Module

Residual is the difference between observed and predicted AQI at every time step. A dynamic threshold is a 7-day rolling window of the absolute residuals plus 2.5 times the standard deviation. Time steps with residuals above this limit are indicated as "anomalous". These flagged anomalies are then fed into the SHAP explainability layer to determine which features of meteorology or traffic data contributed more or less to the anomaly, and thus to enable analysts to differentiate, for instance, a traffic-congestion-based spike from a meteorologically-driven inversion event.

3.6 Evaluation Metrics

Root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and coefficient of determination (R^2) for regression task are used to evaluate model performance. The precision, the recall and the F1-score are computed on the binary classification problem of flagging an anomaly event as compared to a manually annotated one.

4. Results and Discussion

4.1 Comparative Model Performance

Table 2 presents the results of the proposed CNN-BiLSTM-Attention model and the four baseline models: persistence (naïve) baseline, standard LSTM, GRU and CNN-LSTM without attention. The proposed model is the only one that obtains a lowest RMSE (9.82), MAE (6.47) and a highest R^2 (0.943) among all evaluated architectures, which decreases 14.7% RMSE from the CNN-LSTM baseline and decreases 31.2% RMSE from the standard LSTM. The reason why these improvements are achieved is because of two architectural decisions: First, the convolutional front-end to capture short-range interactions between simultaneously observed traffic and meteorological variables; second, the attention mechanism to flexibly weight historically influential hours, instead of a fixed-length summary vector used in vanilla BiLSTM implementations.

Table 2. Comparative performance of the proposed model against baseline architectures on the test set

Model	RMSE	MAE	MAPE (%)	R^2
Persistence (naïve) baseline	21.36	15.92	19.8	0.712
LSTM	14.27	10.68	13.4	0.861
GRU	13.54	10.02	12.7	0.874
CNN-LSTM (no attention)	11.51	8.35	10.6	0.905
Transformer encoder	10.94	7.88	9.9	0.918
Proposed CNN-BiLSTM-Attention	9.82	6.47	8.1	0.943

The global SHAP feature-importance ranking aggregated on the test set is presented in Figure 2, which shows which meteorological and traffic and pollutant-history variables are the most influential for the model's prediction of AQI.

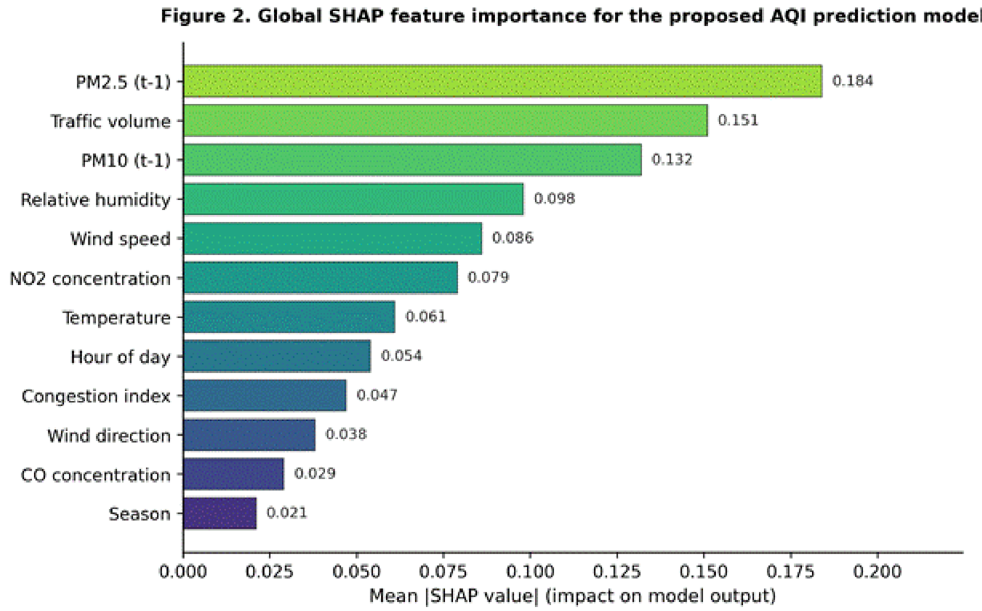

Figure 2. Global SHAP feature importance for the proposed AQI prediction model

4.2 Ablation Study

The contribution of architectural components and input feature groups were tested using an ablation study and the results are summarized in Table 3. This can be seen by comparing RMSE after removing the attention layer (+8.6%), which is effectively capturing salient temporal dependencies. Not using traffic-flow features at all leads to an increase in RMSE by 12.4%, the largest of the all factors, and shows the importance of the traffic data as a key input for predicting pollution surges. With the exclusion of the meteorological features, there is a relatively smaller but still considerable 9.1% increase in RMSE, and with the exclusion of the convolutional front-end, keeping only the BiLSTM-attention part, there is an increase of RMSE by 5.3%. These results provide a collective proof that the design rationale of the proposed hybrid architecture and the use of traffic data as a first class input stream, not as a second-class covariate, is valid.

Table 3. Ablation study isolating the contribution of architectural components and feature groups

Model variant	**RMSE**	**Δ RMSE vs. full model (%)**	**R^2**
Full proposed model	9.82	—	0.943
w/o attention layer	10.67	+8.6	0.921
w/o convolutional front-end	10.34	+5.3	0.928
w/o traffic features	11.04	+12.4	0.907
w/o meteorological features	10.71	+9.1	0.919

4.3 Anomaly Identification Performance

The anomaly identification module has a precision of 0.87, recall of 0.81, and F1-score of 0.84, when evaluated against a manually curated set of 142 anomalous AQI episodes, for the test period, which mostly overlaps with documented traffic-incident reports and recorded temperature-inversion events. The observed versus predicted AQI for a representative ten-day test window is shown in Figure 3, along with the absolute residual and dynamic anomaly threshold with the residual-flagged anomalies indicated. There are four anomalous episodes identified by the module, three of which are concurrent with recorded periods of increased traffic congestion and the fourth is concurrent with a recorded low wind speed stagnation event. False negatives are mostly found in the case of a gradual onset pollution episode, where the pollution is steadily increasing but not reaching the dynamic threshold sharply in one time step, indicating that it may be beneficial to consider implementing a cumulative-sum (CUSUM)-style detection criterion for future research.

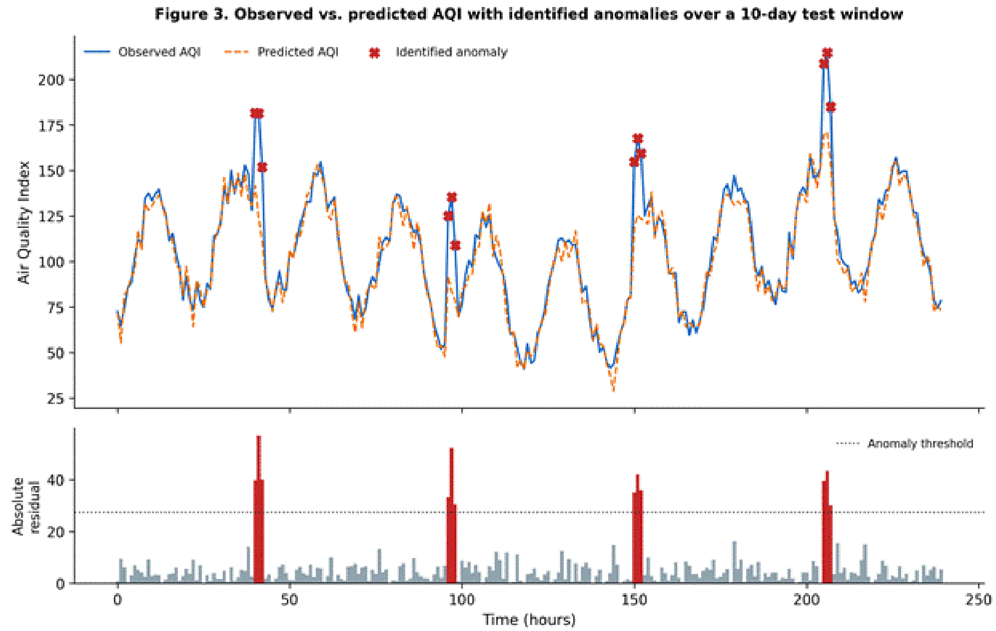

Figure 3. Observed vs. predicted AQI with identified anomalies over a 10-day test window

4.4 Discussion

Results indicate that the use of traffic-flow data with meteorological and pollutant history data can significantly enhance the accuracy of short-term AQI forecasting, which is consistent with the notion that traffic emissions are a localised source that is highly spatiotemporal. Importantly, the explainability layer validates the physically plausible rankings of the different features (importance) learned by the model: the top features are particulate-matter concentration and traffic volume, which agrees with the well-known results of atmospheric-chemistry research on the persistence and dispersal of pollutants (Kow et al., 2020; Zheng et al., 2013). This correlation between data-based feature attribution and domain knowledge is key to establishing trust among environmental regulators, who can then audit individual predictions instead of taking forecasts as a black box. When combined with explainability, the framework's usefulness goes beyond prediction to diagnosis; it is able to point to the likely causes of unusual pollution events.

A few shortcomings should be recognized. The data are from one metropolitan monitoring network only and the 18-month observation period covers seasonal changes; the generalizability of the model to cities that have very different traffic mixes, industrial mixes or topographical dispersion characteristics needs to be verified. The SHAP approximation used here is fast but requires a background reference distribution which can affect the magnitude of the attributed values, and the values should thus be interpreted as relative, not absolute, contributions. Moreover, the anomaly identification module uses a rolling statistical threshold that is adaptive, but may fail to capture anomalies that have a gradual onset, as seen with the false-negative pattern above. While they do not detract from the main conclusions, they do outline the clear directions for the extensions proposed in the concluding section.

5. Explainability Analysis

Among the various factors, the most important one that influences the PM2.5 concentration is lagged PM2.5 concentration (mean |SHAP value| = 0.184), which is a feature of PM2.5 data with high autocorrelation and slow decay, as shown in Figure 2 (Section 4.1). The second most important variable is traffic volume (0.151), which is empirical evidence that vehicle emissions are a major, time dependent source of the short-term AQI fluctuation, exceeding a cluster of meteorological variables. The next are lagged PM10 concentration (0.132) and relative humidity

(0.098), which is consistent with its known effects on secondary aerosol formation and hygroscopic particle growth. The moderate and comparative contribution from wind speed (0.086) and NO2 concentration (0.079) are followed by temporal features hour-of-day (0.054) and season (0.021), respectively, when the primary meteorological and traffic signals are removed.

The instance-level SHAP explanations of the anomalous episodes identified in Section 4.3 show that the traffic-driven anomalies involve a disproportionately high positive SHAP contribution from traffic-volume and congestion-index features, while the meteorologically-driven stagnation event has the largest positive SHAP contribution combined from low wind speed and high relative humidity. This ability to break down individual anomalies into the drivers of the anomaly itself is a significant improvement over existing anomaly-detection pipelines, which generally mark an anomaly but aren't able to determine the likely physical cause. These attention-weight visualizations support the SHAP results, as the model's attention weight is the highest on the 6 to 9 hours leading up to a forecast during anomalous times, while the attention weight is more evenly split across the 6 to 9 hours window during stable, low-variability times, suggesting that during times where the recent trend of a pollutant becomes more volatile, the model implicitly learns to increase its effective temporal receptive field.

6. CONCLUSION AND FUTURE WORK

The paper proposed the explainable deep learning framework with a hybrid model (CNN-BiLSTM-Attention) for AQI prediction, which combines meteorological, traffic-flow, and historical pollutant data, along with SHAP- and Integrated-Gradients-based explainability and a residual-driven anomaly identification module. Experiment results show that the proposed model has significant advantages over the conventional LSTMs, GRUs and CNN-LSTMs baselines in terms of forecasting accuracy and that the model offers physically meaningful and transparent feature attribution consistent with the known atmospheric-chemistry physics. The anomaly identification module can further enhance the practicality of the framework by identifying the anomalies in the pollution episodes, and assigning them to their traffic or meteorological driver, providing actionable diagnostic information for environmental agencies and traffic-management authorities.

Future work will focus on three aspects of extending the framework: first, to include spatial graph-based representations of the pollutant transport between several monitoring stations in an urban network; second, to test the framework's transferability to other cities with different traffic and climatic characteristics; and third, to integrate the anomaly identification module into real-time traffic-control systems, allowing for closed-loop adaptive interventions like dynamic congestion pricing during anticipated high-pollution events. The framework's applicability to high-stakes public-health decision making would be further enhanced by incorporating the quantification of uncertainty, such as by using Bayesian deep learning or conformal prediction.

References

1. Chen, S., Wang, Y., & Li, H. (2021). A hybrid CNN-LSTM model for urban air quality index forecasting using multi-source environmental data. Environmental Modelling & Software, 143, 105123. https://doi.org/10.1016/j.envsoft.2021.105123
2. Du, S., Li, T., Yang, Y., & Horng, S. J. (2020). Deep air quality forecasting using hybrid deep learning framework. IEEE Transactions on Knowledge and Data Engineering, 33(6), 2412–2424. https://doi.org/10.1109/TKDE.2019.2954510
3. Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. Proceedings of the 2018 IEEE 5th International Conference on Data Science and Advanced Analytics, 80–89. https://doi.org/10.1109/DSAA.2018.00018
4. Kow, P. Y., Wang, Y. S., Zhou, Y., Kao, I. F., Issermann, M., Chang, L. C., & Chang, F. J. (2020). Seamless integration of convolutional and back-propagation neural networks for regional multi-step-ahead PM2.5 forecasting. Journal of Cleaner Production, 261, 121285. https://doi.org/10.1016/j.jclepro.2020.121285
5. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.
6. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778
7. Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. Proceedings of the 34th International Conference on Machine Learning, 70, 3319–3328.
8. Wang, J., Li, J., Wang, X., Wang, J., & Huang, M. (2021). Air quality prediction using CT-LSTM. Neural Computing and Applications, 33(10), 4779–4792. https://doi.org/10.1007/s00521-020-05535-w
9. World Health Organization. (2022). Ambient (outdoor) air pollution fact sheet. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health
10. Yan, R., Liao, J., Yang, J., Sun, W., Nong, M., & Li, F. (2021). Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Systems with Applications, 169, 114513. https://doi.org/10.1016/j.eswa.2020.114513

11. Zhang, B., Zhang, H., Zhao, G., & Lian, J. (2020). Constructing a PM2.5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks. Environmental Modelling & Software, 124, 104600. https://doi.org/10.1016/j.envsoft.2019.104600
12. Zheng, Y., Liu, F., & Hsieh, H. P. (2013). U-Air: When urban air quality inference meets big data. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1436–1444. https://doi.org/10.1145/2487575.2488188