https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

BiXiao: An AI-dirven Atmospheric Environmental Forecasting Model with Non-continuous Grids

Shengxuan Ji[1,5], Yawei Qu[2], Cheng Yuan[3], Tijian Wang[1], Bing Liu[4], Lili Zhu[4], Huihui Zheng[4], Zhenfeng Qiu[5], and Pulong Chen[6]

[1]School of Atmospheric Science, Nanjing University, Nanjing 210023, China
[2]College of Intelligent Science and Control Engineering, Jiling Institute of Technology, Nanjing 211169, China
[3]School of Emergency Management, Nanjing University of Information Science & Technology, Nanjing 210044, China
[4]China National Environmental Monitoring Center, Beijing 100012, China
[5]FuYao Intelligence (Beijing) Technology Co., Ltd., Beijing 101108, China
[6]Net Zero Era (Jiangsu) Environmental Technology Co., Ltd., Nanjing 210023, China

Correspondence: Yawei Qu (yawei.qu@jit.edu.cn), Cheng Yuan (cyuan@nuist.edu.cn), and Tijian Wang (tjwang@nju.edu.cn)

Abstract. High-precision and efficient atmospheric environmental forecasting is essential for protecting public health and supporting environmental management. However, traditional physics-based numerical models, while mechanistically interpretable, struggle to balance computational cost and forecast accuracy. Although artificial intelligence (AI) has advanced rapidly in meteorological forecasting, most existing AI models are not optimized for atmospheric environmental prediction and rely heavily on gridded inputs, limiting their ability to integrate site observations and their operational applicability. To overcome these limitations, we develop BiXiao, a new-generation AI-based atmospheric environmental forecasting model. BiXiao features a heterogeneous architecture with non-continuous grids, coupling independent meteorological and environmental modules for synergistic use of multi-source data. The meteorological module employs a 3D Swin Transformer (Swin3D) to process structured meteorological fields, while the environmental module directly assimilates discrete station data, enabling operational urban-scale forecasts. Testing in the Beijing-Tianjin-Hebei region shows that BiXiao completes 72-hour forecasts for six major pollutants across all key cities within 30 seconds. Compared with mainstream numerical models (CAMS and WRF-Chem), BiXiao achieves substantially higher computational efficiency and forecast accuracy, particularly during heavy pollution events.

1 Introduction

High-precision and efficient atmospheric environmental forecasting is essential for protecting public health and supporting environmental management. Traditional physics-based numerical models, while mechanistically interpretable, have long struggled to achieve an effective balance between computational efficiency and forecast accuracy (Meng et al., 2023). For decades, such models have served as the research cornerstone in this field, providing a physically grounded predictive framework by solving complex partial differential equations that describe atmospheric dynamics, physical, and chemical processes (Wang et al., 2008, 2024). However, this mechanism-driven paradigm faces significant challenges in reconciling accuracy with computational cost, posing a fundamental limitation to further model advancement (Düben et al., 2021).

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

The rise of artificial intelligence (AI) has opened a transformative pathway for Earth system modeling. This evolution is particularly evident in meteorological forecasting. Early studies explored the potential of fundamental deep learning architectures: convolutional neural networks (CNNs) have shown strong spatial feature extraction capabilities for short-term precipitation nowcasting (Ayzel et al., 2020), while recurrent neural networks (RNNs) and their variants effectively captured temporal dependencies in dynamic meteorological processes (Xiao et al., 2019; Shi et al., 2015). The development of the ConvLSTM architecture (Shi et al., 2015) marked a milestone by integrating convolutional and recurrent mechanisms to jointly model the spatiotemporal evolution of atmospheric systems.

With advances in deep learning theory and computational capacity, research has entered the era of large-scale pre-trained models. Huawei's Pangu-Weather model (Bi et al., 2023), based on a three-dimensional Transformer, achieved accuracy comparable to traditional numerical weather prediction (NWP) systems while improving computational efficiency by orders of magnitude. Google's GraphCast (Lam et al., 2023) introduced graph neural networks as an alternative paradigm. The MetNet series (Sønderby et al., 2020) demonstrated kilometer-scale precipitation forecasts through multi-source data fusion, and FourCastNet (Pathak et al., 2022) extended data-driven forecasting boundaries with adaptive Fourier neural operators. More recently, generative AI techniques such as diffusion models have shown promising potential—GenCast (Price et al., 2023) enables probabilistic forecasting and uncertainty quantification, while Microsoft's Aurora (Bodnar et al., 2025) extends predictive capabilities to atmospheric composition, revealing the broad applicability of foundation models in Earth system science.

Despite their impressive capabilities, these advanced AI models remain fundamentally constrained by the characteristics of their training datasets. For instance, Aurora relies heavily on CAMS reanalysis data (Inness et al., 2019), resulting in three core limitations: insufficient spatial resolution(~0.4° grids fail to capture urban-scale pollution gradients), incomplete variable coverage (lack of surface concentration fields for key pollutants), and the absence of real-time data assimilation. This dependence on structured gridded data highlights a fundamental disconnect between AI-driven forecasting systems and the actual observational network—precisely the gap that must be bridged for fine-scale operational environmental forecasting. Consequently, existing AI models often fall short of the high demands for precision, timeliness, and reliability required in operational air-quality prediction and management (Zhou et al., 2017; Lin et al., 2020).

These technical constraints, coupled with escalating environmental management needs, form a pronounced contradiction. As pollution control efforts deepen, the demand for refined and accurate forecasts continues to rise (Meng et al., 2023). In the context of China's Air Pollution Prevention and Control Action Plan, the strategic focus has shifted from "weather-driven" to "emission-control-oriented" approaches, imposing stricter requirements on forecasting technology (Meng et al., 2025).

To address these challenges, we propose BiXiao, an innovative AI-based atmospheric environmental forecasting model. BiXiao introduces a heterogeneous architecture with non-continuous grids, inspired by the offline coupling framework of traditional numerical systems such as WRF-CMAQ (Wang et al., 2024). The design incorporates two independent modules: a meteorological module that processes structured grid data, and an environmental module capable of directly assimilating discrete observation data. This architecture fundamentally overcomes the dependence of existing AI models on regular grids and enables effective utilization of monitoring-network data for fully operational forecasts. Moreover, the purely data-driven

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

inference allows BiXiao to achieve second-level prediction speeds, providing exceptional adaptability for real-world applications across diverse operational scenarios.

The remainder of this paper is organized as follows. Section 2 presents the model architecture and training methodology. Section 3 describes the datasets and study region. Section 4 evaluates model performance and compares it with mainstream systems. Section 5 validates BiXiao's capabilities through a representative pollution episode, and Section 6 summarizes the conclusions and outlines future perspectives.

2 Model Description

2.1 Model Name

The name "BiXiao" originates from a classical Chinese term for "blue sky," derived from the famous verse by Liu Yuxi. The model is thus named to symbolize an aspiration toward a cleaner and more beautiful atmospheric environment..

2.2 Model Architecture

2.2.1 Overall Framework

BiXiao adopts an "offline" architecture, consisting of two decoupled components—a meteorological module and an environmental module. This configuration is conceptually similar to the "stand-alone" mode in the WRF-CMAQ numerical modeling framework, in which the environmental module operates independently without feedback to the meteorological computations. Such a design ensures modularity and plug-and-play flexibility.

To mitigate the excessive dependence of conventional AI models on structured gridded data and to better accommodate the characteristics of observational datasets, BiXiao employs a heterogeneous architecture with non-continuous grids. Specifically, the meteorological module follows the regular gridded structure of traditional numerical weather prediction (NWP) systems, while the environmental module operates on discrete grids, enabling it to utilize irregular, station-based observations without requiring complete gridded datasets. This design allows the environmental module to be trained and inferred directly from scattered observation data.

During the training phase, the two modules are trained separately. The meteorological module, implemented as an autoregressive large model, is trained using reanalysis datasets such as ERA5. The environmental module is trained under idealized meteorological backgrounds using both reanalysis data (e.g., ERA5) and in situ air-quality observations.

During the inference phase, the two modules operate sequentially. The meteorological module first generates gridded meteorological forecasts, which provide background fields for the environmental module. The environmental module then integrates these meteorological fields with discrete observational data to predict pollutant concentrations on discrete grids. Specifically, the environmental module takes the initial meteorological state at time T+0 and its one-step prediction at T+1 as inputs, together with environmental observations at T+0, to produce pollutant forecasts at T+1.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

2.2.2 Meteorological Module

The Swin Transformer is a vision-based deep learning architecture built upon the Transformer framework. Its core concept involves partitioning input images into multiple windows and applying self-attention within each window to extract features efficiently (Liu et al., 2021). Compared with the traditional two-dimensional Swin Transformer (Swin2D), the three-dimensional Swin Transformer (Swin3D) extends this approach into 3D space, enabling it to capture volumetric and altitude-dependent features that align well with the three-dimensional structure of the atmosphere. Owing to these advantages, Swin3D is highly suitable for atmospheric modeling and has been successfully implemented in the large-scale meteorological model Pangu-Weather (Bi et al., 2023).

To effectively address the joint prediction of upper-air (3D) and surface (2D) meteorological fields, this study develops a 3D Swin Transformer-based deep neural network following an encoder-decoder architecture. The key idea is to establish a shared latent space that learns the deep representations of both upper-air and surface features, capturing their intrinsic physical relationships for accurate future-state prediction. As illustrated in Fig. 1a, the module comprises three stages: multimodal data embedding and fusion, hierarchical feature extraction and interaction, and feature decoding and output generation.

2.2.3 Multimodal Data Embedding and Fusion

To reconcile the dimensional inconsistencies among input data, parallel Patch Embedding layers are designed at the model entrance. For upper-air meteorological data (dimensions: BatchSize $\times$ Variables $\times$ Levels $\times$ Height $\times$ Width), 3D convolutional layers (Conv3D) are applied for feature extraction and downsampling. For surface data (dimensions: BatchSize $\times$ Variables $\times$ Height $\times$ Width), 2D convolutional layers (Conv2D) are used. The resulting surface feature tensors are then expanded along the vertical dimension and concatenated with upper-air tensors to form a unified 3D feature tensor (BatchSize $\times$ Channels $\times$ Depth $\times$ Height $\times$ Width). This design achieves spatial alignment and preliminary fusion of multi-modal meteorological features.

2.2.4 Hierarchical Feature Extraction and Interaction

The fused feature tensor is processed by the model's core feature extractor, which consists of multiple BasicLayers, each composed of several SwinTransformerBlock3D modules. The main innovation of these blocks lies in the 3D window-based multi-head self-attention mechanism. In this approach, the feature map is partitioned into a set of non-overlapping local 3D windows (e.g., $2 \times 4 \times 4$), and self-attention is computed within each window, substantially reducing computational complexity.

To promote cross-window information exchange, the model alternates between regular and shifted window partitioning strategies in consecutive layers. In addition, learnable relative position bias is introduced to encode 3D spatial relationships, enabling the model to better capture dependencies across longitude, latitude, and altitude. Through hierarchical stacking, the receptive field progressively expands, allowing the model to learn global-scale atmospheric patterns from local features and effectively capture the complex spatiotemporal dynamics of meteorological fields.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

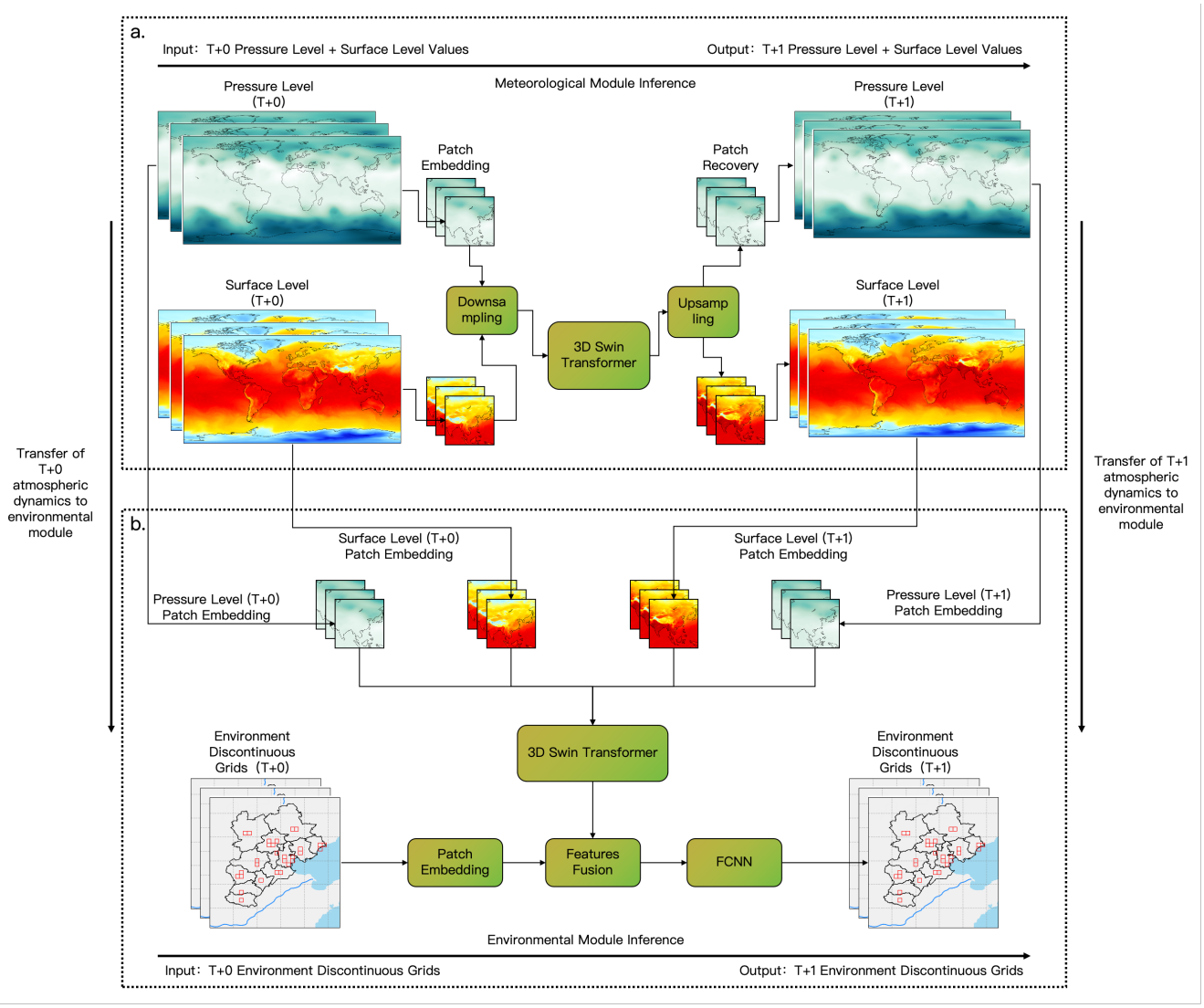

Figure 1. Schematic Diagram of the "BiXiao" Large Model Architecture

2.2.5 Decoding and Output

The decoder (Patch Recovery) performs operations symmetric to those of the encoder. The deeply transformed latent features are separated along the vertical dimension to reconstruct distinct representations for upper-air and surface components. Subsequently, 3D transposed convolutions (ConvTranspose3D) are applied to upsample the upper-air features, restoring their original vertical and horizontal resolutions, while 2D transposed convolutions (ConvTranspose2D) are used for surface features. Finally, the model outputs the predicted upper-air and surface variables simultaneously, achieving end-to-end, multi-variable meteorological forecasting.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

2.2.6 Environmental Module

The environmental module of BiXiao is designed to infer future atmospheric composition by integrating the spatiotemporal evolution of meteorological fields with the current environmental state. The key principle is to extract multi-temporal, three-dimensional meteorological features that capture the evolution of atmospheric dynamics and use these as "background fields" to drive environmental predictions. This approach reflects the underlying physical mechanism in which environmental variables are governed by meteorological forcing, while leveraging deep learning to simulate this interaction.

2.2.7 Multi-temporal Meteorological Feature Extraction

Although the environmental and meteorological modules are trained independently, the environmental module reuses the 3D Swin Transformer architecture from the meteorological component for feature extraction. Instead of focusing on the stepwise inference from T+0 to T+1, this module processes meteorological fields at multiple time points (T+0 and T+1) through the same patch embedding, hierarchical feature extraction, and interaction procedures described in Section 2.2.2. This enables the model to capture the temporal evolution of meteorological conditions and extract consistent dynamic features across time.

2.2.8 Meteorology-Environment Feature Fusion

A major innovation of the environmental module lies in its efficient feature fusion mechanism, which integrates meteorological background information with environmental observations. The fusion process consists of two stages. First, the meteorological features extracted by Swin3D are compressed via dimensionality reduction convolutions to match the dimensionality of environmental data. Then, the compressed meteorological feature vectors are concatenated with environmental observation vectors to form a unified representation of fused features.

This design carries clear physical significance: meteorological features represent the external dynamic forcing that drives environmental changes, while environmental observations describe the current initial state. Their combination allows the model to simultaneously account for both external forcing and initial conditions, aligning with the fundamental physical principles of atmospheric evolution. To enhance cross-domain interactions, a feature interaction mechanism is implemented through nonlinear transformations in fully connected layers, facilitating information exchange between the two feature sources.

2.2.9 Environmental Evolution Network

Building upon the fused feature representation, the environmental module employs a multi-layer perceptron (MLP) network to map these features to future environmental states. The network follows an encoder-decoder structure, where the hidden layers apply SiLU activation and Dropout regularization to balance expressiveness and generalization.

The output layer directly maps predictions onto station-level dimensions, enabling a transformation from continuous feature space to discrete observation sites. This design eliminates the error accumulation often caused by grid interpolation in traditional methods, allowing the model to learn directly from site-level observations and thus substantially improving prediction accuracy.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

3 Model Training and Validation

3.1 Data and Study Area

3.1.1 Study Area

In this study, the Beijing-Tianjin-Hebei region (34.75°N-43°N, 112.5°E-120.75°E) is selected as the research area. This region is one of the most economically developed and densely populated in China. Furthermore, it is one of the first regions in China to deploy atmospheric environment forecasting services. In the future, precise forecasting of air pollutants in the Beijing-Tianjin-Hebei region will continue to be of significant importance, and atmospheric environment forecasting remains a foundational task for environmental protection in the area.

The Beijing-Tianjin-Hebei region is a high-incidence area for haze events, and its pollution characteristics are representative of other major urban areas in China. Conducting model training and testing in this region will help improve the model's adaptability and provide a reliable basis for its broader application across the country. Moreover, compared to direct nationwide training and forecasting, focusing on a regional study reduces the computational resource demands, thus facilitating the efficient development and optimization of the model.

3.1.2 Meteorological Field Data

The "BiXiao" model uses ERA5 reanalysis data as the driving data for the meteorological module. ERA5 is generated using the four-dimensional variational (4D-Var) data assimilation and model forecasting methods from the ECMWF Integrated Forecasting System (IFS), version CY41R2. The data has a horizontal resolution of 0.25°×0.25° and is widely used in atmospheric science (Hersbach et al., 2020). The ERA5 data used in this study includes both three-dimensional and two-dimensional variables.

Three-Dimensional Variables: These primarily include vertical velocity, temperature, and geopotential height, with values obtained from a global model on 137 model layers. The data is then interpolated through the FULL-POS method in IFS to 37 pressure levels.

Two-Dimensional Variables: These include surface (or single-layer) data, such as precipitation and top-of-atmosphere radiation, with values provided as vertical integrals over the entire atmospheric depth.

In terms of temporal coverage, this study uses ERA5 data from January 2014 to March 2024, extracting meteorological variables at UTC times 00, 06, 12, and 18. Since the primary task of the environmental module is to predict various chemical environmental elements at ground-based observation stations, this study focuses on six prressure-level atmospheric variables at six levels (500, 600, 700, 850, 925, and 1000 hPa), as well as five surface variables, totaling 41 elements. These specific variables are listed in Table 1.

In terms of data area selection, considering that this study focuses on the Beijing-Tianjin-Hebei region and to improve training efficiency, the original ERA5 data with a spatial resolution of 0.25° (a total of 721 × 1440 grid points) is cropped. The data driving the meteorological module covers the area from 14.75° N to 55° N and 65° E to 140° E, as shown in the full map in Figure 2 (left), corresponding to 160 × 300 grid points. The data driving the environmental module is further focused on the

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

Table 1. List of meteorological variables used in the model.

short-name	long-name	levels
u	U-component of wind	pressure levels
v	V-component of wind	pressure levels
w	Vertical velocity	pressure levels
t	Temperature	pressure levels
q	Specific humidity	pressure levels
z	Geopotential	pressure levels
t2m	2 m temperature	surface
u10m	10 m U-component of wind	surface
v10m	10 m V-component of wind	surface
d2m	2 m dewpoint temperature	surface
sp	Surface pressure	surface

Beijing-Tianjin-Hebei region (34.75° N-43° N, 112.5° E-120.75° E), corresponding to 34 × 34 grid points, as shown in Figure 2 (right).

In addition, the meteorological variables used in the environmental module are consistent in type with those in the meteorological module, but differ slightly in pressure levels. Specifically, only the four lower atmospheric levels (700, 850, 925, and 1000 hPa) that are closer to the surface are used to better capture near-surface meteorological conditions influencing air quality.

3.1.3 Environmental Field Data

This study uses hourly observational data from 79 environmental monitoring stations in the Beijing-Tianjin-Hebei region, covering the period from January 2021 to March 2024, as training data (specific station information can be found in the Appendix). The observational data includes the six major pollutants specified in China's "Ambient Air Quality Standards," namely sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), and particulate matter ($PM_{2.5}$ and PM_{10}).

To ensure compatibility with the ERA5 meteorological data at a 0.25° resolution, the station observation data is gridded. Specifically, each station's observational data is assigned to the nearest ERA5 grid based on the station's latitude and longitude. For grids with only one station, the atmospheric environmental observation value directly uses the data from that station. If multiple monitoring stations exist within a grid, the air quality level of the grid is represented by the average value of all station data. After data processing, the 79 stations in the Beijing-Tianjin-Hebei region are mapped to 29 valid grids, as shown by the red grid points in Figure 2 (right).

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

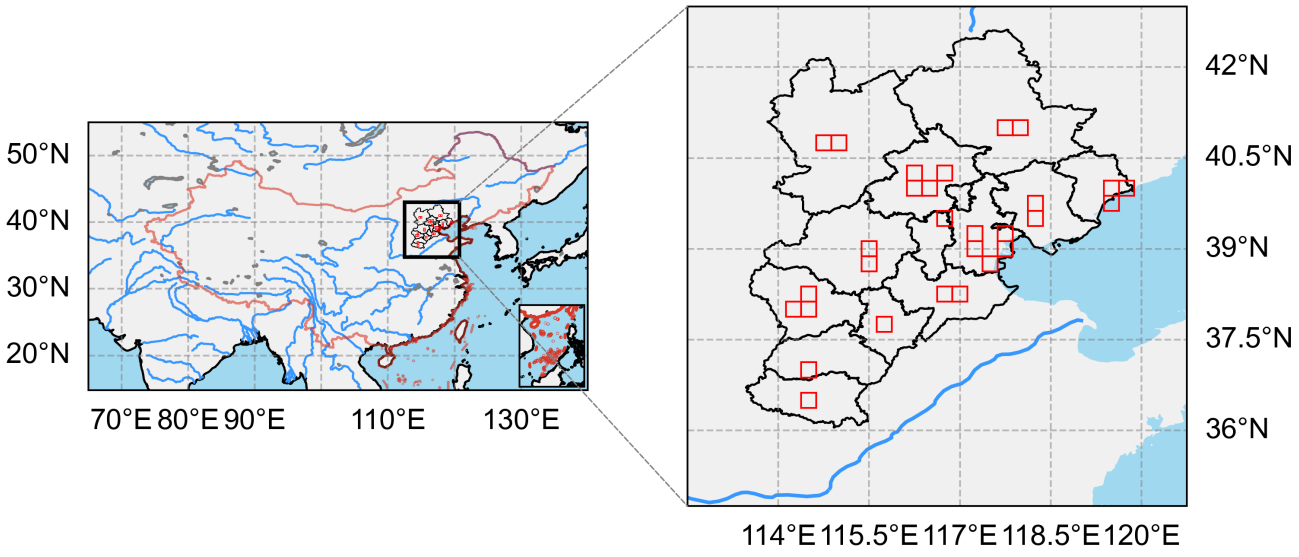

Figure 2. Meteorological Data Range for the Meteorological Module (Left) and Meteorological Data Range and Environmental Grid Distribution for the Environmental Module (Right)

3.2 Model Training

Before training the model, all input data, including ERA5 meteorological fields and environmental station data, need to undergo preprocessing. The preprocessing steps mainly include calculating the mean and variance required for data normalization, ensuring the stability of data distribution during the model training process. After preprocessing, all data are divided into two parts: the "training set" for training and the "test set" for testing.

For the selection of training data, the meteorological module's training data include ERA5 data from January 2014 to March 2024, using data from four time points: UTC 00, 06, 12, and 18. The environmental module's training data use hourly data from January 2021 to March 2024. For the test set selection, to ensure objectivity, five consecutive days of data are randomly selected each month as the test set, and this portion of data does not participate in model training and validation. Additionally, to further evaluate the model's performance in forecasting typical heavy pollution events, extra data from an ozone pollution event in the Beijing-Tianjin-Hebei region in June 2022 and a $PM_{2.5}$ pollution event in October 2023 are included in the "test set," while the remaining data are used as the training set.

The model training uses smoothL1 loss as the loss function and employs supervised learning to optimize the model's single-step autoregressive prediction. The model construction and training are based on the PyTorch framework, and the AdamW optimizer is used in conjunction with a linear decay scheduler to dynamically adjust the learning rate, improving training convergence efficiency. The meteorological module and environmental module are trained independently, with substantial differences in training duration and computational resource requirements. For example, the meteorological module is trained

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

using 8 Nvidia RTX 4090 GPUs, completing 150 epochs of training in several hours, with the final loss reduced to 0.02. The environmental module is trained on a single Nvidia RTX 4090 GPU, requiring 24 hours to complete 200 epochs of training, with the final loss reduced to 0.01. While further training could still lower the loss, the prediction accuracy on the validation set did not show significant improvement, indicating that a lower loss does not necessarily correlate with better generalization ability.

3.3 Model Validation

3.3.1 Validation Metrics

The model performance is evaluated using three metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Pearson Correlation Coefficient (PCC).

- RMSE reflects the average deviation between the model's predicted values and the actual values. It is more sensitive to larger errors and provides a comprehensive measure of the overall deviation of the predicted values.
- MAE is used to assess the average absolute error between the predicted and actual values, focusing on the actual magnitude of errors.
- PCC measures the linear correlation between the predicted values and the actual values. A PCC value closer to 1 or -1 indicates a strong linear correlation between the predicted and actual values, while a value closer to 0 indicates a weak linear correlation.

These three metrics are common model evaluation parameters and have been widely used in many related studies. The specific calculation methods for the three parameters are as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}, \tag{1}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|P_i - O_i|. \tag{2}$$

$$\mathrm{PCC} = \frac{\sum_{i=1}^{N}(P_i - \overline{P})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(P_i - \overline{P})^2}\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}} \tag{3}$$

3.3.2 Validation Experiment

As mentioned earlier, to avoid having data samples in the test set that are similar to those in the training set, this study strictly divides the training and test datasets to ensure their independence. In the validation experiment, simulation predictions are made for the selected data samples from the test set, and the results are compared with the measured data. For the 5-day test set, ERA5

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

reanalysis data for the first two days at UTC 00, 06, 12, and 18 hours are used as the initial input for the meteorological fields (T+0), and the corresponding environmental observation data for the same times are used as the initial state of the environmental fields (T+0). The model uses a 6-hour time step for inference, generating pollution concentration simulation results for the next 72 hours. The simulation results are then compared with the remaining 3 days of measured data. The entire validation experiment includes 308 forecast times, and the model's prediction results are verified using the measured data. The RMSE, MAE, and PCC values for the forecast results at different time steps on all grids are calculated and used to assess the model's forecasting ability.

Figure 3 shows the overall forecast performance of the "BiXiao" model. The forecast results are categorized and statistically analyzed according to the latitude and longitude of the grids and the administrative divisions of the Beijing-Tianjin-Hebei (Beijing-Tianjin-Hebei) region. The analysis focuses on the forecast correlation coefficients (PCC) for O_3, $PM_{2.5}$, and PM_{10} at 6 h, 48 h, and 72 h forecast times for different administrative regions. Figure 4 presents the PCC, RMSE, and MAE for six pollutants (O_3, $PM_{2.5}$, PM_{10}, NO_2, CO, SO_2) at different forecast time steps across all grids.

Among the six pollutants, "BiXiao" performs best in simulating O_3. For forecasts from 6 h to 72 h, the PCC for O_3 consistently outperforms other pollutants, with the average O_3 PCC across 29 grids reaching as high as 0.91 at 6 h. The RMSE and MAE for O_3 forecasts remain moderate, at 26±6 μg m-3 and 20±5 μg m-3, respectively. Regional analysis shows that the forecast performance for O_3 is better in the central areas of Beijing-Tianjin-Hebei, such as Beijing, Tianjin, Langfang, and Baoding, where the PCC at all forecast times is slightly higher than in other areas of the region.

For particulate matter concentration forecasts, the performance of the "BiXiao" model for $PM_{2.5}$ and PM_{10} is similar, with slightly better performance for $PM_{2.5}$. At the 6-hour forecast, the PCC for $PM_{2.5}$ and PM_{10} are 0.86 and 0.79, respectively. As the forecast time increases, the correlation gradually decreases, with the PCC for PM_{10} showing a more pronounced drop after 48 h and remaining slightly lower than for $PM_{2.5}$. Since PM_{10} concentration includes $PM_{2.5}$, the RMSE and MAE for PM_{10} are higher than those for $PM_{2.5}$ across all forecast periods, with a noticeable increase in RMSE and MAE for PM_{10} after 48 h. The reasons behind this trend require further investigation. Spatial distribution analysis indicates that "BiXiao" performs better in simulating particulate matter in the southern and central regions of Beijing-Tianjin-Hebei.

For other gaseous pollutants, the "BiXiao" model shows acceptable forecast ability for NO_2 and CO, with a relatively steady decrease in PCC as the forecast time increases. The RMSE and MAE for these two pollutants also increase gradually with forecast time, though the increase is relatively small. Specifically, the RMSE for NO_2 and CO are 14±3 μg m-3 and 0.32±0.04 mg m-3, respectively, while the MAE for NO_2 and CO are 11±2 μg m-3 and 0.22±0.04 mg m-3, respectively. Compared to other pollutants, the forecast correlation for SO_2 is the lowest, but its RMSE and MAE show relatively stable changes.

Overall, these results suggest that "BiXiao" excels in simulating O_3 and particulate matter, while there is still considerable room for improvement in simulating SO_2.

Pearson Correlation Coefficient (PCC)

Root Mean Square Error (RMSE)

Mean Absolute Error (MAE).

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

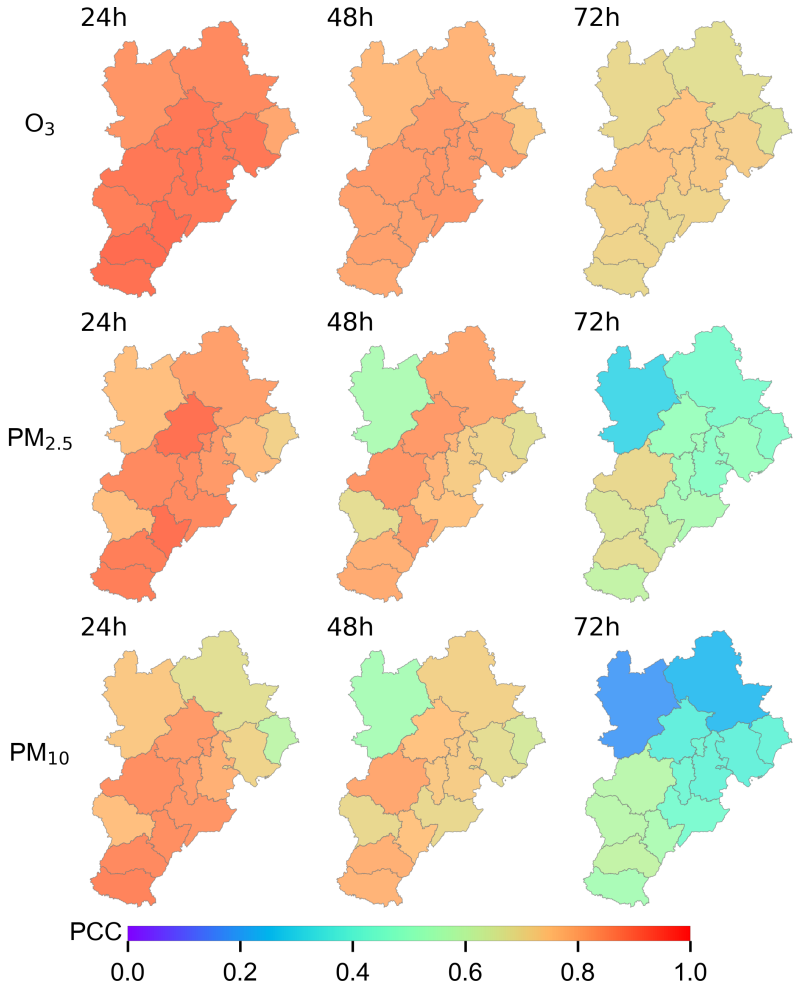

Figure 3. Pearson Correlation Coefficient (PCC) between the simulated and observed values for "BiXiao," where the first column represents the 24-hour forecast, the second column represents the 48-hour forecast, and the third column represents the 72-hour forecast.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

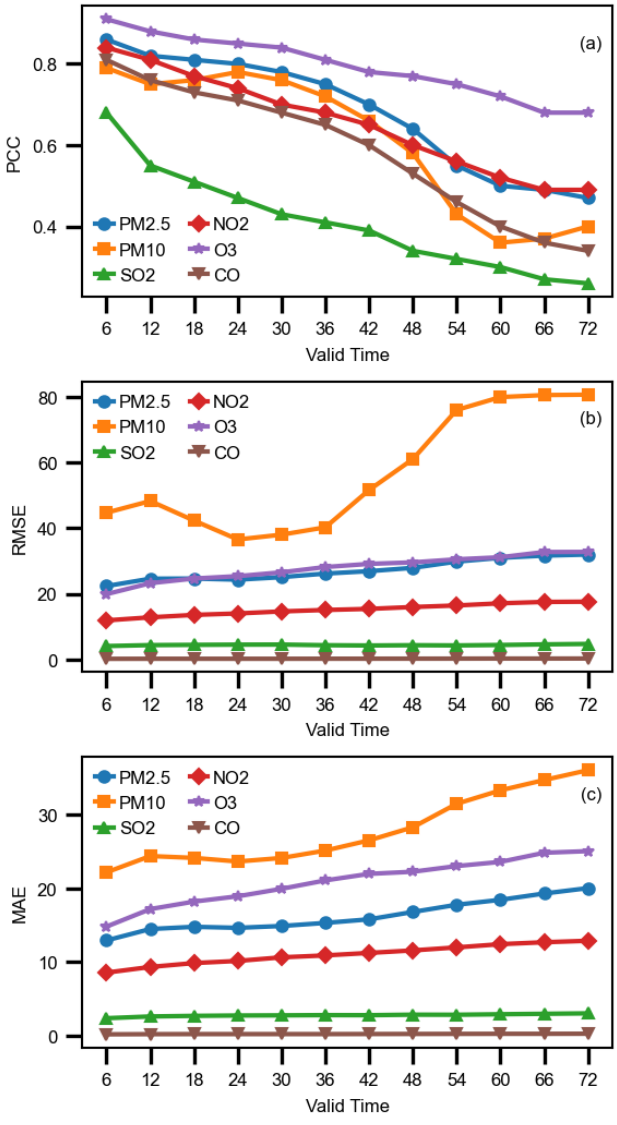

Figure 4. the following between the predicted values of "BiXiao" and the actual observed values:

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

4 Routine Environmental Forecasting — "BiXiao" vs CAMS

4.1 Experimental Setup

ECMWF's CAMS provides global atmospheric environmental monitoring and forecasting, widely recognized in the industry for its high forecasting capability and broad representativeness (Inness et al., 2019). Therefore, this study selects CAMS forecast data as a benchmark to evaluate the performance of the "BiXiao" model in operational environmental forecasting.

When comparing, two main issues must be considered: spatial and temporal matching. In terms of spatial matching, environmental forecast results from CAMS are extracted from the corresponding locations based on the latitude and longitude of the 29 discrete environmental grids in "BiXiao." For temporal matching, both CAMS and BiXiao provide 6-hour forecast intervals, and the comparison is made over the 72-hour forecast period.

It is worth noting that since CAMS only provides forecast results at UTC 00:00 and 12:00 each day, the comparison is conducted using only the common start times between CAMS and BiXiao. Additionally, as CAMS provides surface concentrations of $PM_{2.5}$ and PM_{10} only, whereas "BiXiao" forecasts column concentrations of pollutants, the comparison is focused only on particulate matter ($PM_{2.5}$ and PM_{10}). The experiment uses observed environmental grid data to validate the forecast results from both CAMS and BiXiao to assess the differences in forecasting performance between the two models.

4.2 Experimental Results

4.2.1 Grid-based Statistical Comparison

The comparison of $PM_{2.5}$ and PM_{10} forecast results between "BiXiao" and CAMS across the 29 valid grids is shown in Figure 5. In the 6-hour forecast, "BiXiao" outperforms CAMS, with a Pearson Correlation Coefficient (PCC) of 0.87 for $PM_{2.5}$ and 0.82 for PM_{10}, significantly higher than CAMS's PCC of 0.60 for $PM_{2.5}$ and 0.45 for PM_{10}. The RMSE (MAE) for $PM_{2.5}$ and PM_{10} in "BiXiao" are 21.41 $\mu g/m^3$ (12.45 $\mu g/m^3$) and 41.55 $\mu g/m^3$ (20.94 $\mu g/m^3$), respectively, much lower than the RMSE (MAE) of 40.86 $\mu g/m^3$ (27.7 $\mu g/m^3$) and 72.03 $\mu g/m^3$ (43.25 $\mu g/m^3$) in CAMS.

As the forecast time increases, the forecast performance of both models decreases. Overall, for most forecast times, the PCC between the forecasted pollutant concentrations and observations decreases as forecast duration increases, with "BiXiao" maintaining higher correlation than CAMS. In the longer forecast period, i.e., the 72-hour forecast, "BiXiao" achieves PCC values of 0.44 for $PM_{2.5}$ and 0.40 for PM_{10}, while CAMS's PCC for $PM_{2.5}$ and PM_{10} are 0.45 and 0.31, respectively.

From the error perspective, the RMSE and MAE for $PM_{2.5}$, as well as the MAE for PM_{10}, consistently show that "BiXiao" has lower deviations compared to CAMS across all forecast times. This indicates that "BiXiao" has an overall superior forecasting ability for particulate pollutants, particularly in the short-term forecasts.

(a) $PM_{2.5}$ Pearson Correlation Coefficient (PCC)

(b) PM_{10} Pearson Correlation Coefficient (PCC)

(c) $PM_{2.5}$ Root Mean Square Error (RMSE)

(d) PM_{10} Root Mean Square Error (RMSE)

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

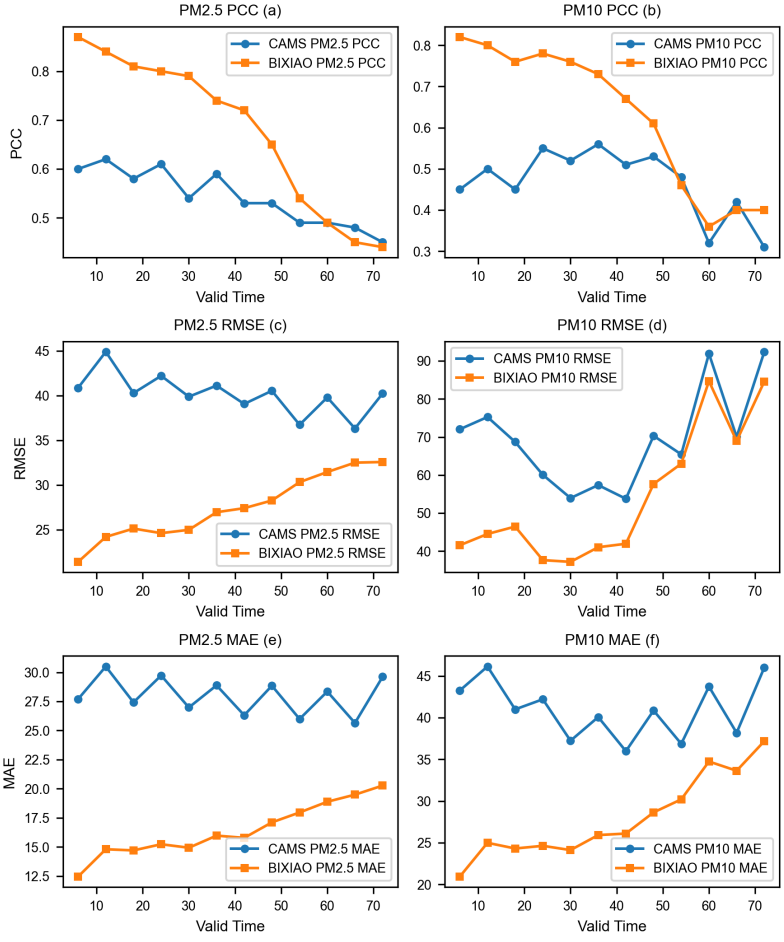

Figure 5. Comparison of forecast accuracy between the "BiXiao" model and CAMS:

(e) $PM_{2.5}$ Mean Absolute Error (MAE)

(f) PM_{10} Mean Absolute Error (MAE)

4.2.2 Effective Grid Comparison

On the 29 valid grids, the forecast performance of the "BiXiao" model for $PM_{2.5}$ and PM_{10} was compared with the numerical model CAMS at the same forecast time, as shown in Figure 6. The comparison results are presented using interpolation, where the difference in the Pearson Correlation Coefficient (PCC) is calculated as "BiXiao" minus "CAMS," and the difference in the Root Mean Square Error (RMSE) is calculated as "CAMS" minus "BiXiao."

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

For $PM_{2.5}$ forecasting, within the first 48 hours, BiXiao consistently outperforms CAMS with higher PCC and smaller RMSE across all grids. After 48 hours, BiXiao's PCC is slightly lower than that of CAMS in the central and eastern regions of the Beijing-Tianjin-Hebei (Beijing-Tianjin-Hebei) area, but the RMSE is still smaller than CAMS in most grids.

For PM_{10} forecasting, the evaluation results in the first 24 hours show similar characteristics to those of $PM_{2.5}$, with BiXiao achieving higher PCC and smaller RMSE in all grids compared to CAMS. After 24 hours, BiXiao continues to outperform CAMS in PCC, and most regions also show smaller RMSE than CAMS, with a few grids in the southern region showing slightly worse performance. After 48 hours, BiXiao's PCC in the southern and northern regions is slightly lower than CAMS, with similar trends observed in the RMSE.

This trend is speculated to be related to the method of combining environmental grids, where each environmental grid represents the average data of all the monitoring stations within that grid. In the central urban areas, where monitoring stations are more concentrated, the data in each grid is more stable, leading to better model performance in these areas. Overall, BiXiao shows a significant advantage in the first 48 hours of forecasting. After 48 hours, the forecast ability in some regions gradually declines to the point where it becomes comparable to or slightly worse than CAMS.

5 Heavy Pollution Case Forecast—"BiXiao" vs WRF-Chem

5.1 Experimental Setup

In this study, the WRF-Chem model, developed jointly by the NOAA Forecast Systems Laboratory (FSL) and the Pacific Northwest National Laboratory (PNNL), is selected as a representative atmospheric environment model for comparison. The analysis aims to compare the performance of WRF-Chem and the "BiXiao" model in predicting typical fine particulate matter ($PM_{2.5}$) and ozone (O_3) pollution events.

The two pollution events selected for this study are the O_3 pollution event from June 24-27, 2022, and the $PM_{2.5}$ pollution event from October 28-31, 2023. Both WRF-Chem and "BiXiao" models are used to simulate these two events, and the results are compared with observed data.

The WRF-Chem model is set up with two nested grids. The first grid consists of 91×74 horizontal grid points, while the second grid, which covers the Beijing-Tianjin-Hebei (Beijing-Tianjin-Hebei) region, has a resolution of 27 km and includes 97×106 grid points. The boundary conditions for the second grid are provided by the first grid. The simulations for the two pollution events start at 00:00 on June 21, 2022, and October 25, 2023, respectively. The first 72 hours are used for model integration, and the subsequent 72 hours serve as the effective forecast period.

In this study, WRF-Chem version 4.1 is used with the RADM2 chemical mechanism and the MADE/SOGARM aerosol parameterization scheme. Anthropogenic emission sources are based on the 2023 release of the Chinese Multiscale Emission Inventory (MEIC v1.4) (http://www.meicmodel.org/).

The "BiXiao" model uses ERA5 data for meteorological initial conditions at 00:00 on June 24, 2022, and October 28, 2023, respectively, and uses environmental observational data from the same times as environmental initial conditions. The model runs

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

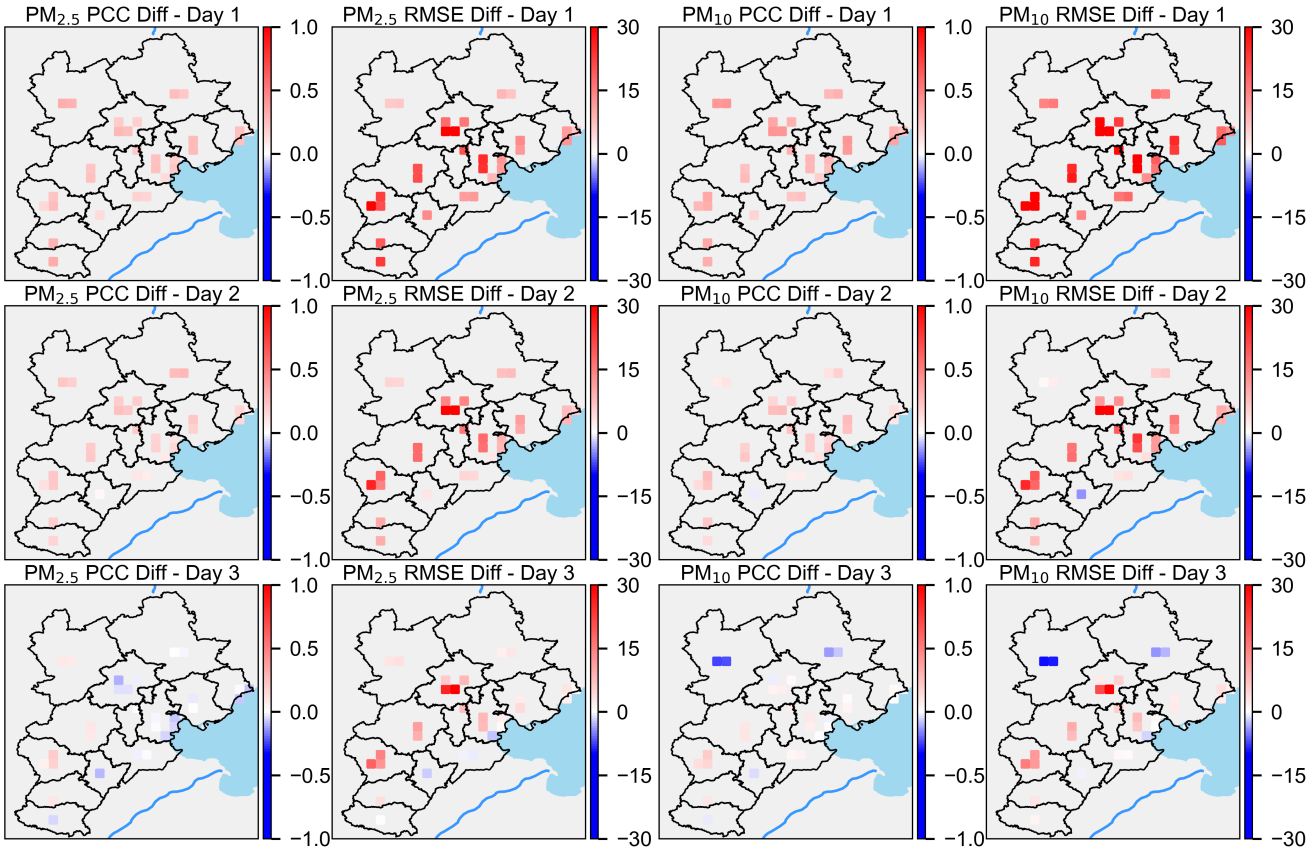

Figure 6. Comparison of daily average forecast performance for $PM_{2.5}$ and PM_{10} between "BiXiao" and CAMS across different grids for 1-day, 2-day, and 3-day forecast periods: Pearson Correlation Coefficient (PCC) difference (BiXiao minus CAMS) and Root Mean Square Error (RMSE) comparison (CAMS minus BiXiao).

with a 6-hour time step to forecast the next 72 hours. The experimental results are compared with observed environmental grid data to evaluate the forecast performance of both models during these typical pollution events.

5.2 Experimental Results

5.2.1 Ozone Pollution Case Study

In June 2022, frequent extreme heat events occurred across northern China, coinciding with increased ozone (O_3) pollution. The temporal and spatial distribution of ozone pollution, including its intensity, significantly intensified during this period, with a strong correlation between temperature anomalies and O_3 pollution in the Beijing-Tianjin-Hebei (BTH) region (Yang et al., 2025). This study focuses on a specific ozone pollution event from June 24 to 27, 2022, to assess the forecasting capability of the BiXiao model. During this event, peak O_3 concentrations in the BTH region occurred at 14:00 local time (UTC 06:00) on June

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

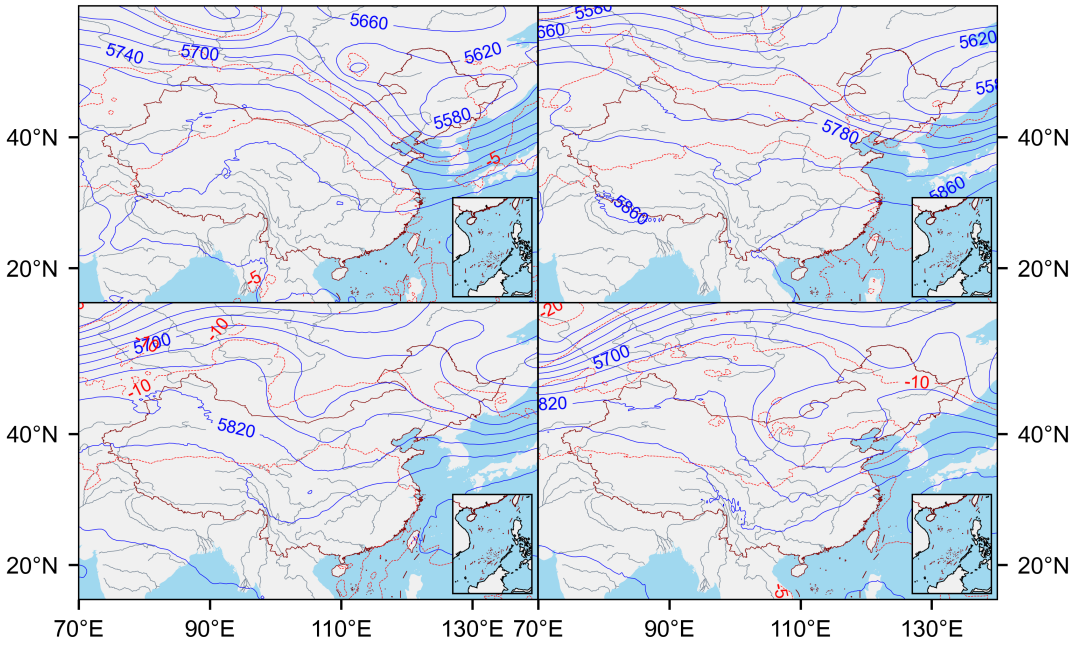

Figure 7. 500 hPa geopotential height and temperature map for June 24-27, 2022, 00:00 UTC.

25 and 26, with grid-averaged O_3 concentrations reaching 270.0 μg m^{-3} and 163 μg m^{-3}, respectively, surpassing the Chinese air quality standards. During this period, the region was under the influence of a weather pattern following an upper-level trough (Figure 7), with clear skies, no precipitation, high temperatures, and weak winds. Afternoon surface temperatures in the grid averaged above 30°C, peaking at 37°C on June 25, and the relative humidity at the surface was 76% in the afternoon of the same day. Strong sunlight and high temperatures accelerated photochemical reactions, fostering O_3 formation. The near-surface wind speed peaked at 4.8 m s^{-1} at 06:00 UTC on June 25 and 26, with wind speeds not exceeding 3.7 m s^{-1} at other times during the event, which hindered pollutant dispersion.

Figure 8 presents the forecast results of the BiXiao and WRF-Chem models, initialized at 00:00 UTC on June 24, 2022. Both models predict the trend of O_3 concentration well, though WRF-Chem tends to underestimate the values compared to observations. Throughout all forecast hours, the mean forecast values from BiXiao are closer to the observed values than those from WRF-Chem. During this pollution event, at 14:00 local time on June 25, 2022 (UTC 06:00), which corresponds to the 30th hour of the model forecast, O_3 concentration reached its peak, with a correlation coefficient of 0.82 between BiXiao and observations, while WRF-Chem had a value of -0.34. Six hours later, at the 36th hour of the forecast, O_3 concentration reached its secondary peak of 201 μg m-3, with a correlation coefficient of 0.55 for BiXiao and 0.02 for WRF-Chem. At these two times, the mean absolute errors of BiXiao were 85.33 μg m-3 and 56.37 μg m-3, respectively, lower than WRF-Chem's 181.38 μg m-3 and 91.28 μg m-3. For all other forecast times between 0 and 72 hours, the average absolute error between BiXiao

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

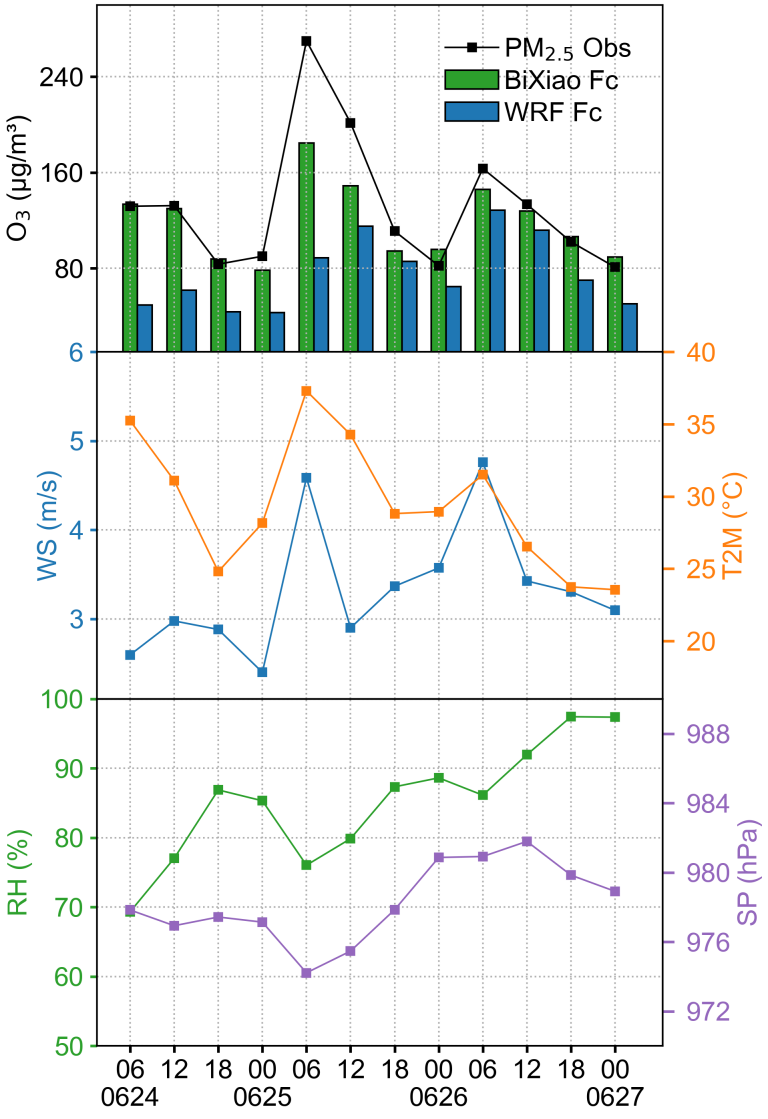

Figure 8. Time series of average pollutant concentrations during the O_3 pollution event and forecast values from BiXiao and WRF-Chem (top), surface wind speed and temperature time series (middle), and surface relative humidity and pressure time series (bottom).

and observations was approximately 21.43 μg m-3, indicating the high accuracy of the BiXiao model in forecasting the ozone pollution event.

Figure 9 shows the average pollution levels across all grids during the entire pollution event, as well as the average forecast errors of BiXiao and WRF-Chem in different grids. BiXiao's average absolute forecast error is consistently lower than that of WRF-Chem across all grids. Moreover, the distribution of BiXiao's forecast errors is more uniform across different grids,

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

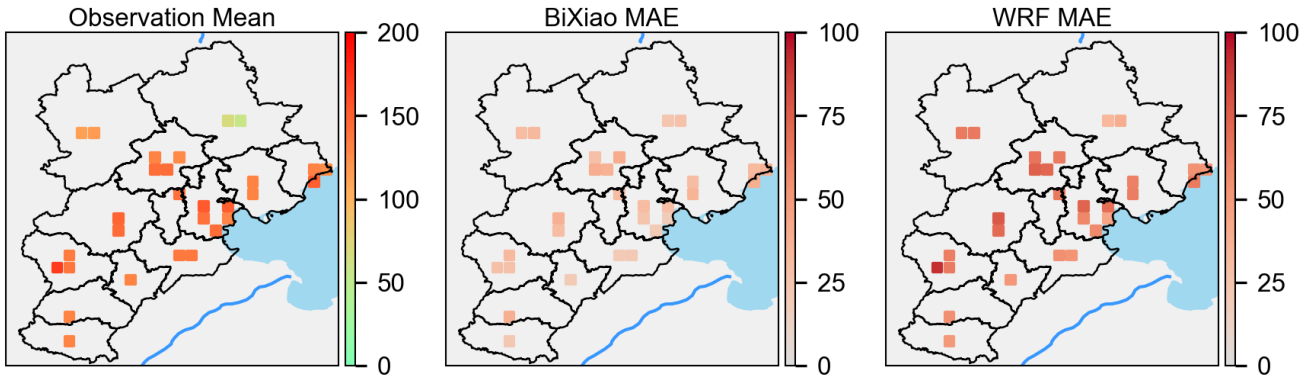

Figure 9. Distribution of average O_3 observed concentration (left), BiXiao forecast average absolute error (middle), and WRF-Chem forecast average absolute error (right) across different grids for the 6-72 hour forecast period starting from 00:00 UTC, June 24, 2022.

whereas WRF-Chem exhibits larger forecast errors in the central region, where O_3 pollution is more severe. Overall, BiXiao outperforms WRF-Chem significantly during this O_3 pollution event.

5.2.2 $PM_{2.5}$ Pollution Case

From October 29 to November 1, 2023, the Beijing-Tianjin-Hebei (Beijing-Tianjin-Hebei) region experienced a severe $PM_{2.5}$ pollution event, influenced by both local emissions (such as motor vehicle exhaust, industrial furnaces, and straw burning) and regional transport. Several cities, including Beijing, Tianjin, Shijiazhuang, and Anyang, issued heavy pollution weather warnings. This study selected this pollution episode for a comparative analysis, with a forecast start time of 00:00 UTC, October 29, 2023, and a total forecast period of 72 hours, to evaluate the forecasting capability of the BiXiao model and compare it with numerical forecast models.

Meteorologically, starting from October 29, the Beijing-Tianjin-Hebei region and surrounding areas were under the control of a saddle-shaped pressure pattern, characterized by low surface pressure and persistent southerly winds near the surface with an average wind speed of only 2 m/s. Due to the static weather pattern, atmospheric diffusion conditions in the region were unfavorable. As shown in Figure 10, this pollution event was also accompanied by prolonged high temperatures and humidity. The highest near-surface temperatures exceeded 23°C, while the relative humidity at night reached 97%, nearly saturated, which favored the accumulation and transformation of particulate pollution. During this pollution event, the $PM_{2.5}$ concentration peaked at 20:00 local time (12:00 UTC) on both October 29 and 31, with the average grid concentrations in the Beijing-Tianjin-Hebei region reaching 151.4 $\mu g/m^3$ and 147.1 $\mu g/m^3$, respectively.

Figure 11 shows the average observed $PM_{2.5}$ concentrations across all grids during the forecast period, as well as the forecasted concentrations from the BiXiao and WRF-Chem models. Comparing the forecast results every 6 hours with observed concentrations, BiXiao's forecasts are closer to the observed values, with the highest correlation coefficient of 0.82 at the 0h and

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

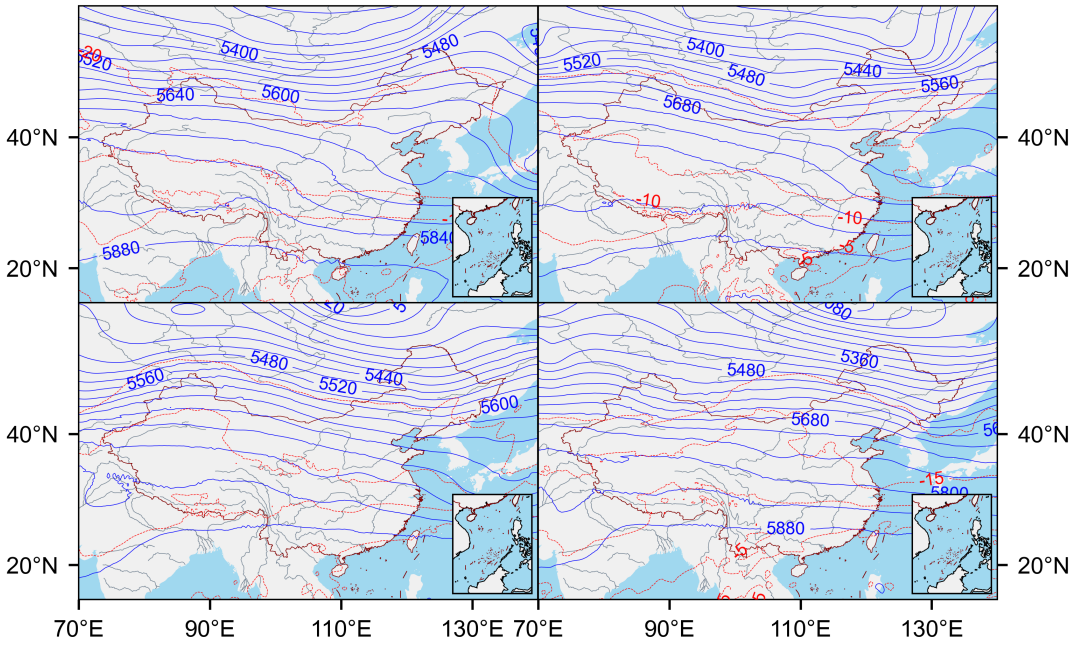

Figure 10. 500 hPa geopotential height and temperature map from 00:00 UTC, October 29 to November 1, 2023.

30h forecast times. The correlation coefficients are lower at the 24h and 72h forecasts, at 0.49 and 0.56, respectively, with the average absolute error around 37.07±16.65 μg m-3. In contrast, WRF-Chem's forecasts exhibit an underestimation, with lower correlation coefficients, peaking at 0.67 at the 30h forecast time.

Figure 12. Average pollutant concentrations in different grids during the forecast period, along with the average absolute forecast errors from BiXiao and WRF-Chem in various grids. BiXiao exhibits higher forecast accuracy for fine particulate concentrations in northern regions of the Beijing-Tianjin-Hebei area compared to the southern regions, with the northern average absolute error for $PM_{2.5}$ around 32 μg m^{-3} (e.g., Zhangjiakou, Beijing, Chengde). Compared to BiXiao, WRF-Chem shows larger forecast errors in most grids, with significantly higher errors in the heavily polluted central and southern regions. Overall, BiXiao performs significantly better than WRF-Chem in forecasting this $PM_{2.5}$ pollution event.

6 Conclusion

To address the limitations of traditional numerical forecasting models, such as low computational efficiency and insufficient resolution in atmospheric environmental element predictions, this study proposes the "BiXiao" AI model for atmospheric environmental forecasting based on a non-uniform grid design. Leveraging the rapid development of AI-driven meteorological large models, the "BiXiao" model innovatively integrates the weather and environmental modules in a "heterogeneous" architecture. This approach enables the direct use of discrete station observation data, overcoming the current dependency of AI large models on gridded data.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

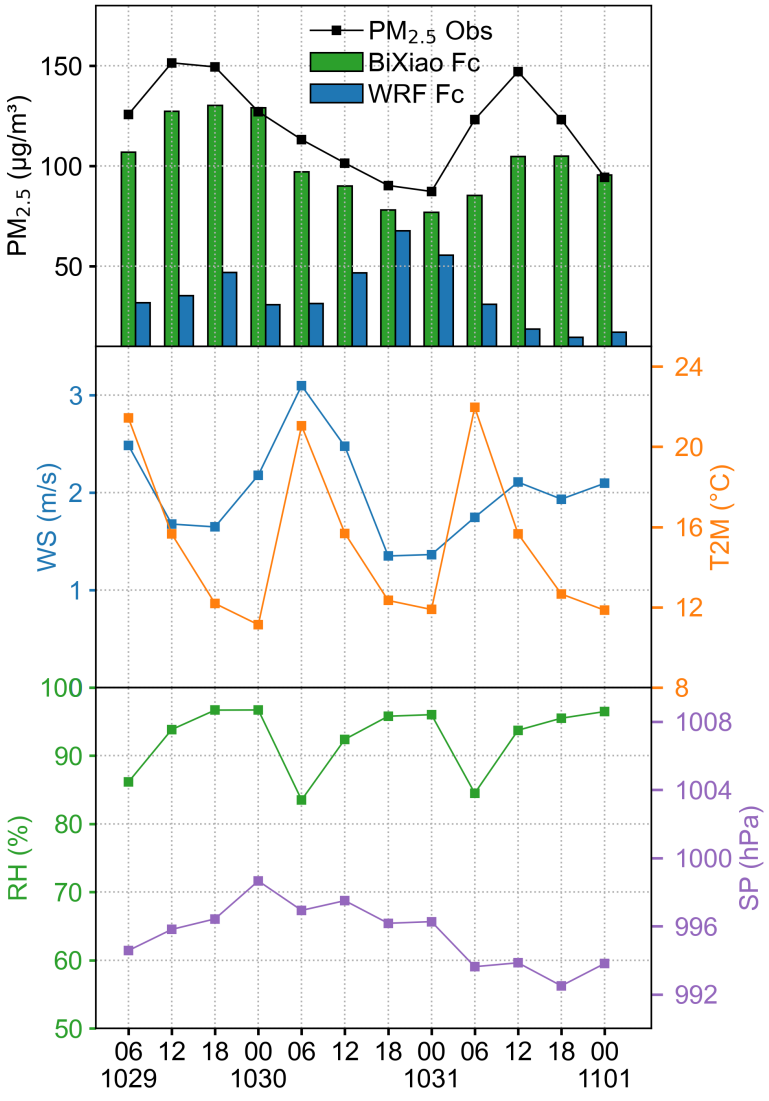

Figure 11. Time series of average pollutant concentrations during the $PM_{2.5}$ pollution event and forecast values from BiXiao and WRF-Chem (top), surface wind speed and temperature time series (middle), surface relative humidity and pressure time series (bottom).

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

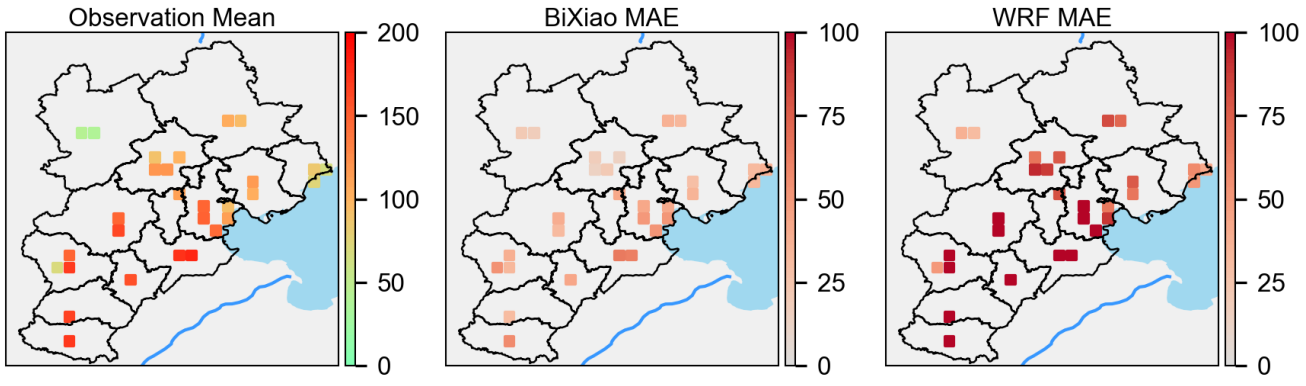

Figure 12. Distribution of average $PM_{2.5}$ observed concentrations (left), BiXiao forecast average absolute errors (middle), and WRF-Chem forecast average absolute errors (right) for different grids over the 6-72 hour forecast period, starting from 00:00 UTC on October 29, 2023.

The meteorological module employs a 3D Swin Transformer as the backbone network to extract three-dimensional atmospheric dynamic features. The environmental module combines the evolution characteristics of the meteorological fields with discrete environmental field data to accurately predict six major pollutants. Additionally, the model can perform inference tasks using a single GPU, significantly improving computational efficiency compared to traditional numerical models.

In validation experiments using the test dataset, the BiXiao model demonstrated strong performance, with the best 6-hour O_3 forecast, achieving a correlation coefficient (PCC) of 0.91. The PCC for $PM_{2.5}$ and PM_{10} forecasts were 0.86 and 0.79, respectively. In comparisons with the mainstream numerical model CAMS for 72-hour forecasts, BiXiao outperformed CAMS in the first 48 hours in terms of both PCC and RMSE for particulate matter forecasts. After 48 hours, BiXiao's performance slightly lagged behind CAMS in small regional areas, though it still outperformed CAMS in most regions. The model reduced short-term RMSE errors for $PM_{2.5}$ and PM_{10} by over 50% compared to CAMS.

In case study simulations, the BiXiao model showed better performance in simulating $PM_{2.5}$ and O_3 pollution events, with significantly lower average errors compared to the WRF-Chem model, demonstrating stronger robustness. This positions the BiXiao model as a new paradigm for fine-scale urban atmospheric environment forecasting.

The development of the BiXiao model marks an initial attempt to apply AI technologies in atmospheric environmental research. Future extensions will focus on the following directions: Firstly, incorporating satellite remote sensing, mobile monitoring data, and other observational data to build a multi-modal dataset that enhances the model's ability to capture local pollution sources. Secondly, expanding the model's research into pollution source tracing capabilities, in addition to forecasting. Finally, addressing the simulation needs for different emission reduction scenarios in atmospheric environmental forecasting, incorporating emission source impact factors into the BiXiao model. The model will also be extended from the Beijing-Tianjin-Hebei region to nationwide applications, with a focus on its suitability in complex terrains (such as plateaus and basins) and typical pollution areas (such as the Yangtze River Delta and Pearl River Delta). With these improvements, the

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

BiXiao model is expected to become an important technological tool in the "pollution reduction and carbon reduction" strategy, providing a Chinese solution for global atmospheric environmental governance.

Acknowledgements

This research was supported by the National Key Basic Research Development Program of China (Grant No. 2024YFC3711905) and the National Natural Science Foundation of China (Grant No. 42477103). The authors used OpenAI's GPT-4/5 language model to assist in translating the initial Chinese draft into English and improving the linguistic clarity of the manuscript. All scientific content, analysis, and conclusions are the sole responsibility of the authors.

Code and data availability

The current version of the BiXiao model is available from https://github.com/jishengxuan1023/BiXiao_model_v1.0 under the GNU Affero General Public License v3.0 only (AGPL-3.0-only). The exact version of the code, model executables, sample input data, observational data for figure reproduction, and the inference and visualization scripts used in the demonstration package have been archived on Zenodo under the DOI: https://doi.org/10.5281/zenodo.17630145 (Ji, 2025). The input fields included in the archive correspond to the heavy $PM_{2.5}$ pollution episode analyzed in the manuscript.

Author contributions

S. Ji designed the research framework, developed the model code, and performed the experiments. All authors discussed the results and contributed to the final version of the paper.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

References

Ayzel, G., Scheffer, T., and Heistermann, M.: RainNet v1. 0: a convolutional neural network for radar-based precipitation nowcasting, Geoscientific Model Development, 13, 2631–2644, 2020.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q.: Accurate medium-range global weather forecasting with 3D neural networks, Nature, 619, 533–538, 2023.

Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Allen, A., Brandstetter, J., Garvan, P., Riechert, M., Weyn, J. A., Dong, H., et al.: A foundation model for the Earth system, Nature, pp. 1–8, 2025.

Düben, P., Modigliani, U., Geer, A., Siemen, S., Pappenberger, F., Bauer, P., Brown, A., Palkovic, M., Raoult, B., Wedi, N., et al.: Machine learning at ECMWF: A roadmap for the next 10 years, ECMWF Technical Memoranda, 878, 24, 2021.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly journal of the royal meteorological society, 146, 1999–2049, 2020.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., et al.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, 2019.

Ji, S.: BiXiao Model, https://doi.org/10.5281/zenodo.17630145, 2025.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., et al.: Learning skillful medium-range global weather forecasting, Science, 382, 1416–1421, 2023.

Lin, H., Feng, X., Fu, T.-M., Tian, H., Ma, Y., Zhang, L., Jacob, D. J., Yantosca, R. M., Sulprizio, M. P., Lundgren, E. W., et al.: WRF-GC (v1. 0): online coupling of WRF (v3. 9.1. 1) and GEOS-Chem (v12. 2.1) for regional atmospheric chemistry modeling–Part 1: Description of the one-way model, Geoscientific Model Development, 13, 3241–3265, 2020.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF international conference on computer vision, pp. 10 012–10 022, 2021.

Meng, J., Han, W., and Yuan, C.: Seasonal and multi-scale difference of the relationship between built-up land landscape pattern and PM2. 5 concentration distribution in Nanjing, Ecological Indicators, 156, 111 079, 2023.

Meng, J., Han, W., Yuan, C., Yuan, L., and Li, W.: The capacity of human interventions to regulate PM2. 5 concentration has substantially improved in China, Environment International, 195, 109 251, 2025.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., et al.: Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators, arXiv preprint arXiv:2202.11214, 2022.

Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T. R., El-Kadi, A., Masters, D., Ewalds, T., Stott, J., Mohamed, S., Battaglia, P., et al.: Gencast: Diffusion-based ensemble forecasting for medium-range weather, arXiv preprint arXiv:2312.15796, 2023.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c.: Convolutional LSTM network: A machine learning approach for precipitation nowcasting, vol. 28, 2015.

Sønderby, C. K., Espeholt, L., Heek, J., Dehghani, M., Oliver, A., Salimans, T., Agrawal, S., Hickey, J., and Kalchbrenner, N.: Metnet: A neural weather model for precipitation forecasting, arXiv preprint arXiv:2003.12140, 2020.

Wang, T. J., Li, M. M., Han, Z. W., Zhang, H., Zhou, C. H., Xie, M., Li, S., Zhuang, B. L., Wu, H., and Qu, Y. W.: Research progress and future prospects of atmospheric environment models in China, China Environ. Sci., 44, 5921–5933, (in Chinese), 2024.

https://doi.org/10.5194/egusphere-2025-5589
Preprint. Discussion started: 10 February 2026
© Author(s) 2026. CC BY 4.0 License.

Wang, Z. F., Pang, C. M., Zhu, J., An, J. L., Han, Z. W., and Liao, H.: IAP Progress in Atmospheric Environment Modeling Research, Chin. J. Atmos. Sci., 32, 987–995, (in Chinese), 2008.

Xiao, X., Kuang, Q., Xiang, S., Hu, J., and Pan, C.: Precipitation forecasting via multi-scale deconstructed convlstm, arXiv preprint arXiv:1912.09425, 2019.

Yang, Z. J., Li, K., Liao, H., and Chen, L.: Analysis of Surface Ozone Pollution and Its Meteorological Causes in China during the Record Summertime Extreme Heat of 2022, Chin. J. Atmos. Sci., 49, 1–12, (in Chinese), 2025.

Zhou, G., Xu, J., Xie, Y., Chang, L., Gao, W., Gu, Y., and Zhou, J.: Numerical air quality forecasting over eastern China: An operational application of WRF-Chem, Atmospheric Environment, 153, 94–108, 2017.