Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-15
DOI: *10.15244/pjoes/213855* ONLINE PUBLICATION DATE: 2026-03-04

Original Research

STAMGCN: A Spatio-Temporal Attention-Based Multi-Graph Convolution Model for Fine-Grained Air Quality Analysis

Huixiang Liu°, Jiacheng Wang, Lingpeng Meng, Wenbai Chen*

College of Automation (College of Artificial Intelligence), Beijing Information Science and Technology University, Beijing 102206, China

Received: 14 August 2025
Accepted: 02 November 2025

Abstract

Air pollution poses significant global health challenges, underscoring the need for highly detailed spatial predictions to inform effective environmental policy and safeguard public health. This study introduces a novel Spatio-Temporal Attention-based Multi-Graph Convolutional Network (STAMGCN) designed for fine-grained air quality prediction. The framework constructs spatial, atmospheric pollution-pattern, and meteorological-pattern graphs to represent complex, non-Euclidean regional relationships. Through graph convolutional networks, the model aggregates contextual information from adjacent nodes, followed by a fine-grained attention mechanism that emphasizes interactions between the target and nearby monitoring stations. By leveraging gated recurrent units with temporal attention, STAMGCN effectively captures evolving air quality changes. Experiments conducted on the Beijing dataset demonstrate that the model improves prediction accuracy by 10.18% for immediate forecasts and 15.56% for 6-hour forecasts compared to baseline spatio-temporal models. These results highlight the model's potential to support urban air-quality management and provide a robust scientific foundation for pollution-control strategies.

Keywords: fine-grained air quality analysis, graph convolutional networks, gated recurrent units, spatiotemporal attention mechanism

Introduction

In recent decades, the rapid progression of urbanization has led to increasingly severe challenges in urban air quality, profoundly affecting both human health and environmental sustainability. The World Health Organization (WHO) reported that, in 2016, the global average life expectancy decreased by one year, while the average life expectancy in polluted countries in Asia and Africa decreased by 1.2 to 1.9 years [1]. Annually, deteriorating air quality is linked to approximately seven million deaths worldwide [2]. Moreover, more than 90% of the global population breathes air containing pollutant concentrations that exceed WHO-recommended limits. This situation is particularly acute in developing countries, especially in China, which has a population of 1.4 billion people [3].

*e-mail: chenwb@bistu.edu.cn
Tel.: +86-10-8018-7488
°ORCID iD: 0000-0002-4101-6768

Consequently, air pollution has been widely studied. The assessment of air quality involves a comprehensive evaluation of the concentrations of various pollutants in the environment to determine their potential impact on human health and ecosystems. Among these assessments, the Air Quality Index (AQI) serves as a crucial indicator for the evaluation of air quality. A higher AQI value indicates poorer air quality, which increases the likelihood of related diseases among the population [4]. To obtain accurate AQI measurements, several air quality monitoring stations have been established worldwide. These stations employ automatic monitoring equipment to detect pollutants in the air in real-time, process the collected data, and release the results to the public. However, due to factors such as high construction costs, complex maintenance requirements, and significant labor demands, it is not feasible to establish enough monitoring stations for urban air pollution monitoring. Furthermore, the coverage of monitoring stations in each area is often relatively limited, making it difficult to achieve comprehensive monitoring across the region. Therefore, predicting the AQI in these unmonitored target areas remains a challenging task. Since air quality can vary considerably over short distances [5], it is not feasible to rely solely on data from nearby monitoring stations for inference. This creates a critical challenge: how to predict the AQI in unmonitored target areas accurately. Developing specialized models that leverage data from existing stations to perform such fine-grained predictions is essential for tracking air quality trends and supporting timely pollution control measures.

To collect air pollution information from different urban areas, including production, commercial, and entertainment zones, it is often necessary to develop specialized models that use data gathered from existing monitoring stations to predict the current or future AQI of target regions. This approach enables more accurate analysis of air quality conditions. Using this approach, air quality trends can be effectively tracked, and the prompt implementation of appropriate pollution control measures, such as traffic restrictions, wearing masks outdoors, and adjusting indoor ventilation, can be achieved. These preventive measures are crucial for mitigating the harmful impacts of environmental pollution on human health.

In fact, the level of air pollution at any specific location is often strongly influenced by its historical data as well as the transport and dispersion of pollutants from surrounding areas. This relationship suggests that air quality is affected by interacting temporal and spatial dynamics originating from other locations, reflecting the spatio-temporal dependency characteristics of AQI distribution [6]. Although many existing methods have achieved significant progress in air quality forecasting, they continue to encounter several critical challenges in the "fine-grained prediction" of unmonitored locations. These limitations mainly involve insufficient attention to local spatial correlations between target and neighboring sites, as well as a restricted capacity to infer conditions in areas without direct monitoring. Such challenges have prompted the development of more complex methods to effectively integrate multiple graph structures and fine-grained attention mechanisms. For fine-grained prediction, it is essential not only to capture the overall spatial landscape but also to precisely quantify the individual influence of surrounding stations on a target location. Therefore, to achieve accurate prediction of both current and future AQI levels in target areas, this paper proposes the Spatio-Temporal Attention Multi-Graph Convolutional Network (STAMGCN) technique.

Therefore, the main contributions of this paper are summarized as follows:

(1) A Multi-Graph Convolutional Network (MGCN) is introduced for spatial dependency modeling to integrate atmospheric pollution and meteorological information from monitoring nodes. This enables a preliminary understanding of the overall spatial correlations among site-level data.

(2) A fine-grained attention mechanism is incorporated to help the model more effectively focus on localized spatial features of the target monitoring nodes and their nearby stations.

(3) Temporal dependencies are captured by combining Gated Recurrent Units (GRUs) with a temporal attention layer. In addition, spatio-temporal feature representations from both the target and nearby sites are used to predict the future AQI values of the target locations.

(4) Extensive experiments are conducted on the collected Beijing dataset, and the results indicate that STAMGCN outperforms all baseline methods across multiple evaluation metrics, demonstrating its superior efficiency and accuracy in fine-grained air quality analysis.

In summary, this paper leverages the strengths of multi-graph neural networks in capturing spatial features for non-topological prediction tasks, while introducing a novel fine-grained attention mechanism to further refine the model's learning strategy. As a result, the proposed approach effectively performs fine-grained air quality estimation and employs temporal networks combined with historical observations to predict future air quality.

This paper is organized as follows: in next Section, the related works are introduced, whereas the structure and principles of the STAMGCN techniques are proposed in "Materials and Methods". "Results and Discussion" Section presents the experimental setup along with the analysis of the results, as well as the study area and adopted dataset. Finally, the research conclusions of this study are presented along with some future work.

Related Works

Existing fine-grained air quality prediction methods can generally be divided into numerical methods

and data-driven approaches. Numerical approaches rely on atmospheric physics and chemical mechanisms, combined with mathematical models, to simulate and infer the dynamic processes governing urban air quality.

With continuous progress in atmospheric environmental research, researchers have increasingly combined empirical assumptions with observational datasets to describe the diffusion and dispersion of pollutants in the atmosphere under real-world conditions. This integration has enabled the development of physics-based models capable of predicting pollution concentrations at specific locations and time intervals. Guided by studies on the turbulent behavior of the atmospheric boundary layer, numerous indoor numerical simulations and field experiments have been performed, allowing more complex meteorological components to be integrated into physical models. In addition, Chemical Transport Models (CTMs) employ numerical methods to solve the mathematical equations governing these chemical and physical reactions. Compared to models used in traditional monitoring networks, numerical models can provide more comprehensive and widespread air quality information.

For instance, Hood et al. [7] proposed a coupled regional (EMEP4UK) and urban (ADMS-Urban) model system that integrates adjusted emission factors and chemical dispersion models to simulate and evaluate the air quality in London. Moreover, Kim et al. [8] introduced a multi-scale urban air pollution model that combines a Chemical Transport Model (CTM) with a street network model to simulate and evaluate NO_x and O_3 concentrations in the suburbs of Paris. This study achieved a spatial resolution of about 1 km. Moreover, Computational Fluid Dynamics (CFD) was used to infer the aerodynamic physical processes, especially concerning air flow, turbulence, and pollutant dispersion at various scales. It is commonly employed in high-resolution urban environmental studies. For instance, Santiago et al. [9] proposed a CFD-based urban street model that uses the Reynolds-Averaged Navier-Stokes (RANS) method to infer pollutants' spatial distribution in traffic-heavy areas. Other models have been developed, such as the Gaussian plume models [10], Lagrangian particle models [11], and convection-diffusion models used for fine-grained air quality analysis [12]. These models consist of a series of equations requiring extensive and diverse data to determine their parameters, and leading to increased computational load. In addition, parameter settings may not be applicable to all atmospheric environments as they highly depend on detailed pollutant emission data, therefore introducing substantial uncertainty in the predictions.

Moreover, data-driven air quality analysis methods are categorized into air quality prediction, fine-grained air quality estimation, and their combination [13]. First, air quality prediction involves developing models that use historical data to forecast air pollution levels in a specific area for a future time point or period. Several data-driven air quality modeling methods have been proposed. Early studies typically employed statistical methods [14, 15], which did not consider the complexity of atmospheric processes. It has been shown that statistical models, such as Autoregressive Integrated Moving Average (ARIMA) [16], SVM [17], as well as their variants, are effective for short-term predictions. However, their linear assumptions limit their prediction accuracy when working with non-linear air pollution data.

Machine Learning (ML) methods have been applied to predict air quality by learning model parameters from datasets, without the need to directly understand the underlying changing mechanisms. This approach does not rely on physical or chemical analysis processes. It rather uses the patterns of pollution changes hidden within the data. Some ML methods, such as Support Vector Regression (SVR) [18], Random Forest (RF) [19], and Multi-Layer Perceptron (MLP) [20], have been adopted for air quality prediction. These approaches can effectively handle non-linear dependencies in datasets. However, they often overlook deeper non-linear relationships and long-term sequential patterns within this data.

Deep learning represents a cutting-edge ML technology with significant potential in the field of air quality prediction. Recurrent Neural Networks (RNNs) and their variants, including Long Short-Term Memory (LSTM) [21] and GRU [22], have recently been applied to extract temporal dependencies from air quality data. These methods have consistently outperformed traditional models, such as ARIMA and SVR, in terms of prediction accuracy [23]. However, due to the dynamic behavior of urban air pollutants and their complex spatial interactions across regions, researchers have increasingly employed Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs) to capture spatial features, leading to the emergence of novel spatio-temporal model architectures. However, because air pollutant monitoring stations are distributed in a non-Euclidean manner, CNN-based methods, which have been highly effective in image analysis, are less capable of modeling complex non-Euclidean spatio-temporal dependencies. For instance, it has been shown that a CNN-LSTM hybrid model does not outperform a standalone LSTM in air quality prediction tasks [24]. In contrast, GNNs possess a strong capability to model relational structures and graph topologies, making Spatio-Temporal Graph Neural Networks (STGNNs) a major focus of current research [25, 26]. Chen et al. [27] proposed a Spatiotemporal Multi-Graph Convolutional Network (STMGCN), which constructs three types of graphs (spatial, air pollution pattern, and meteorological pattern) to effectively capture multidimensional correlations within the monitoring station network. When combined with a GRU network, the model successfully captures temporal evolution and demonstrates the effectiveness of the multi-

graph approach for city-level air quality prediction. Nevertheless, STMGCN remains a supervised learning model tailored to a fixed set of monitoring stations. Its architecture is primarily designed to forecast future values for stations within the existing network, and therefore cannot be directly extended to infer air quality in unmonitored or external regions. This constraint significantly limits its suitability for fine-grained air quality estimation tasks.

Furthermore, attention mechanisms, which enable models to focus selectively on salient or informative features, have been extensively explored in recent research. For instance, Liang et al. [28] introduced a Transformer-based architecture, AirFormer, which employs self-attention mechanisms to learn spatio-temporal representations and effectively capture data uncertainty. This approach achieved high-accuracy air quality predictions across large-scale datasets. However, Transformer-based frameworks generally demand substantial training data and computational resources. More importantly, their global spatial aggregation strategies, designed to improve overall prediction efficiency, may inadvertently reduce the model's ability to capture fine-grained, high-fidelity local environmental information around specific target regions, which is precisely the core requirement for fine-resolution interpolation prediction tasks.

Fine-grained estimation involves utilizing existing monitoring data to infer pollutant concentrations in target areas lacking direct observations, for either current or future time points. The most widely adopted estimation methods estimate and predict fine-grained air quality in such target areas using spatial interpolation techniques, such as spatial averaging, nearest neighbor, and Inverse Distance Weighting (IDW) algorithms. Other studies have explored deep learning-based interpolation models to enhance estimation accuracy. For instance, Xu et al. [29] proposed a two-stage inference framework that integrates remote sensing and urban auxiliary data to estimate fine-grained air quality through artificial neural networks combined with tensor decomposition. Moreover, research has also examined the direct prediction of future air quality indices in target areas. For instance, Zhao et al. [30] proposed a spatio-temporal fusion graph convolutional network based on attention mechanisms. This approach models the historical spatio-temporal patterns of air pollution data in target areas for predicting the future AQI at those locations.

In recent years, several studies have focused on simultaneously performing air quality prediction and fine-grained estimation. For instance, Qi et al. [31] proposed the Deep Air Learning (DAL) method that integrates feature selection and semi-supervised learning within deep network layers to address air quality interpolation, prediction, and feature analysis. Similarly, Hsieh et al. [13] developed a spatio-temporal attention-based model capable of estimating the current AQI for target areas and forecasting values for the next 24 hours. In addition, Han et al. [32] introduced the Self-Supervised Hierarchical Graph Neural Network (SSH-GNN), which employs a multi-view fusion module to combine spatio-temporal features and environmental information. SSH-GNN uses hierarchical recursive GNNs to achieve city-wide air quality prediction and demonstrates strong performance under semi-supervised scenarios. Nevertheless, the framework also exhibits certain limitations, including the risk of error accumulation caused by multi-stage processing and the focus on long-range dependence.

While recent advances such as AirFormer and SSH-GNN have achieved promising results in air quality prediction, these models are designed for different scenarios. In fact, AirFormer leverages Transformer architecture for large-scale datasets with hundreds of monitoring stations, whereas SSH-GNN emphasizes city-wide prediction in situations where ground-truth data for target areas are unavailable for evaluation. In contrast, our STAMGCN approach is specifically designed for fine-grained prediction in cities with limited monitoring coverage (typically 10-20 stations). The key challenge is to accurately estimate air quality at specific unmonitored locations based on available monitored data. Therefore, the proposed multi-graph structure with learnable fusion weights and a fine-grained attention mechanism is tailored to this specific scenario.

Specifically, our proposed STAMGCN employs a supervised fine-grained attention mechanism specifically designed to directly capture the relationship between a target area's features and the data from surrounding monitored stations, providing a more direct approach to the fine-grained estimation problem.

Materials and Methods

Problem Statement

The fine-grained air quality prediction task aims to model historical air pollution data collected from monitoring stations over a time period T to predict current and future pollution levels in unmonitored locations. Two sets of site-level graphs are defined: $S = s_1, s_2, ..., s_N$ and $S = s_1, s_2, ..., s_M$, where N and M represent the total numbers of monitoring and non-monitoring station areas, respectively. Let $X_t^n \in R^{(N\times V)}$ denote the measurements of V variables, including air pollutants (e.g., AQI, CO, and O_3), meteorological elements, geographical coordinates (latitude and longitude), and source types for the monitoring stations at time t. Similarly, let $X_t^m \in R^{(M\times B)}$ represent the measurements of B meteorological and spatial attributes for non-monitoring station areas at time t. The goal of this study is to use historical observation $X^n = X^n_{(t-T+1)}, ..., X_t^n, X^n \in R^{(N\times T\times V)}$ and $X^m = X^m_{(t-T+1)}, ..., X_t^m, X^m \in R^{(M\times T\times B)}$ collected over T consecutive hours across all $N+M$ sites to predict the AQI values $Y = X_{(t+1)}, ..., X_{t+L}$, $Y \in R^{(M\times L)}$ for M target sites over the next L hours.

Note that the AQI index, which is considered a comprehensive measure of the urban air quality, is mainly adopted for the prediction task in this work.

Overall Framework

The overall architecture of the proposed STAMGCN framework is depicted in Fig. 1. The framework is composed of four blocks, described as follows:

BLOCK A (Input Data): It consists of input data, such as air quality, geographical location, meteorological, and source type data.

BLOCK B (Input Data Processing): In this block, the raw data undergoes a series of preprocessing operations and is divided into monitoring station and target area data.

BLOCK C (Spatial Information Modeling): This block represents the core component of the proposed framework. Monitoring station data is fed into the MGCN to extract overall spatial features, while target area data is processed through a fully connected layer for feature dimensionality reduction. These features are then jointly inserted into the spatial fine-grained attention mechanism to perform a fine-grained estimation of air quality.

BLOCK D (Temporal Information Modeling): In this block, temporal variation patterns are learned from the historical features derived from spatial information modeling to predict future air quality in the target area. This block consists of a GRU, temporal attention mechanisms, and fully connected layers.

Modelling of Spatial Dependency

Constructing graphs is a crucial step in graph convolution, as it enables the model to capture spatial interconnections among different locations. In general, monitoring stations that are geographically closer and share similar pollution and meteorological patterns tend to exhibit comparable pollution conditions. Consequently, the edge weights between certain node pairs may be stronger than others within the graph. To represent these spatial dependencies, the spatial graph is denoted as $G_n = (S, E_n, A_n)$, where S indicates the set of sites consisting of N nodes and E is the set of edges within the graph. The adjacency matrix A, which corresponds to the graph, represents the geographical spatial relationships between the underlying sites. Moreover, A_n is constructed based on the Euclidean distance between the sites, where $\alpha \in A_n$ denotes the connectivity between sites s_i and s_j. To select nearby stations for a given location s_i, the distance d_{ij} between locations s_i and s_j must be smaller than a predefined threshold $d_{k\text{-}lg}$. The threshold $d_{k\text{-}lg}$ is defined as the maximum distance among the k nearest neighboring stations to site s_i.

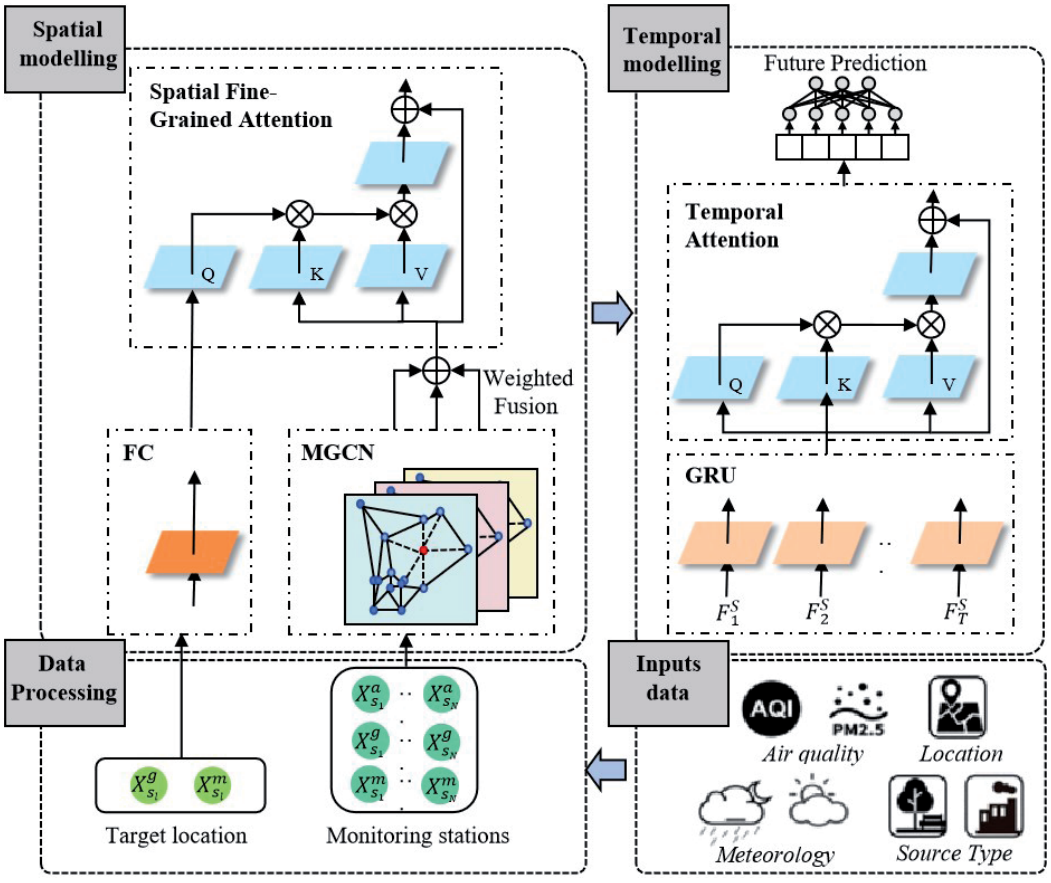

Fig. 1. Overall framework of the proposed STAMGCN for the prediction of air pollution.

$$a_{ij} = \begin{cases} 1/d_{ij}, & if \quad d_{ij} < d_{\mathrm{k-lg}} \\ 0, & else \end{cases} \tag{1}$$

Typically, adjacent regions exhibit strong correlations in air pollution levels, often reflected in similar trends within the observed data. However, even non-adjacent regions that share similar geographic environments and economic development conditions may also exhibit similar pollution behaviors. To capture these relationships, this paper constructs an air pollution pattern graph that models the similarity of pollution characteristics among monitoring stations. The weight $a_{ij} \in A_p$ represents the degree of similarity in pollutant patterns between two stations. Based on the historical monitoring data, the similarity between sites is calculated using correlation methods, and the resulting correlation indices are employed to generate the air pollution pattern graph.

In practical applications, constructing the pollution pattern requires extracting $N \times V$ time series from the air pollution characteristic data of each monitoring site. Each sequence is arranged hourly and exhibits a length denoted by L. To comprehensively capture the pollution patterns of each site, data from one representative week per month can be selected from the historical air pollution characteristic data and then concatenated to form continuous sequences. Furthermore, the Pearson Correlation Coefficient (PCC) is employed to measure the strength and direction of the linear relationship between two random variables. It quantifies how changes in one variable are associated with changes in another. Consequently, the PCC values are calculated to assess the similarity of air pollution characteristics among reference sites. Finally, an average correlation index of different air pollution characteristics is computed to generate a pollution pattern map.

For instance, in the first step, a pollution sequence is constructed for each monitoring station using historical pollutant data, where each sequence represents one-dimensional data. These sequences are then aggregated to form an $N \times L$ dimensional dataset. In the second step, the PCCs are computed between the pollution sequence of each station and those of all other stations, resulting in a similarity matrix for each pollutant. This matrix is symmetric with a principal diagonal of ones, where each off-diagonal element quantifies the degree of similarity in pollution patterns between station pairs. In the third step, similarity matrices for the remaining pollutants are computed in the same manner, yielding multiple matrices corresponding to the number of pollutant types. These matrices are then averaged element-wise to produce an $N \times N$ air pollution pattern matrix. The detailed workflow of air pollution pattern analysis is illustrated in Fig. 2.

Existing studies have shown that meteorological conditions play a crucial role in influencing the dispersion and accumulation of air pollutants. When significant temperature gradients occur within an area, they can substantially affect the direction and speed of pollutant transport, thereby impacting the accuracy of the performed predictions. Conversely, extreme weather conditions, such as sandstorms and heavy rainfall, can dramatically alter the current air pollution status. Therefore, areas exhibiting similar meteorological patterns often exhibit comparable pollution conditions. To capture this relationship, this paper constructs

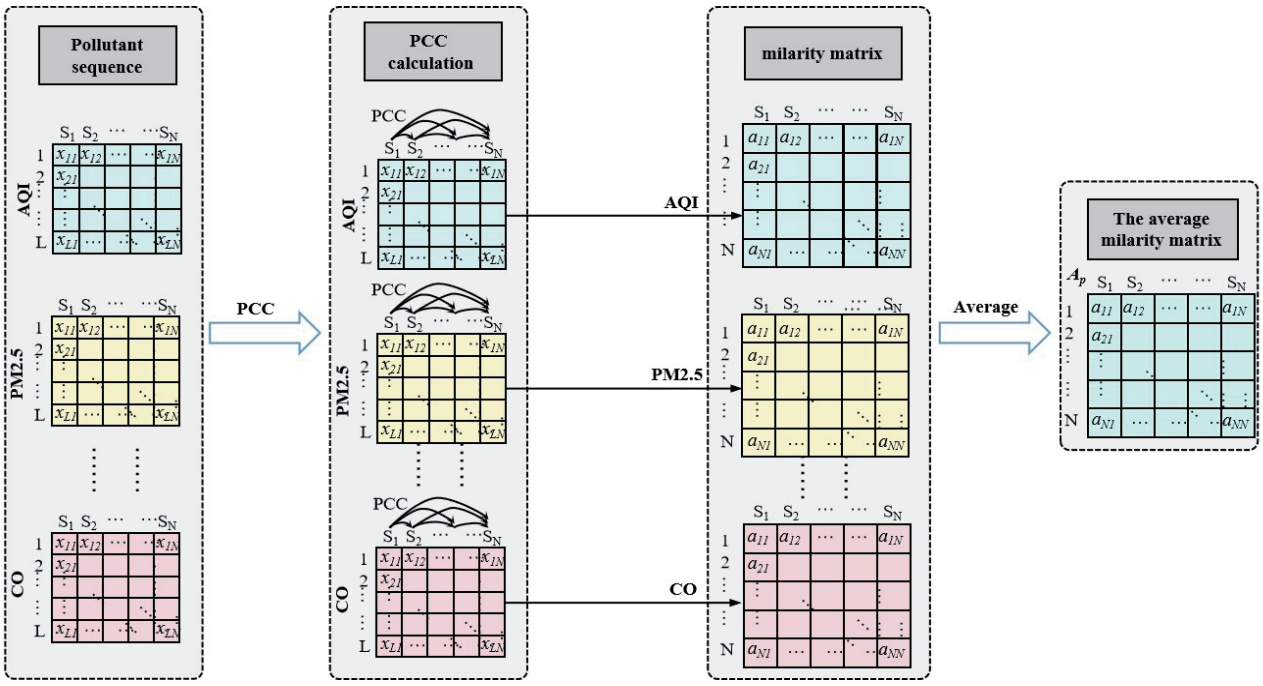

Fig. 2. Construction of the pattern graph of the air pollutant.

meteorological pattern graphs following the same process used for generating pollution pattern graphs. The meteorological pattern graph is denoted as $G_m = (S, E_m, A_m)$. The weight $a_{ij} \in A_m$ represents the similarity in meteorological patterns between the two sites s_i and s_j.

Three types of graph adjacency matrices, A_n, A_p, and A_m, are constructed to represent spatial, pollution, and meteorological patterns, respectively. To jointly integrate the spatial feature information derived from these three graph structures, each processed through graph convolution, the proposed framework employs graph convolutional layers with an identical structure for processing. The resulting feature matrices, denoted as H_n, H_p, and H_m, capture the spatial representations corresponding to the three graphs. Learnable weighting coefficients are then used to perform weighted fusion of spatial features. Subsequently, learnable weighting coefficients λ'_n, λ'_p, and λ'_m are employed to perform weighted fusion of these spatial features.

The layer-by-layer propagation rules for the multi-graph convolutional layers are expressed as follows:

$$H^{(l+1)} = \sigma(\tilde{D}^{(-1/2)}\hat{A}\tilde{D}^{(-1/2)}H^{(l)}W^{(l)}) \quad (2)$$

$$\lambda'_n, \lambda'_p, \lambda'_m = softmax(\lambda_n, \lambda_p, \lambda_m) \quad (3)$$

$$H_{fusion} = \lambda'_n \odot H_n + \lambda'_p \odot H_p + \lambda'_m \odot H_m \quad (4)$$

where $H^{(l)} \in R^{(C\times N)}$ denotes the output of the l-th layer, and C represents the feature embedding dimension of the GCN. Moreover, H_{fusion} is the output of graph convolution computations performed on the three distinct graph structures, integrating the results from all three graph structures.

It is worth noting that X_t is the input to the GCN. The equation $\hat{A} = A + I$ represents the weighted adjacency matrix of the graph, including self-connections, where I represents the identity matrix, A denotes the spatial graph matrix A_n that has been constructed, $\tilde{D}$ is the diagonal degree matrix of $\hat{A}$, $W^{(l)}$ indicates the trainable matrix of weights for the filter parameters within the graph convolutional layer, and $\sigma(x)$ represents the activation function.

Due to the necessity of estimating and predicting air quality at target sites, this study performs learning of the spatial dependencies between the target and nearby sites after employing multi-graph convolutional layers for spatial feature extraction from monitored site data. Consequently, a spatial fine-grained attention mechanism is used to address this need.

To this end, all features are first normalized. The key and value vectors in the attention layer are then derived from the multi-graph convolutional output layers. Afterwards, a linear layer generates query vectors from target site data, aiming to maintain a consistent dimensionality with the key and value vectors. The spatial attention layer is then used to determine the spatial relationships between the query site and its corresponding region. Consequently, the refined spatial features represent the final output $P_h \in R^{(C/N_h\times N\times T)}$. These outputs from the multiple attention heads (N_h) are first concatenated along the feature dimension and then passed through a final linear projection layer to produce the unified output representation. This process is illustrated in Fig. 3, representing the spatial fine-grained attention workflow, where N_{trg} denotes the number of target sites and M_{ner} is the number of neighboring sites around the target sites, which is equivalent to the number of neighboring sites (as defined in the spatial graph construction mentioned above).

The mathematical formulation for spatial fine-grained attention is defined as follows:

$$\begin{cases} P_h = Softmax(\alpha Q_h K_h^T + B_h)V_h \\ M_{head} = Concat(P_1, P_2, \ ..., \ P_h)W^o \end{cases} \quad (5)$$

where $B_h \in R^{(N\times M)}$ represents a learnable relative position encoding [20] used to incorporate positional information into the attention mechanism. The matrices $Q_h = HW_Q$, $K_h = HW_k$, and $V_h = HW_v$ denote the query, key, and value projections, respectively, whereas W_q, W_k, and $W_v \in R^{(C\times C/N_h)}$ are learnable linear transformations, where N_h is the number of attention heads, and α serves as a scaling factor. Specifically, the relative position encoding B_h is implemented as a learnable embedding that captures the spatial displacement between the query site and key sites, allowing the model to learn translation-invariant and spatially generalizable attention patterns. B_h is initialized as a learnable parameter matrix, where each element b_{ij} represents the relative positional bias between

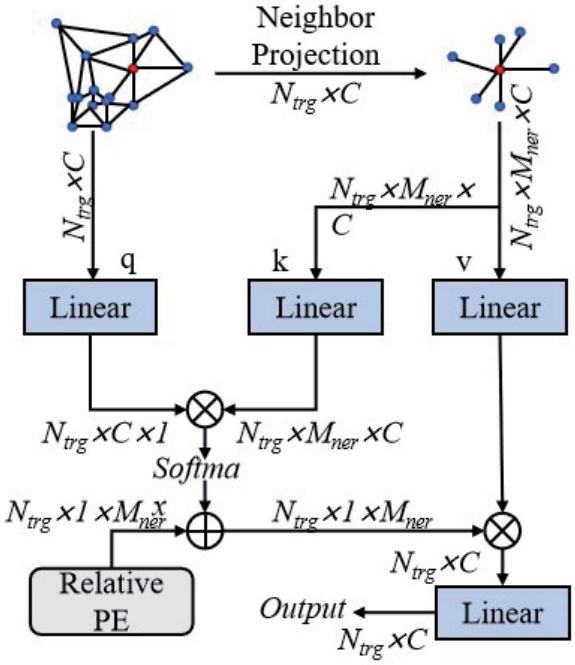

Fig. 3. Process of the fine-grained attention mechanism.

target site i and monitoring site j. During training, this matrix is jointly optimized with other model parameters through backpropagation, enabling it to encode complex, non-linear spatial relationships. This learnable formulation provides greater flexibility compared to fixed encoding schemes, such as sinusoidal position encoding used in transformers, as it allows the model to adaptively discover and emphasize the most relevant spatial dependencies for accurate air quality prediction. Through optimization, B_h effectively captures non-linear and context-dependent spatial patterns that may not be fully represented by simple distance metrics alone. Finally, the multiple outputs are concatenated into a matrix $M_{\text{head}} \in R^{(C\times N\times T)}$ through a concatenation operation, where W^o denotes the linear transformation matrix.

Modelling of Temporal Dependency

Forecasting air pollution is fundamentally a time series prediction problem. The GRU effectively addresses the gradient problems associated with RNNs while also enabling faster training compared to the LSTM network. In a standard GRU cell, the presence of a reset gate r_t and an update gate z_t allows the model to adaptively combine recent and historical inputs to forecast the output. The mathematical modelling of this process is defined as follows:

$$z_t = \sigma(W_z X' + U_z h_{t-1}) \tag{6}$$

$$r_t = \sigma(W_r X_t' + U_r h_{t-1}) \tag{7}$$

$$\tilde{h}_t = tanh\big(W_h X_t' + U_h(h_{t-1} \odot r_t)\big) \tag{8}$$

$$\mathrm{h}_t = (1 - z_t) \odot \tilde{h}_t + z_t \odot h_{t-1} \tag{9}$$

$$y_t = \sigma(W_o h_t) \tag{10}$$

After extracting temporal features from a historical T-hour window using GRU, a temporal attention layer is introduced to further enhance the predictive ability of the model and more accurately capture key temporal dependencies within the time series. The temporal multi-head self-attention mechanism enables the model to adaptively focus on different time steps, thereby significantly improving its ability to capture evolving trends in air quality changes. Following relative position encoding, the feature series $R = R_{(t-T+1)}, ..., R_t$, $R \in R^{(U\times N\times T)}$ are input into the temporal attention layer, where U denotes the number of hidden units in the last GRU layer. The integrated features are then fed into an output layer composed of fully connected layers, which produce the air quality forecasts for the current time step and the next 6 hours.

Results and Discussion

Study Area and Dataset

As Beijing serves as the economic, political, and cultural center of China, its high population density and intensive industrial activities have led to persistent and complex air pollution issues, attracting extensive research interest in urban air quality management. This study selects Beijing as the research area for fine-grained air quality analysis, aiming not only to reveal the characteristics of air pollution in large metropolitan environments but also to provide a scientific foundation for policymakers in developing effective pollution control strategies. Furthermore, Beijing's well-established air quality monitoring network supplies comprehensive and reliable dataset support for this research.

The dataset used in this study for model training and evaluation comprises air quality data, meteorological data, geographic coordinates, and regional classifications. The air quality data were obtained from the Beijing Municipal Environmental Monitoring Center, consisting of hourly historical measurements collected from air quality monitoring stations across 15 administrative regions of Beijing between January 1, 2017, and December 31, 2019. Each region corresponds approximately to an administrative district, with one representative monitoring station selected per region. The locations of these stations are illustrated in Fig. 4. Table 1 lists the detailed geographic coordinates and regional types of the selected stations, including latitude, longitude, and the overall environmental classification of each region. Each monitoring station continuously records the AQI and the concentrations of 6 air pollutants on an hourly basis. Additionally, since meteorological conditions play a critical role in influencing air quality, meteorological data for the same time period were collected from weather stations of the China Meteorological Administration. This data includes five types of meteorological data: temperature, wind direction, wind speed, precipitation, and humidity. It is worth noting that, for each air quality monitoring station, corresponding meteorological parameters were obtained from its nearest weather station to ensure spatial consistency. Finally, the detailed descriptions of the data variables used in this study are presented in Table 2.

Due to various natural factors (e.g., disasters) and human activities, air quality datasets often contain missing values. Simply removing these missing records can lead to temporal discontinuities and a reduction in prediction accuracy. Therefore, a two-step imputation strategy was adopted for data preprocessing. First, when a value was missing at a specific station for a given time step, it was imputed using the average of the same variable from all other stations at that particular time. Second, if the data for a certain variable were missing across all stations at a specific time, the gap was filled by calculating the average of the values

from the immediately preceding and following time steps for each station individually. To eliminate the effects of dimensional inconsistencies among variables and to accelerate model convergence, data were transformed to the range [0, 1] using the following normalization Equation:

$$X' = \frac{X - min(X)}{max(X) - min(X)} \tag{11}$$

This temporal interpolation strategy preserves the temporal evolution patterns of the data while preventing the introduction of artificial discontinuities. The combined use of spatial and temporal imputation ensures data completeness and maintains both the spatial variability and temporal continuity patterns in the dataset.

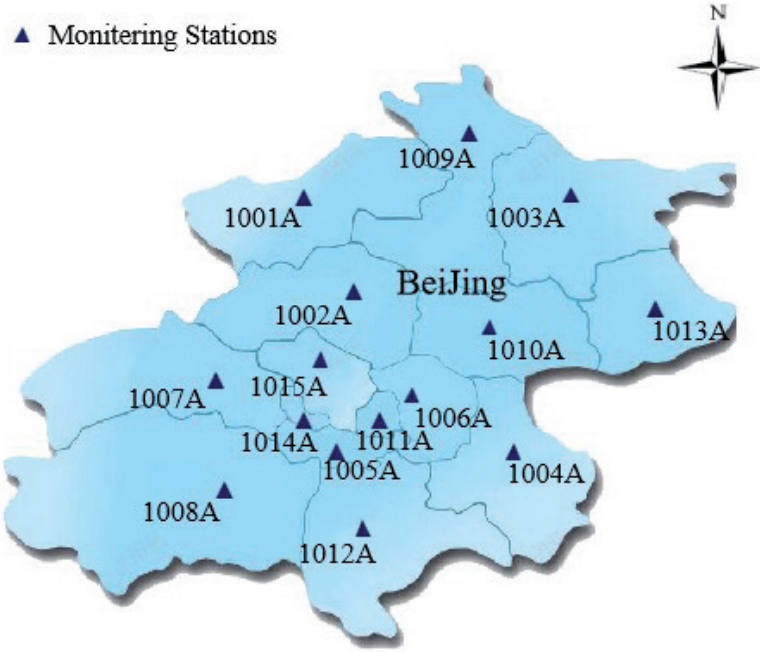

Fig. 4. Distribution map of the monitoring stations.

Table 1. Air quality monitoring sites in Beijing.

Monitoring station	Latitude	Longitude	District	Type
1001A	40.453	115.972	Yanqing	Tourism and Mountainous Area
1002A	40.217	116.23	Changping	Agriculture and Technology
1003A	40.37	116.832	Miyun	Tourism
1004A	39.886	116.663	Tongzhou	Commerce
1005A	39.863	116.279	Fengtai	Light Industry and Commerce
1006A	39.929	116.417	Dongsi	Commerce and Culture
1007A	39.937	116.106	Mentougou	Tourism and Mountainous Area
1008A	39.742	116.136	Fangshan	Light Industry and Tourism
1009A	40.328	116.628	Huairou	Tourism
1010A	40.127	116.655	Shunyi	Light Industry
1011A	39.899	116.395	Xicheng	Commerce and Culture
1012A	39.718	116.404	Daxing	Light Industry and Agriculture
1013A	40.143	117.1	Pinggu	Tourism and Agriculture
1014A	39.914	116.184	Shijing	Heavy Industry and Commerce
1015A	39.987	116.287	Haiding	Technology

Table 2. Description of the air pollutants and meteorological factors.

Factor	Unit of Measurement	Min	Max
AQI	μg/m^3	0.1	510.0
$PM_{2.5}$	μg/m^3	3	999
PM_{10}	μg/m^3	3	999
NO_2	μg/m^3	2	300
SO_2	μg/m^3	2	307
CO	μg/m^3	0.1	15.0
O_3	μg/m^3	1	504
Temperature	℃	-19.9	39.3
Precipitation	mm	0.0	52.1
Wind speed	m/s	0.0	11.2
Wind direction	°	0.0	337.5

Baseline Models

To evaluate the performance of the proposed model, a series of comparative experiments was conducted against several benchmark models, described as follows:

(1) K-Nearest Neighbors *(KNN):* This method predicts by selecting the K nearest locations and calculating their average values. The distance metric parameter P determines how the distances between neighbors are measured. In this study, $K = 3$ and $P = 5$.

(2) Gradient Boosting Decision Tree *(GBDT)* regression*:* A state-of-the-art tree-based model capable of capturing complex interaction features. The number of trees and the maximum depth were set to 64 and 2, respectively.

(3) Long Short-Term Memory *(LSTM):* A recurrent neural network that employs memory cells and gates to preserve and regulate information, enabling the model to learn long-term temporal dependencies. The number of units and layers is configured as 64 and 2, respectively.

(4) Gated Recurrent Unit *(GRU):* Compared to LSTM, this method represents a simplified Recurrent Neural Network (RNN) architecture that can also capture long-term dependencies in time-series data. Due to its fewer parameters, this method often demonstrates faster convergence and higher efficiency. The number of units and layers was set to 64 and 2, respectively.

(5) Attentional Deep Air Quality Inference Network *(ADAIN)* [33]*:* This method employs fully connected layers and LSTM networks for spatio-temporal modeling of historical data, integrating an attention mechanism to optimize spatio-temporal feature weighting.

(6) Multi-Channel Attention Model *(MCAM)* [34]*:* This approach utilizes GCN and LSTM to perform spatio-temporal modeling of historical data and similarly introduces an attention mechanism to dynamically adjust the model's focus across different spatial regions.

Parameter Settings and Evaluation Metrics

For the experiment design, each of the 15 monitoring station areas was treated as an individual target site. The air quality data corresponding to the target station were excluded from the model's input, while its latitude, longitude, and regional environmental type were embedded into the feature set. Accordingly, a total of 15 experiments were conducted, and the average of the evaluation results was calculated. The hyperparameter configuration of the STAMGCN model was carefully optimized through a combination of systematic experiments and theoretical analysis.

Regarding the model structure design, the STAMGCN model mainly consists of four layers: GCN, spatial fine-grained attention, GRU, and temporal attention layers. The GCN layer consists of Chebyshev graph convolutional layers. The maximum order K of the Chebyshev polynomial and the embedding dimension of the GCN layer are set to 1 and 20, respectively.

The GCN embedding dimension was set to 20, a relatively modest size compared to typical graph neural network applications. This configuration was intentionally selected to prevent overfitting given the limited number of nodes (15 monitoring stations) in the spatial graph. The Chebyshev polynomial order was fixed at $K = 1$ to restrict the graph convolution operation to immediate neighboring nodes only, avoiding over-smoothing effects that may arise when information propagates too far in small-scale graphs. This design choice ensures that each node learns primarily from its direct neighbors while preserving distinct local characteristics.

Given the low dimensionality of the extracted spatial features, the spatial attention layer was configured with two attention heads. A single attention head was insufficient to capture the diversity of spatial relationships (geographical, pollution-based, and meteorological), while increasing beyond two heads yielded diminishing returns and increased computational cost without significant accuracy improvements. Moreover, the two-head configuration effectively balances model capacity with computational efficiency, where each head is designed to specialize in different spatial relationship patterns.

The GRU layer is implemented as a two-level hierarchical structure, with the first and second layers comprising 64 and 128 neurons, respectively. This hierarchical structure allows the model to initially capture simpler, short-term temporal patterns using the smaller GRU layer, followed by the extraction of more complex, long-term dependencies through the larger layer. Therefore, the selected two-layer design offers an optimal trade-off between model complexity and prediction accuracy.

The learning rate was set to 0.001, determined through grid search experiments testing values between 0.0001 and 0.01. Higher learning rates led to training instability, while lower rates resulted in excessively slow

convergence. Moreover, the batch size of 256 was selected to balance training stability and memory efficiency, after being validated through experiments with batch sizes ranging from 64 to 512. To mitigate overfitting, the L1 regularization coefficient (set to 0.001) promotes sparsity in learned weights, while the L2 regularization coefficient (also set to 0.001) controls weight magnitude. These regularization strengths were optimized using the validation set performance monitoring. Early stopping was implemented during model training to prevent overfitting. This strategy ensured optimal generalization while avoiding unnecessary computational costs from prolonged training.

In the field of air quality analysis, the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are effective metrics used for evaluating the predictive performance of models. They are defined as follows:

Table 3. AQI prediction results of training sets obtained using different methods.

Method	Evaluation Metrics	T_0	T_1	T_2	T_3	T_4	T_5	T_6
KNN	MAE	51.91	51.69	53.47	54.75	57.47	61.58	60.28
	RMSE	62.67	61.25	63.58	66.28	67.52	69.52	68.76
GBDT	MAE	36.83	38.05	39.81	43.68	47.62	49.67	52.73
	RMSE	46.67	47.69	51.69	53.96	57.88	61.7	64.6
LSTM	MAE	18.42	18.01	19.63	20.52	22.69	22.90	23.16
	RMSE	23.97	22.30	24.05	24.46	25.98	26.46	27.25
GRU	MAE	17.36	17.12	18.99	19.22	20.08	21.64	22.53
	RMSE	23.05	21.78	23.80	24.36	24.39	26.00	26.45
ADAIN	MAE	15.65	15.23	15.98	16.57	17.66	18.62	19.27
	RMSE	20.67	20.48	20.98	21.22	22.68	23.74	24.69
MCAM	MAE	11.95	12.03	12.26	12.14	13.36	14.38	17.66
	RMSE	16.82	16.75	16.97	17.24	18.16	19.78	21.33
STAMGCN	MAE	9.84	9.61	10.09	9.90	10.82	12.09	14.14
	RMSE	14.22	13.78	14.04	14.01	14.62	16.69	18.53

Table 4. AQI prediction results of validation sets obtained using different methods.

Method	Evaluation Metrics	T_0	T_1	T_2	T_3	T_4	T_5	T_6
KNN	MAE	53.09	53.81	56.38	57.60	61.31	64.37	65.20
	RMSE	63.48	62.39	67.91	68.05	70.24	69.76	71.18
GBDT	MAE	39.36	37.85	41.81	43.76	47.74	51.05	54.18
	RMSE	49.75	51.26	53.47	57.28	58.76	63.48	68.43
LSTM	MAE	20.28	19.86	20.36	22.37	23.76	23.90	24.63
	RMSE	25.38	24.36	25.74	26.89	26.98	27.40	29.10
GRU	MAE	17.93	17.88	19.28	20.56	21.39	21.48	22.89
	RMSE	23.95	21.67	23.93	24.92	25.76	26.40	28.36
ADAIN	MAE	17.02	16.78	16.99	17.25	18.02	20.56	22.40
	RMSE	21.89	20.85	21.64	21.86	22.73	23.98	24.88
MCAM	MAE	12.42	12.30	12.48	12.94	13.59	14.88	16.97
	RMSE	17.03	16.88	18.27	18.69	19.28	20.98	22.36
STAMGCN	MAE	10.05	9.94	9.95	10.20	10.73	12.05	14.09
	RMSE	14.65	13.79	14.04	13.46	15.28	17.35	18.79

$$RMSE = \sqrt{\frac{1}{K}\sum_{i=1}^{K}(y_i - y_i')^2} \quad (12)$$

$$MAE = \frac{1}{K}\sum_{i=1}^{K} \| y_i - y_i' \| \quad (13)$$

where y_i and y'_i represent the observed and predicted values of AQI, respectively, whereas K denotes the total number of all samples.

Prediction Results Obtained Using Different Methods

AQI prediction experiments were conducted using the STAMGCN method and benchmark approaches for both the current time and the next 6 hours. The evaluation results are presented in Tables 3 and 4. Specifically, on the training set, the MAE and RMSE values for AQI prediction at the current time are 9.84 and 14.22, respectively, while the average MAE and RMSE values for the next 6 hours are 11.11 and 15.28, respectively. Regarding the validation set, the MAE and RMSE values for the current time are 10.05 and 14.65, respectively, whereas the average MAE and RMSE values for the next 6 hours are 11.16 and 15.45, respectively. Clearly, these values are significantly lower than those obtained by the benchmark methods, highlighting that the proposed STAMGCN method outperforms all comparative approaches. In addition, other methods based on spatio-temporal networks, such as ADAIN and MCAM, also exhibit superior predictive performance compared to traditional techniques.

Continuous Prediction Results of Different Methods

In people's daily lives, the AQI can be significantly influenced by severe weather conditions, often resulting in sudden changes. In such scenarios, simple model evaluation metrics may not fully capture a model's ability to respond to these situations. Therefore, this study extracted a long continuous sequence with a pronounced trend from the dataset to conduct further analysis. Specifically, the experiment selects station 1015A as the target site for long-term continuous prediction, covering the period from November 1, 2019, to December 1, 2019. The data were input into the trained model for inference, and the results were compared with the original values. Fig. 5 illustrates the long-term continuous AQI prediction capability of the STAMGCN method in comparison with baseline models under a steep trend scenario. It is evident that the XGBT, LSTM, and GRU models perform poorly when abrupt changes occur in the ground truth values. In contrast, the proposed STAMGCN model demonstrates superior

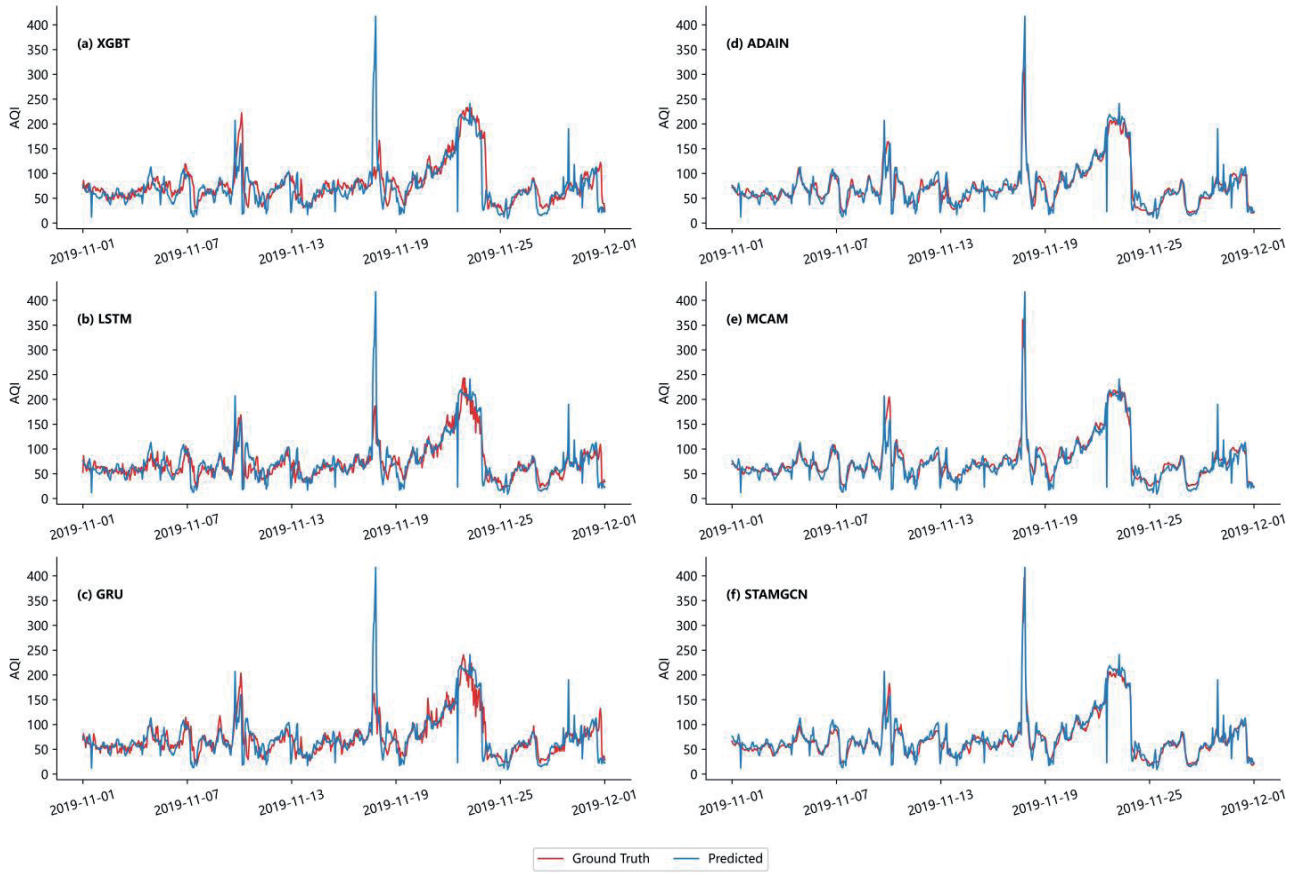

Fig. 5. Long-term prediction of a target location.

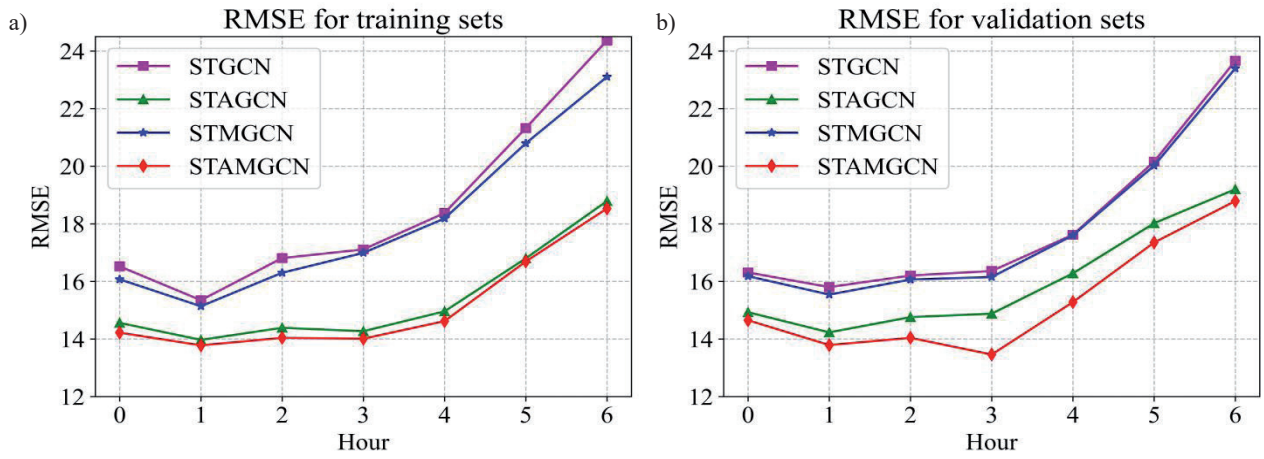

Fig. 6. Comparison of RMSE values calculated from the ablation experiment results using a) the training set and b) validation set.

performance compared to all baseline models under this challenge, with the predicted AQI values closely overlapping the ground truth curve and maintaining consistent trends.

Model Ablation Experiment

In the fine-grained air quality analysis task, the Spatio-Temporal Graph Convolutional Model (STGCN) is capable of learning spatial relationships among monitoring sites through its graph convolutional layer. For the subsequently extracted temporal features, this method can further learn contextual relationships, effectively capturing underlying potential variation patterns. Based on this model, this paper introduces a spatial fine-grained attention module specially designed for fine-grained prediction tasks, which guides the model to focus on deeper correlations between the target site and its neighboring sites. It also integrates multi-graph structures to enhance prior information in the graph convolution module, ultimately achieving excellent results. In this section, ablation experiments were conducted on both the attention module and the multi-graph convolution module to comprehensively evaluate their impact on the model.

Fig. 6 illustrates the ablation experiment results for the current time and the next 6 hours. It is clearly observed that both modules designed in this study exert a positive impact on the model, significantly enhancing prediction performance at the target site. Specifically, the RMSE values of the STGCN and STAMGCN models at the current time are 16.31 and 14.65, respectively, whereas the average RMSE values for the next 6 hours are 18.3 and 15.45, respectively. Therefore, compared to STGCN, STAMGCN improves the AQI prediction accuracy by 10.18% for the current time and by 15.56% for the next 6 hours.

Conclusions

This paper presents an in-depth study on fine-grained air pollution analysis, aiming to estimate current and predict future air quality in target areas lacking direct air quality observations. Given the spatio-temporal correlations between target areas and nearby monitored regions, effective spatio-temporal modeling serves as the core focus of this research. Accordingly, this paper proposes a novel model, termed STAMGCN, which exhibits superior spatial feature encoding ability for non-Euclidean air quality data through its MGCN module. The model's multi-graph structure, which comprises spatial, pollutant, and meteorological correlation graphs, enriches the module with more comprehensive prior knowledge. Furthermore, to make the model more attentive to target area feature information, a spatial fine-grained attention module is designed that allows the model to extract potential influencing patterns between the target area and its nearby monitoring stations. To model temporal dependencies, the GRU recurrent network is employed to extract temporal features from historical data. Combined with temporal attention, it further captures the context of the entire temporal feature set while focusing on significant temporal characteristics. Experimental results show that the proposed model significantly outperforms all baseline methods in terms of MAE and RMSE for the current time and the next 6 hours.

While this study has certain limitations, the model was validated using a dataset from a single city (Beijing), and its generalization to other cities with different climatic conditions and pollution sources should be evaluated. The scarcity of comparable high-resolution air quality datasets with detailed regional type classifications, consistent meteorological measurements, and comprehensive spatial coverage from other regions restricts broader validation.

Finally, for future work, the research scope and targets will be substantially expanded. First, we will investigate multi-city transfer learning by applying

domain adaptation and meta-learning approaches to accelerate model expansion to new cities with limited historical data. Second, we will integrate multi-source remote sensing data, including satellite-derived Aerosol Optical Depth (AOD), land use classifications, and vegetation indices, with ground-based observations via graph neural networks or multi-modal attention mechanisms to effectively address sparse monitoring coverage in suburban and rural areas. Third, we will optimize real-time deployment using model compression techniques (e.g., knowledge distillation, pruning, and quantization) and incremental learning mechanisms to achieve sub-second inference latency for operational forecasting. These add-ons aim to enhance the model's generalization capability and practical value for large-scale air quality management applications.

Acknowledgment

This work was supported by the R&D Program of Beijing Municipal Education Commission (KM202311232018), the Beijing Municipal Natural Science Foundation (Z220015), and the National Natural Science Foundation of China (62276028).

Conflict of Interest

The authors declare no conflict of interest.

References

1. APTE J.S., BRAUER M., COHEN A.J., EZZATI M., POPE C. Ambient $PM_{2.5}$ reduces global and regional life expectancy. Environmental Science & Technology Letters. **5** (9), 546, **2018**.
2. VALLERO D.A. Fundamentals of air pollution. Academic press. 154, **2014**.
3. KUMAR P. A critical evaluation of air quality index models (1960–2021). Environmental Monitoring and Assessment. **194** (5), 1, **2022**.
4. ORGANIZATION W.H. WHO ambient air quality database, 2022 update: status report. World Health Organization, 100, **2023**.
5. PRASHANT K., KUMAR P., PATTON A.P., DURANT J.L., CHRISTOPHER F.H. A review of factors impacting exposure to PM2. 5, ultrafine particles and black carbon in Asian transport microenvironments. Atmospheric Environment. **187**, 301, **2018**.
6. CHEN W., TANG H., ZHAO H. Diurnal, weekly and monthly spatial variations of air pollutants and air quality of Beijing. Atmospheric Environment. **119**, 21, **2018**.
7. HOOD C., MACKENZIE I., STOCKER J., JOHNSON K., CARRUTHERS D., VIENO M., DOHERTY R. Air quality simulations for London using a coupled regional-to-local modelling system. Atmospheric Chemistry and Physics. **18** (15), 11221, **2018**.
8. KIM Y., WU Y., SEIGNEUR C., YOUNGSEOB K., CHRISTIAN S., YELVA R. Multi-scale modeling of urban air pollution: Development and application of a Street-in-Grid model (v1. 0) by coupling MUNICH (v1. 0) and Polair3D (v1. 8.1). Geoscientific Model Development. **11** (2), 611, **2018**.
9. SANTIAGO J.L., BORGE R., MARTIN F., PAZ D., MARTILLI A., LUMBRERAS J., SANCHEZ B. Evaluation of a CFD-based approach to estimate pollutant distribution within a real urban canopy by means of passive samplers. Science of The Total Environment **576**, 46, **2017**.
10. POKHREL R., LEE H. Comparison of Gaussian plume model and lagrangian particle model for the application of coastal air quality modelling. American Journal of Environmental and Resource Economics. **4** (4), 152, **2019**.
11. GARIAZZO C., PAPALEO V., PELLICCIONI A., CALORI G., RADICE P., TINARELLI G. Application of a Lagrangian particle model to assess the impact of harbour, industrial and urban activities on air quality in the Taranto area, Italy. Atmospheric Environment. **41** (30), 6432, **2007**.
12. HUANG H., AKUTSU Y., ARAI M., TAMURA M. A two-dimensional air quality model in an urban street canyon: evaluation and sensitivity analysis. Atmospheric Environment. **34** (5), 689, **2000**.
13. HSIEH H.-P., WU S., KO C.-C., SHEI C., YAO Z.-T., CHEN Y.-W. Forecasting fine-grained air quality for locations without monitoring stations based on a hybrid predictor with spatial-temporal attention based network. Applied Sciences. **12** (9), 4268, **2022**.
14. GUPTA N.S., MOHTA Y., HEDA K., ARMAAN R., VALARMATHI B., ARULKUMARA N. Prediction of air quality index using machine learning techniques: a comparative analysis. Journal of Environmental and Public Health. **2023**, 1, **2023**.
15. LIU B., JIN Y., LI C. Analysis and prediction of air quality in Nanjing from autumn 2018 to summer 2019 using PCR–SVR–ARMA combined model. Scientific Reports. **11** (1), 348, **2021**.
16. ABHILASH M., THAKUR A., GUPTA D., SREEVIDYA B. Time series analysis of air pollution in Bengaluru using ARIMA model. Proceedings of the Ambient Communications and Computer Systems: RACCCS. **2017**, 413, **2018**.
17. GU K., ZHOU Y., SUN H., ZHAO L., LIU S. Prediction of air quality in Shenzhen based on neural network algorithm. Neural Computing and Applications. **32**, 1879, **2020**.
18. KUMAR D. Evolving Differential evolution method with random forest for prediction of Air Pollution. Procedia Computer Science. **132**, 824, **2018**.
19. ABDULLAH S., ISMAIL M., AHMED A. Multi-layer perceptron model for air quality prediction. Malaysian Journal of Mathematical Sciences. **13**, 85, **2019**.
20. LIU X., GUO H. Air quality indicators and AQI prediction coupling long-short term memory (LSTM) and sparrow search algorithm (SSA): A case study of Shanghai. Atmospheric Pollution Research. **13** (10), 101551, **2022**.
21. WANG X., YAN J., WANG X., WANG Y. Air quality forecasting using GRU model based on multiple sensors nodes. IEEE Sensors Letters. **7** (7), 1, **2023**.
22. LI X., PENG L., YAO X., CUI S., HU Y., YOU C., CHI T. Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. Environmental Pollution. **231**, 997, **2017**.
23. GILIK A., OGRENCI A S., OZMEN A. Air quality prediction using CNN+ LSTM-based hybrid deep learning architecture. Environmental Science and Pollution Research. **29** (8), 11920, **2022**.

24. YAN R., LIAO J., YANG J., SUN W., LI F. 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, **2021**.
25. LIAO H., YUAN L., WU M., CHEN H. Air quality prediction by integrating mechanism model and machine learning model. Science of The Total Environment. **899**, 165646, **2023**.
26. WU C.-L., SONG R.-F., ZHU X.-H., PENG Z.-R., FU Q.-Y., PAN J. A hybrid deep learning model for regional O_3 and NO_2 concentrations prediction based on spatiotemporal dependencies in air quality monitoring network. Environmental Pollution. **320**, 121075, **2023**.
27. CHEN J., YUAN C., DONG S., FENG J., WANG H. A novel spatiotemporal multigraph convolutional network for air pollution prediction. Applied Intelligence. **53** (15), 18319, **2023**.
28. LIANG Y., XIA Y., KE S., WANG Y., WEN Q., ZHANG J., ZHENG Y., ZIMMERMANN R.H. Airformer: Predicting nationwide air quality in china with transformers. proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence. **37** (12), 14329, **2023**.
29. XU Y., ZHU Y. When remote sensing data meet ubiquitous urban data: Fine-grained air quality inference. Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), 1252, **2016**.
30. ZHAO Y., FAN S., XIA K., JIA Y., WANG L., YANG W. ASTGC: Attention-based spatio-temporal fusion graph convolution model for fine-grained air quality analysis. Air Quality, Atmosphere & Health. **16** (9), 1761, **2023**.
31. QI Z., WANG T., SONG G., HU W., ZHANG Z. Deep air learning: Interpolation, prediction, and feature analysis of fine-grained air quality. IEEE Transactions on Knowledge and Data Engineering. **30** (12), 2285, **2018**.
32. HAN J., LIU H., XIONG H., YANG J. Semi-supervised air quality forecasting via self-supervised hierarchical graph neural network. IEEE Transactions on Knowledge and Data Engineering. **35** (5), 5230, **2022**.
33. CHENG W., SHEN Y., ZHU Y., HUANG L. A neural attention model for urban air quality inference: Learning the weights of monitoring stations. Proceedings of the AAAI Conference on Artificial Intelligence. **32** (1), 2151, **2018**.
34. HAN Q., LU D., CHEN R. Fine-Grained Air Quality Inference via Multi-Channel Attention Model. In Proceedings of the International Joint Conference on Artificial Intelligence, 2512, **2021**.