Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Baseline-relative predictability of hourly PM2.5 and O3 in a low-emission Sou th Chinese city: an attention-LSTM with feature attribution

Chu Wu

wuchu182@126.com

Heyuan Eco-environmental Monitoring Station of Guangdong Province

Jiajun Ou

Guangdong University of Technology

Jianwei Zeng

Key Laboratory of Regional Air Quality Monitoring, Ministry of Ecology and Environment, Guangdong Environmental Protection Key Laboratory of Secondary Air Pollution Research

Xixi Cai

Guangdong Academy of Agricultural Sciences

Luan Yuan

Key Laboratory of Regional Air Quality Monitoring, Ministry of Ecology and Environment, Guangdong Environmental Protection Key Laboratory of Secondary Air Pollution Research

Research Article

Keywords: Air quality prediction, Attention-LSTM, SHAP, Temporal evaluation, PM2.5, O3

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11616075/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Baseline-relative predictability of hourly PM2.5 and O3 in a low-emission South Chinese city: an attention-LSTM with feature attribution

Chu Wu[1,2*], Jiajun Ou[3], Jianwei Zeng[2], Xixi Cai[4], Luan Yuan[2]

1. Heyuan Eco-environmental Monitoring Station of Guangdong Province, Heyuan 517000, China
2. Key Laboratory of Regional Air Quality Monitoring, Ministry of Ecology and Environment, Guangdong Environmental Protection Key Laboratory of Secondary Air Pollution Research, Guangdong Ecological Environmental Monitoring Center, Guangzhou 510308, China
3. School of Automation, Guangdong University of Technology, Guangzhou 510006, China
4. Guangdong Key Laboratory of Ornamental Plant Germplasm Innovation and Utilization, Environmental Horticulture Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

*Corresponding author:

Dr. Chu Wu

E-mail: wuchu182@126.com; Tel.: +86-762-3883573; Fax: +86-762-3882367

Abstract: Recent observations can support short-term air quality prediction, but temporal persistence complicates the interpretation of high predictive agreement. We applied an attention-based long short-term memory (attention-LSTM) model to hourly observations from two urban monitoring stations in Heyuan during 2019–2023. SHapley Additive exPlanations (SHAP) were used to examine model-specific feature contributions. Reported test-set R^2 values ranged from 0.9796 to 0.9877 for PM2.5 and from 0.9921 to 0.9975 for O3. Corresponding RMSE values were 1.28–2.02 μg/m³ and 1.77–2.83 μg/m³. Predictions followed the displayed temporal patterns, while errors varied across concentration intervals. Recent pollutant observations received substantial SHAP importance, with humidity and temperature also contributing to O3 predictions. The analysis describes local prediction behavior and statistical feature associations. These associations do not establish pollutant formation mechanisms.

Keywords: Air quality prediction; Attention-LSTM; SHAP; Temporal evaluation; PM2.5; O3

1 Introduction

Fine particulate matter with an aerodynamic diameter of 2.5 μm or less (PM2.5) and ozone (O3) remain important air quality concerns (Mabahwi et al., 2014; Wang et al., 2017). Short-term concentration predictions can describe expected local conditions, but their usefulness depends on the forecast horizon and available observations. This study considers monitoring records from Heyuan, Guangdong, where relatively low pollutant concentrations have been reported (Wu et al., 2022). Here, low concentration describes the study setting rather than a distinct atmospheric chemical regime.

Three-dimensional air quality models represent atmospheric transport, emissions and chemical processes. Their predictions depend on emission inventories, meteorological inputs and process parameterizations (Gao & Zhou, 2024; Liu et al., 2023). Data-driven models provide a complementary approach to predicting concentrations from monitoring records.

Previous research has examined hybrid approaches combining numerical air quality modeling and machine learning (Li et al., 2023; Liao et al., 2023). Emission inventory development and ground-based inversion provide further avenues for improvement (Zheng et al., 2019; Kim et al., 2023; Zhu et al., 2024). These approaches emphasize the i

mportance of input information and motivate evaluation beyond model architecture al one.

Long short-term memory (LSTM) networks have been applied to air quality time series, including encoder–decoder and hybrid architectures (Bui et al., 2018; Huang & Kuo, 2018; Qin et al., 2019). Attention-based extensions weight historical representations during prediction (Li et al., 2019; Wen & Li, 2023). Their incremental contribution requires comparison with an otherwise matched model without attention.

Recent pollutant observations often contain predictive information about subsequent concentrations. Using historical concentrations is legitimate when they are available at the forecast origin. Temporal persistence can explain high prediction agreement without constituting data leakage. Leakage occurs when later observations or evaluation data influence inputs, fitted preprocessing or model selection (Hewamalage et al., 2023). Chronological splitting alone does not exclude these routes.

Feature attribution provides a complementary description of how a fitted model uses its inputs. SHAP summarizes estimated contributions relative to a reference distribution (Ponce-Bobadilla et al., 2024). Correlated pollutants, lagged values and rolling summaries can share information, which complicates interpretation. Temporal attribution methods also require attention to sequence structure (Bento et al., 2021). Feature importance therefore should not be equated with chemical causation or emission-control effectiveness.

This study examines short-term PM2.5 and O3 prediction at two neighboring urban monitoring stations in Heyuan. Its objectives are to describe local prediction agreement, examine concentration-dependent errors and summarize model-specific feature contributions. The application uses an established attention-LSTM architecture. The geographic scope is limited to the two evaluated stations.

2 Data and methods

2.1 Monitoring sites and observations

The DP and LC monitoring stations are located in urban Heyuan, northeastern Guangdong, approximately 4 km apart. They provide a local comparison between nearby monitoring sites (Wu et al., 2022). Their proximity limits geographic generalization.

Observation coverage

The study covers hourly observations from 1 January 2019 to 31 December 2023. The listed measurements include pressure, temperature, humidity, wind direction, wind speed, SO2, NO, NO2, NO_x, O3, CO, PM10 and PM2.5.

The reported preprocessing procedure comprised screening with the 3σ criterion and interquartile ranges, linear interpolation for gaps shorter than three hours, and seasonal or time-matched imputation for longer gaps (Rangel-Heras et al., 2023). Potential future-information use in interpolation and preprocessing is considered among the evaluation limitations.

The reported input construction included calendar indicators, lagged pollutant and meteorological observations, rolling summaries and empirical interaction terms (Dai et al., 2021; Naz et al., 2024; Tran et al., 2023). The interaction terms serve as empirical predictors and are not direct measurements of atmospheric stability or transport.

2.2 Model architecture and training

Figure 1 depicts an encoder–decoder LSTM with attention. The encoder contains layers with 64 and 32 hidden units and dropout of 0.2. Softmax-normalized scores weight encoder hidden states to form a context representation. The decoder and output layers map this representation to a concentration prediction. This schematic uses an established modeling approach (Li et al., 2019; Wen & Li, 2023).

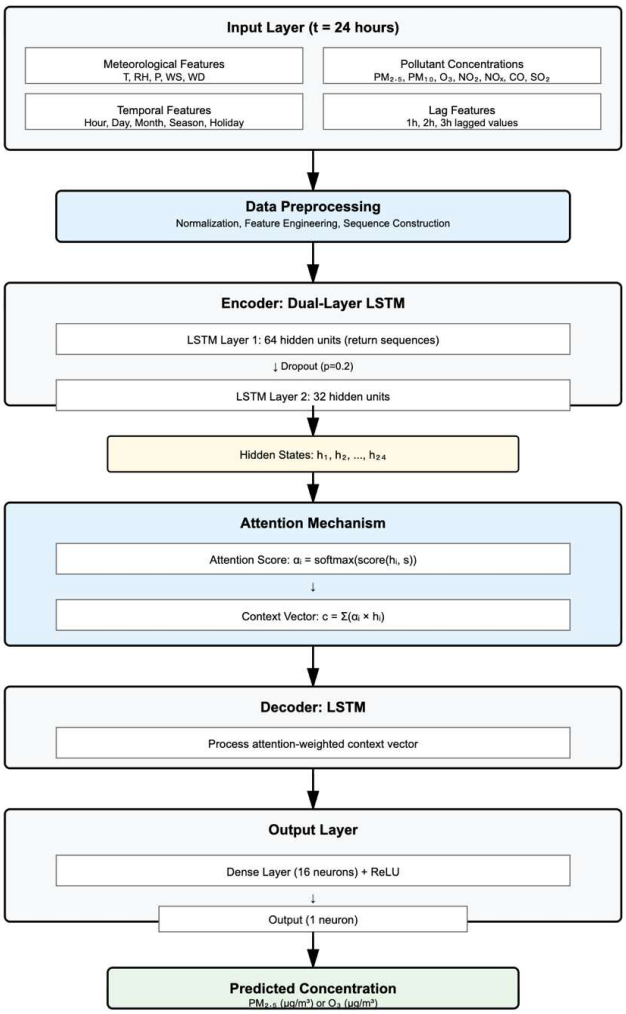


Fig. 1 Architecture of the attention-LSTM concentration prediction model

The reported training configuration specified a 16-neuron fully connected layer, a single-neuron output, the Adam optimizer with an initial learning rate of 0.001, mean squared error loss, learning-rate decay and early stopping after 15 epochs without validation improvement.

The reported configuration used a 24-hour input history and a batch size of 32. A diurnal cycle motivates consideration of a 24-hour history, but this choice does not establish that the history length is optimal.

2.3 Evaluation metrics

Prediction errors are described using mean absolute error (MAE), root mean square error (RMSE) and the coefficient of determination (R^2). MAE and RMSE are reported in $\mu g/m^3$, while R^2 is dimensionless and can be negative. Performance is reported separ

ately for each station and pollutant. Near-zero concentrations can make percentage errors unstable, so MAPE is not used as the principal measure.

For n observed targets, residuals are defined as $e_i = \hat{y}_i - y_i$. The corresponding metrics are $MAE = \Sigma|e_i|/n$, $RMSE = \sqrt{(\Sigma e_i^2/n)}$, and $R^2 = 1 - \Sigma e_i^2/\Sigma(y_i - \bar{y})^2$, where $\hat{y}_i$ and y_i denote predicted and observed concentrations and $\bar{y}$ is the mean observed concentration. R^2 is undefined when the observed target variance is zero.

Table 1 summarizes the reported training, validation and test results. These metrics are interpreted descriptively; incremental skill over persistence and the contribution of attention have not been established by matched comparisons.

2.4 SHAP feature attribution

SHAP was used to summarize contributions of pollutant and meteorological inputs to fitted model predictions. Global importance was represented by mean absolute SHAP values, alongside signed summary and individual prediction plots. These summaries describe model behavior relative to a background distribution.

Correlated inputs can produce non-unique attribution patterns. Perturbing lagged and rolling features independently may also produce combinations unlike observed sequences. Rankings are therefore interpreted as exploratory descriptions of fitted models. They do not identify atmospheric formation pathways, precursor limitation or causal effects.

Figures 14 and 15 summarize global and individual feature contributions. SHAP values represent concentration differences in μg/m³ only when the explained output is expressed on that scale.

3 Results

3.1 PM2.5 prediction agreement

At DP, the reported test-set R^2 for PM2.5 was 0.9877, with RMSE of 1.28 μg/m³ and MAE of 0.74 μg/m³ (Table 1). At LC, R^2 was 0.9796, RMSE was 2.02 μg/m³ and MAE was 0.89 μg/m³. Training-set R^2 values were 0.9629 and 0.9584, respectively. Validation-set values were 0.9880 and 0.9885. These differences should be interpreted alongside concentration distributions and the temporal composition of each subset.

Table 1 Reported model performance at the DP and LC monitoring stations

Station	Pollutant	Training R^2	Validation R^2	Test R^2	Test RMSE (μg/m³)	Test MAE (μg/m³)
DP	$PM_{2.5}$	0.9629	0.9880	0.9877	1.28	0.74
DP	O_3	0.9636	0.9942	0.9921	2.83	1.98
LC	$PM_{2.5}$	0.9584	0.9885	0.9796	2.02	0.89
LC	O_3	0.9658	0.9913	0.9975	1.77	1.28

Predictions and observations were closely aligned in the displayed PM2.5 scatterplots (Fig. 2). The predicted time series followed the observed variation in Fig. 3. These plots summarize the displayed samples without separately quantifying seasonal or episode-specific performance.

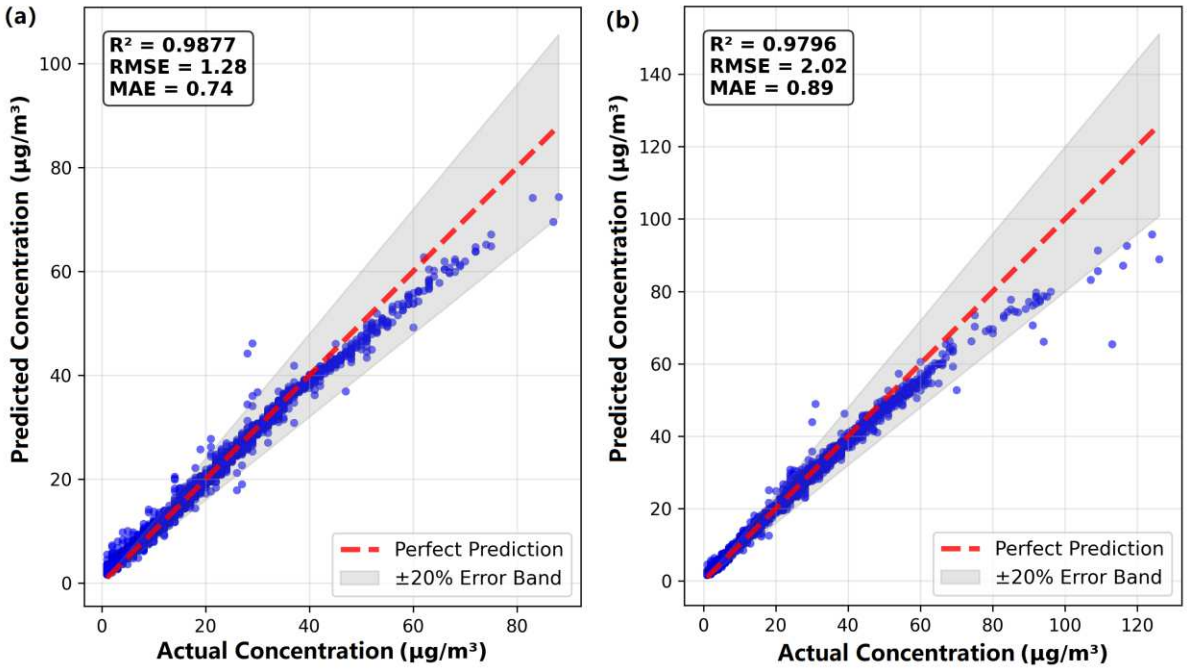
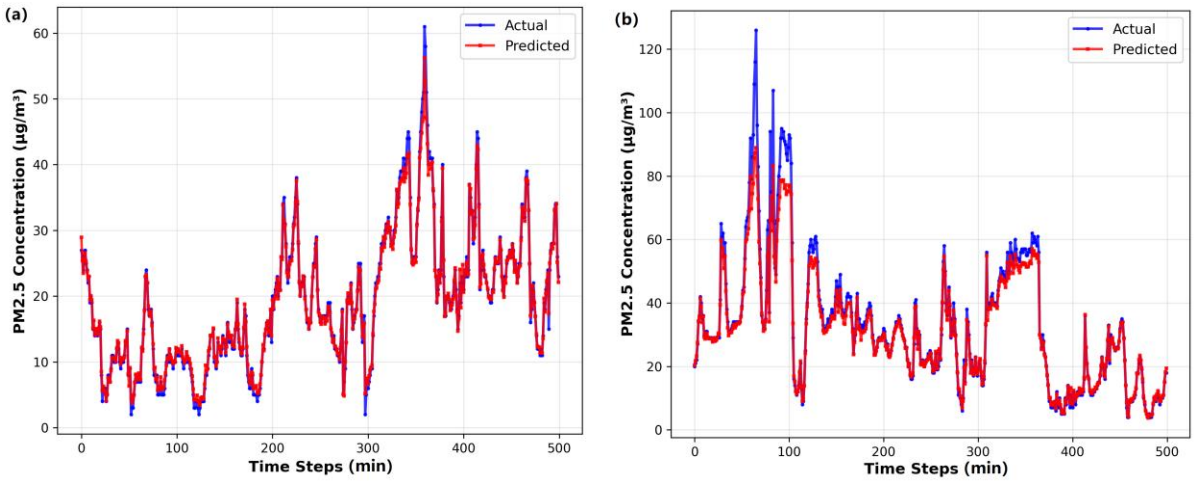


Fig. 2 *Predicted versus observed $PM_{2.5}$ concentrations (a, DP; b, LC)*

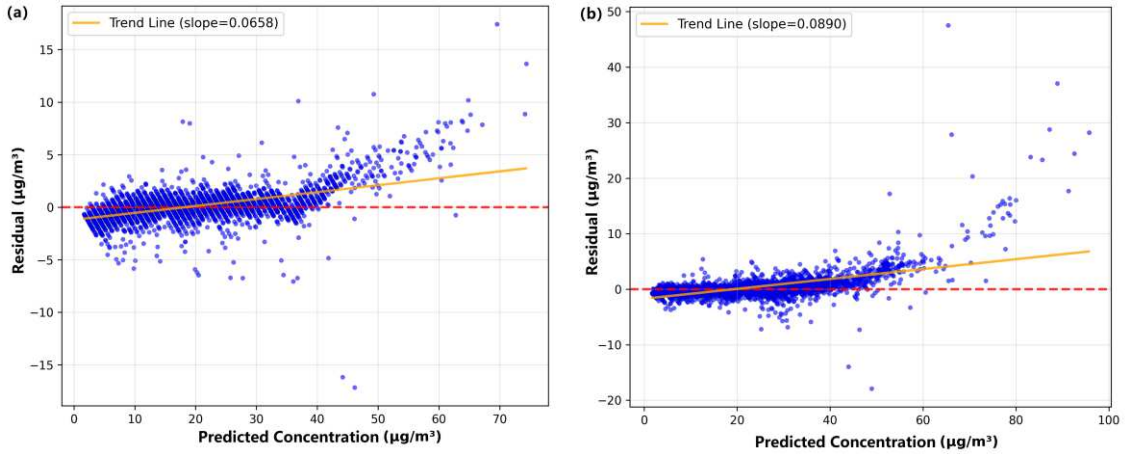


170

Fig. 3 *Predicted and observed temporal changes in $PM_{2.5}$ concentration (a, DP; b, LC)*

171

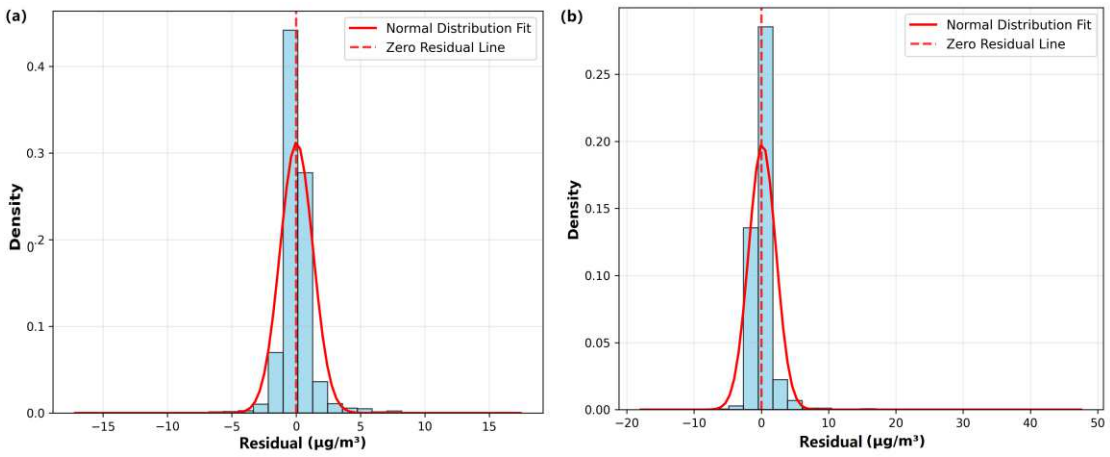
172 The displayed PM2.5 residuals were concentrated near zero (Figs. 4 and 5). Figure 5 i
173 ncludes fitted normal curves, which do not establish residual normality. Mean absolut
174 e errors increased in the highest displayed concentration interval, Q5 (Fig. 6). These d
175 istributions do not establish residual independence or negligible bias.



176

Fig. 4 *Residuals of $PM_{2.5}$ concentration predictions (a, DP; b, LC)*

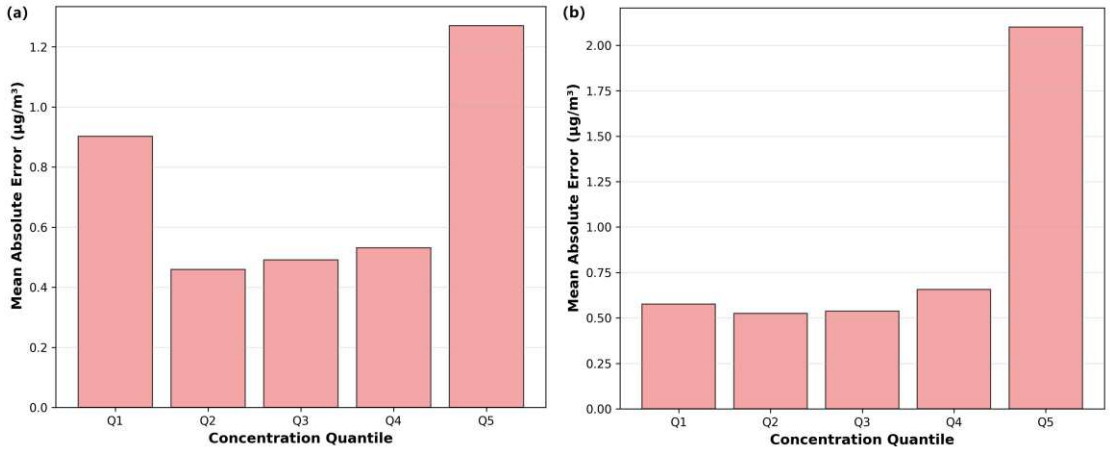
177



178

Fig. 5 PM2.5 residual histograms with fitted normal curves (a, DP; b, LC)

179

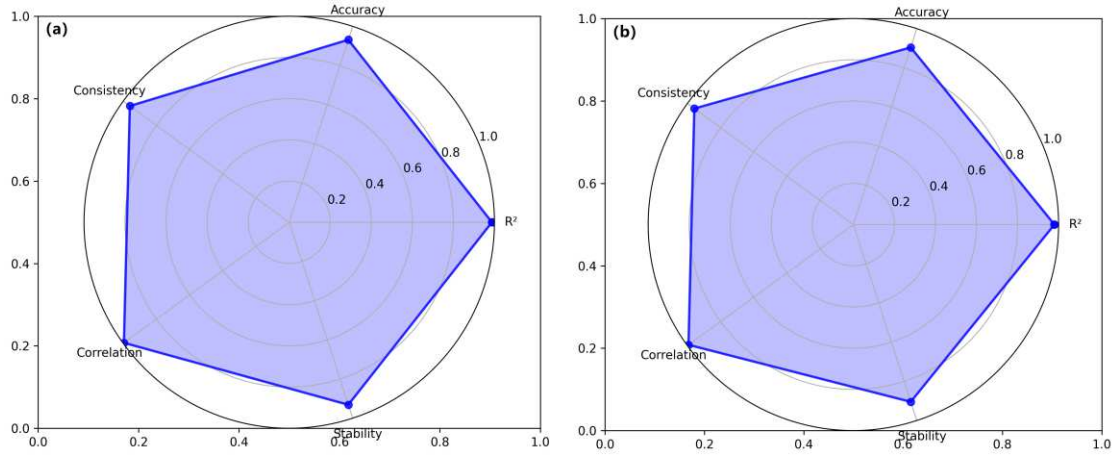


180

Fig. 6 Mean absolute errors of PM2.5 predictions across concentration groups (a, DP; b, LC)

181

182 Figure 7 summarizes the reported PM2.5 metrics. DP had lower test RMSE and MAE
183 than LC.



184

Fig. 7 *$PM_{2.5}$ concentration prediction performance of the model (a, DP; b, LC)*

3.2 O3 prediction agreement

At DP, the reported O3 test-set R^2 was 0.9921, RMSE was 2.83 μg/m³ and MAE was 1.98 μg/m³ (Table 1). At LC, the corresponding values were 0.9975, 1.77 μg/m³ and 1.28 μg/m³. The reported O3 R^2 values exceeded the PM2.5 values, although R^2 depends on target variability.

The O3 scatterplots showed close prediction–observation agreement (Fig. 8). Predicted concentrations followed the displayed temporal variation in Fig. 9.

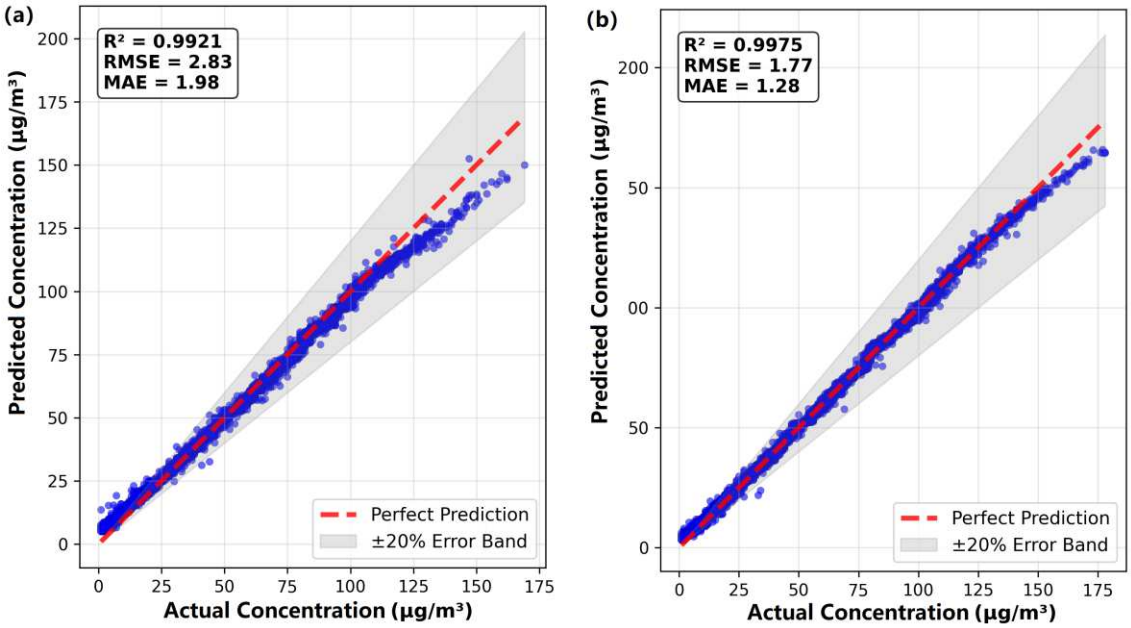


Fig. 8 *Predicted versus observed O_3 concentrations (a, DP; b, LC)*

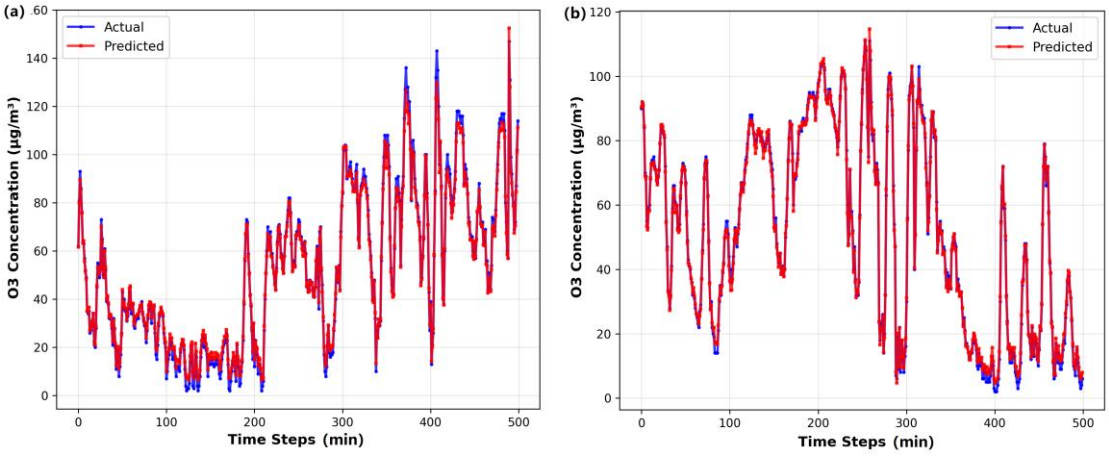


Fig. 9 Predicted and observed O3 time series (a, DP; b, LC)

The O3 residual distributions were centered near zero in the displayed plots (Figs. 10 and 11). MAE was lower in middle concentration intervals and higher at the lowest and highest intervals (Fig. 12). These patterns describe error variation, rather than the statistical significance of differences between intervals.

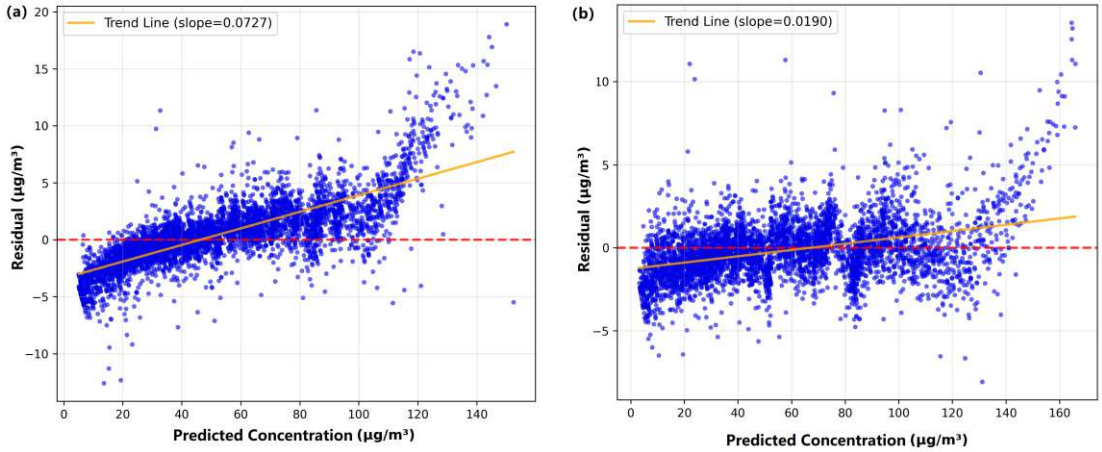


Fig. 10 Residuals of O3 predictions (a, DP; b, LC)

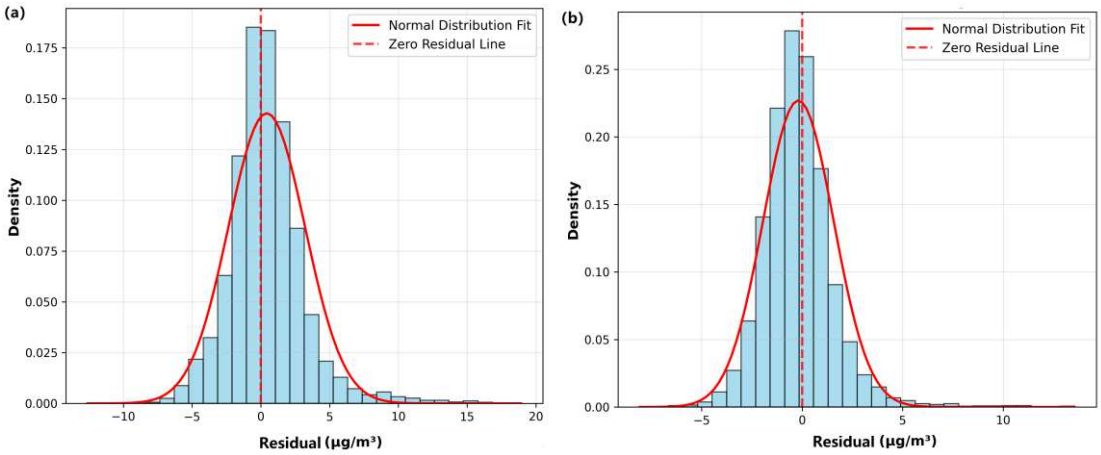


Fig. 11 O3 residual histograms with fitted normal curves (a, DP; b, LC)

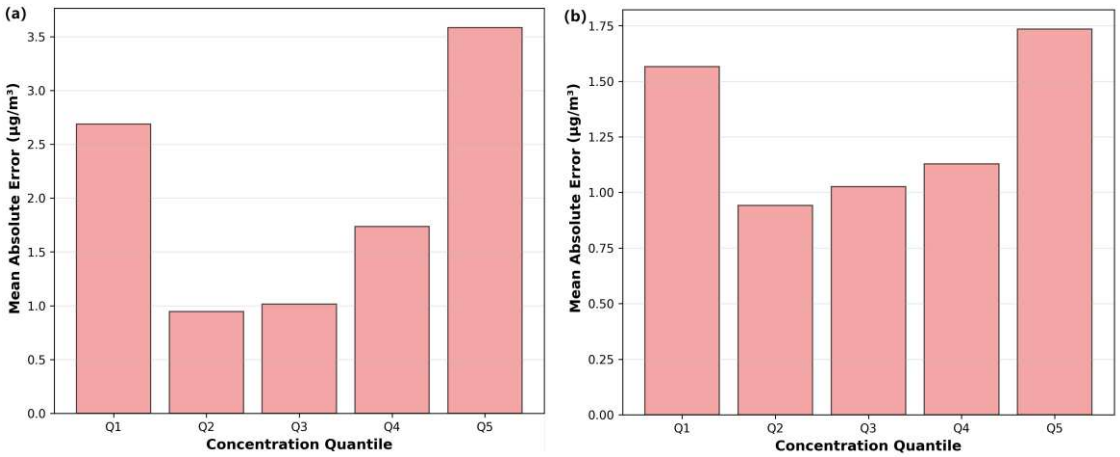


Fig. 12 Mean absolute errors of O3 predictions across concentration groups (a, DP; b, LC)

The plotted O3 predictions followed peaks and troughs in the displayed series. This agreement describes concentration patterns rather than photochemical reaction mechanisms.

Figure 13 summarizes the reported O3 metrics. Test RMSE and MAE were lower at LC than at DP.

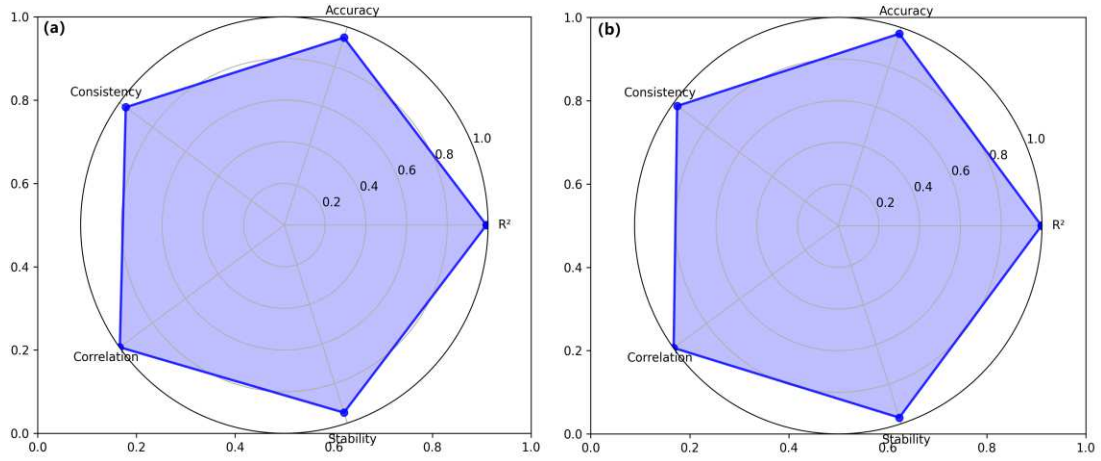


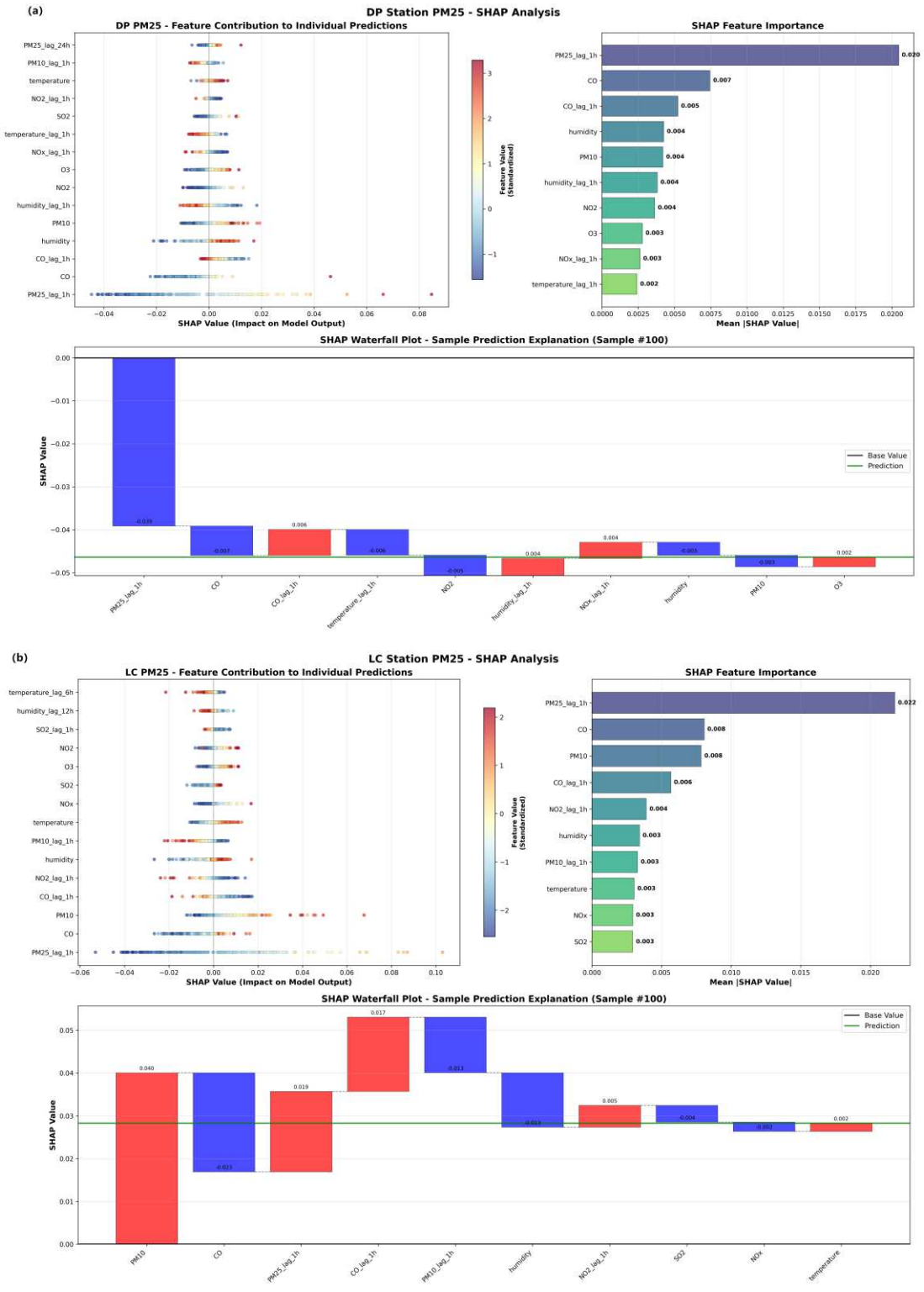
Fig. 13 *O_3 concentration prediction performance of the model (a, DP; b, LC)*

Differences in concentration distributions, monitoring conditions and temporal composition can affect comparisons between stations. The present analysis does not isolate the contributions of local emissions or microclimate.

3.3 Model-specific feature attribution

The SHAP plots summarize feature contributions within the fitted models. Lagged concentrations received substantial importance for both pollutants. Rankings are interpreted as model-specific descriptions because the inputs share correlated information.

Recent PM2.5 observations were important at both stations (Fig. 14). CO, PM10 and nitrogen-related features also contributed to predictions. These contributions describe shared predictive information among pollutants, without distinguishing accumulation, transport and chemical production.



225

226

Fig. 14 SHAP feature attribution for PM2.5 predictions (a, DP; b, LC)

227

228 Feature rankings differed between DP and LC. Attribution magnitudes can depend on
229 feature distributions, correlations and background selection. These rankings do not ide
230 ntify station-specific emission sources or secondary aerosol pathways.

For O3, lagged O3, humidity and temperature contributed to predictions at both stations (Fig. 15). Higher humidity values generally had negative contributions in the displayed summaries. The sign describes an association within the fitted model, rather than a direct humidity effect on ambient O3.

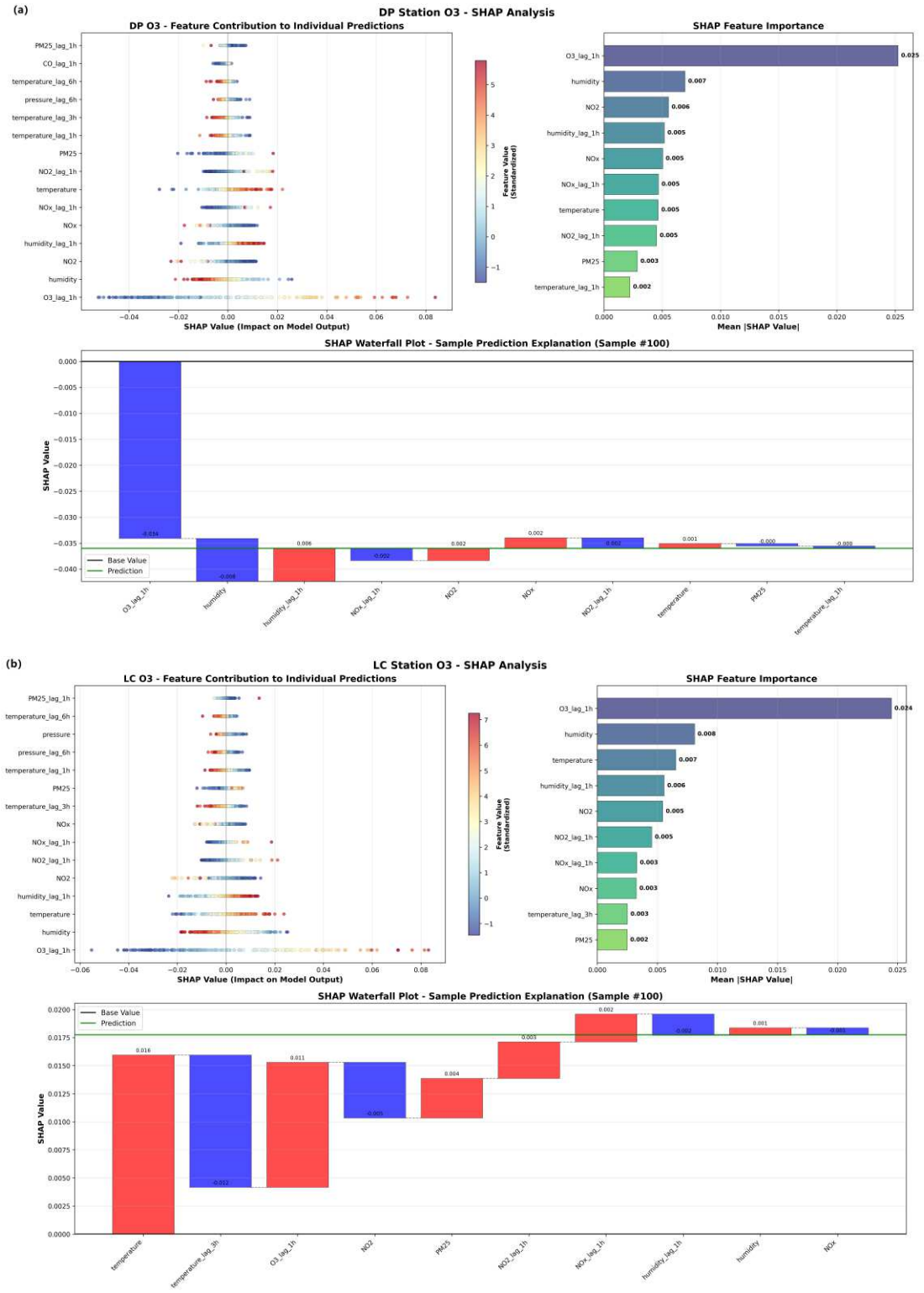


Fig. 15 SHAP feature attribution for O3 predictions (a, DP; b, LC)

Some nitrogen-related features and PM2.5 had smaller contributions, while pressure-related contributions were near zero in the displayed summaries. Small attribution values do not establish small physical or chemical effects. Correlated features can share information or substitute for one another within a fitted model.

The available predictor information does not establish VOC reactivity, solar radiation or photolysis rates. The attribution analysis therefore cannot identify a VOC-limited or NO_x-limited chemical regime.

Overall, the available attribution plots indicate model reliance on recent pollutant observations and selected meteorological inputs. This is an exploratory description of prediction behavior. Chemical mechanisms would require independent observations or process-based analyses.

4 Discussion

4.1 Interpretation of prediction agreement

The reported metrics describe close prediction–observation agreement in the evaluated samples. Substantial attribution to recent concentrations suggests that temporal persistence may explain part of this agreement. Differences between training and validation R^2 values alone establish neither leakage nor its absence.

4.2 Scope and limitations

The two nearby stations provide a local application of an established architecture. Their proximity limits geographic conclusions, and the results do not demonstrate transferability to other cities or chemical environments. Correlated inputs also constrain the interpretation of SHAP rankings. Temporal evaluation and reproducible input construction are necessary for distinguishing model-specific forecasting skill from persistence.

The present evaluation does not establish that future-information use was excluded from preprocessing or input construction. The exact forecast horizon, subset boundaries and preprocessing endpoints are not sufficiently documented to reconstruct the temporal evaluation. The reported metrics therefore should not be interpreted as independently validated operational forecasting skill. Matched persistence baselines, an attention-free LSTM comparison and uncertainty estimates are needed to quantify any increment

al benefit. The SHAP input mapping, output scale and waterfall additivity also require reproducibility checks before quantitative attribution magnitudes can be interpreted. These limitations restrict the conclusions to a descriptive analysis of the reported predictions and feature-attribution plots.

5 Conclusions

This study examined short-term PM2.5 and O3 prediction and model-specific feature attribution at two urban stations in Heyuan. Reported test R^2 values were 0.9796–0.9877 for PM2.5 and 0.9921–0.9975 for O3. Prediction errors varied with pollutant, station and concentration interval.

SHAP highlighted recent pollutant concentrations and selected meteorological inputs as contributors to the fitted predictions. These attributions describe statistical model behavior and do not establish pollutant formation mechanisms or emission-control responses.

The analysis characterizes local prediction behavior and highlights the contribution of recent pollutant observations. Its scope is limited to the evaluated stations and concentration targets. The findings do not establish multi-day forecasting or regional operational performance.

References

Bento, J., Saleiro, P., Cruz, A. F., Figueiredo, M. A. T., & Bizarro, P. (2021). TimeSHAP: Explaining recurrent models through sequence perturbations. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2565–2573. https://doi.org/10.1145/3447548.3467166

Bui, T.-C., Le, V.-D., & Cha, S.-K. (2018). A deep learning approach for forecasting air pollution in South Korea using LSTM. arXiv preprint, arXiv:1804.07891. https://doi.org/10.48550/arXiv.1804.07891

Dai, H., Huang, G., Wang, J., Zeng, H., & Zhou, F. (2021). Prediction of Air Pollutant Concentration Based on One-Dimensional Multi-Scale CNN-LSTM Considering Spatial-Temporal Characteristics: A Case Study of Xi'an, China. *Atmosphere, 12*(12), 1626.

Gao, Z., & Zhou, X. (2024). A review of the CAMx, CMAQ, WRF-Chem and NAQPMS models: Application, evaluation and uncertainty factors. *Environmental Pollution, 343*, 123183.

Hewamalage, H., Ackermann, K., & Bergmeir, C. (2023). Forecast evaluation for data scientists: Common pitfalls and best practices. Data Mining and Knowledge Discovery, 37(2), 788–832. https://doi.org/10.1007/s10618-022-00894-5

Huang, C. J., & Kuo, P. H. (2018). A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities. *Sensors, 18*(7), 2220.

Kim, S., Kim, J., Hu, H., Jang, M., Lee, J.-B., Hong, S. C., Kim, O., & Woo, J.-H. (2023). Update of the year 2019 modeling emission inventory in China. Asian Journal of Atmospheric Environment, 17, 20. https://doi.org/10.1007/s44273-023-00012-x

Li, B., Hu, Q., Gao, M., Liu, T., Zhang, C., & Liu, C. (2023). Physical informed neural network improving the WRF-CHEM results of air pollution using satellite-based remote sensing data. *Atmospheric Environment, 311*, 120031.

Li, Y., Zhu, Z., Kong, D., Han, H., & Zhao, Y. (2019). EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowledge-Based Systems, 181*, 104785.

Liao, H., Yuan, L., Wu, M., & Chen, H. (2023). Air quality prediction by integrating mechanism model and machine learning model. *Science of The Total Environment, 899*, 165646.

Liu, C., Wu, C., Kang, X., Zhang, H., Fang, Q., Su, Y., Li, Z., Ye, Y., Chang, M., & Guo, J. (2023). Evaluation of the prediction performance of air quality numerical forecast models in Shenzhen. *Atmospheric Environment, 314*, 120058.

Mabahwi, N. A. B., Leh, O. L. H., & Omar, D. (2014). Human health and wellbeing: Human health effect of air pollution. Procedia - Social and Behavioral Sciences, 153, 221–229. https://doi.org/10.1016/j.sbspro.2014.10.056

Naz, F., Fahim, M., Cheema, A. A., Viet, N. T., Cao, T.-V., Hunter, R., & Duong, T. Q. (2024). Two-Stage Feature Engineering to Predict Air Pollutants in Urban Areas. *IEEE Access, 12*, 114073–114085.

Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. Clinical and Translational Science, 17(11), e70056. https://doi.org/10.1111/cts.70056

Qin, D., Yu, J., Zou, G., Yong, R., Zhao, Q., & Zhang, B. (2019). A novel combined prediction scheme based on CNN and LSTM for urban PM2.5 concentration. IEEE Access, 7, 20050–20059. https://doi.org/10.1109/ACCESS.2019.2897028

Rangel-Heras, E., Zuniga, P., Alanis, A. Y., Hernandez-Vargas, E. A., & Sanchez, O. D. (2023). Seasonal time-series imputation of gap missing algorithm (STIGMA). Kybernetika, 59(6), 861–879. https://doi.org/10.14736/kyb-2023-6-0861

Tran, H. D., Huang, H.-Y., Yu, J.-Y., & Wang, S.-H. (2023). Forecasting hourly $PM_{2.5}$ concentration with an optimized LSTM model. *Atmospheric Environment, 315*, 120161.

Wang, T., Xue, L., Brimblecombe, P., Lam, Y. F., Li, L., & Zhang, L. (2017). Ozone pollution in China: A review of concentrations, meteorological influences, chemical precursors, and effects. Science of The Total Environment, 575, 1582–1596. https://doi.org/10.1016/j.scitotenv.2016.10.081

Wen, X., & Li, W. (2023). Time Series Prediction Based on LSTM-Attention-LSTM Model. *IEEE Access, 11*, 48322–48331.

Wu, C., Li, J., Lin, X., Mo, Y., Zhang, H., Liao, Z., Wang, Y., & Chen, D. (2022). Characteristics of typical air pollutants in a valley-basin city in South China. Meteorology and Atmospheric Physics, 134, 14. https://doi.org/10.1007/s00703-021-00850-w

Zheng, H., Cai, S., Wang, S., Zhao, B., Chang, X., & Hao, J. (2019). Development of a unit-based industrial emission inventory in the Beijing–Tianjin–Hebei region and resulting improvement in air quality modeling. A*t*mosphe*ri*c *C*hemistry *an*d Physi*cs, 19*(6), 3447–3462. https://doi.org/10.5194/acp-19-3447-2019

Zhu, L., Tang, X., Yang, W., Zhao, Y., Kong, L., Wu, H., Fan, M., Yu, C., & Chen, L. (2024). Improvement of PM2.5 Forecast in China by Ground-Based Multi-Pollutant Emission Source Inversion in 2022. *Atmosphere, 15*(2), 181. https://doi.org/10.3390/atmos15020181

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Data will be made available on request.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was financially supported by the Social Development S&T Planning Fund of Heyuan City, China (No. 241010101471315).

Author contributions

Chu Wu: Conceived the main idea and wrote the manuscript. **Jiajun Ou:** Performed the experiments and collected the data. **Jianwei Zeng:** Performed the experiments and collected the data. **Xixi Cai:** Contributed to the data analysis and prepared Figures 1–3.