Computer Modeling in
Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.083813

ARTICLE

Frequency-Aware Spatiotemporal Graph Modeling of Multi-Pollutant Dynamics in Industrial Air Quality Systems

Chia-Hui Liu* **and Chen-Chuan Cheng**

Department of Electronic Engineering, National Formosa University, Yunlin, Taiwan

*Corresponding Author: Chia-Hui Liu. Email: liuch@nfu.edu.tw

Received: 11 April 2026; Accepted: 30 June 2026; Published: 28 August 2026

ABSTRACT: Industrial air quality forecasting remains challenging due to nonlinear pollutant formation, localized emissions, meteorological variability, and nonstationary spatiotemporal dependencies among monitoring stations. This study proposes FFTGNet, a frequency-aware spatiotemporal graph neural network for multi-pollutant forecasting in industrial air quality systems. It integrates an FFT-guided dominant-period estimation and period-folding module with a temporal-to-spatial graph backbone composed of TemporalGLU and Chebyshev graph convolution. The frequency-guided module reorganizes input sequences into intra-period and inter-period representations, TemporalGLU adaptively filters nonlinear temporal fluctuations and short-term spikes, and ChebGCN propagates information across inter-station spatial dependencies. Experiments were conducted using five years of hourly observations from ten Photochemical Assessment Monitoring Stations in Yunlin, Taiwan, covering O_3, CO, PM_{10}, meteorological variables, and PMF-derived VOC source features. Compared with conventional baselines and recent spatiotemporal models, including MGCGRU-SAN and AirNN-Adapted, the proposed method achieves the most consistent performance across pollutants and forecasting horizons. For 1 h forecasting, it obtains R^2 values of 0.9509, 0.8574, and 0.8556 for O_3, CO, and PM_{10}, respectively. For O_3, it further maintains R^2 values of 0.8116 and 0.6797 at the 3 and 6 h horizons, while showing smooth degradation rather than abrupt performance collapse up to 24 h. Ablation results confirm the complementary contributions of frequency-guided period folding, gated temporal modeling, and graph-based spatial propagation. Computational analysis further indicates that FFTGNet remains feasible for server-side near-real-time deployment, with an inference latency of 71.7 ms per sample. These results demonstrate that coupling frequency-domain priors with spatiotemporal graph learning improves forecasting accuracy, stability, and applicability for industrial multi-pollutant air quality systems.

KEYWORDS: Frequency-aware learning; spatiotemporal graph models; multi-pollutant; industrial air quality systems; fast fourier transform

1 Introduction

With the advancement of sensing and automation technologies, engineering systems have become increasingly data-driven, making real-time monitoring and accurate prediction of dynamic processes essential yet challenging. Air quality forecasting exemplifies this complexity due to nonlinear interactions among emission sources, meteorological factors, and regional transport, as well as strong nonstationarity and multi-scale temporal dynamics. Traditional statistical and physical models are limited in capturing such complex relationships, while classical time series approaches (e.g., ARIMA and VAR) struggle with nonlinear dependencies and cross-station interactions. Although deep learning models, including CNNs, LSTMs, and graph convolutional networks (GCNs), have demonstrated improved performance, existing methods often focus on single pollutants or short-term prediction and lack robustness under multi-step forecasting settings, cross-seasonal variations, and distributional shifts. In particular, most spatiotemporal models operate solely in the time domain and fail to explicitly capture frequency-domain characteristics, limiting their ability to model periodic temporal patterns and maintain forecasting stability. These limitations highlight the need for a unified framework that jointly models temporal, spatial, and frequency-domain information for robust air quality forecasting. Recent studies and reviews have shown that air-quality forecasting involves nonlinear temporal dependencies, spatial transport, meteorological variability, and pollutant-specific formation mechanisms, making deep learning and spatiotemporal modeling increasingly important for environmental prediction [1–3].

To address the above challenges, this study proposes FFTGNet (Fast Fourier Transform Temporal Graph Network), a deep learning architecture that integrates temporal, spatial, and frequency-domain information. From an AI perspective, FFTGNet introduces a Fast Fourier Transform mechanism within the temporal branch to adaptively adjust feature representations, compensating for the limitations of conventional time-domain models in capturing periodic signals. Meanwhile, a spectral weighting strategy is incorporated into the graph convolutional network to enhance stable spatial information integration across monitoring stations. From an engineering standpoint, hourly pollutant data collected from multiple photochemical monitoring stations in the Yunlin Industrial Zone, Taiwan (2020–2024), are used as a case study. Unlike conventional air quality stations, photochemical stations simultaneously measure multiple precursors and meteorological variables, allowing the model to learn potential relationships between ozone formation and pollutant transport from multivariate inputs. The high dimensionality and multiscale spatiotemporal interactions inherent in these data provide an ideal testbed for evaluating deep learning models under nonstationary and complex environmental forecasting scenarios.

In addition, the Yunlin region in Taiwan represents a typical high-emission industrial environment due to intensive petrochemical activities. Characterized by elevated levels of volatile organic compounds (VOCs) and nitrogen oxides (NOx), the region exhibits complex pollution dynamics. Under the influence of land–sea breeze circulation and solar radiation, pronounced afternoon ozone accumulation is frequently observed, making it a representative and challenging scenario for evaluating multi-pollutant spatiotemporal forecasting models. The availability of multi station observational data and well defined industrial emission patterns makes it a suitable verification setting for spatiotemporal forecasting models involving multiple pollutants. To evaluate the proposed model's capability to maintain consistent predictive performance under nonstationary environmental conditions, this study adopts photochemical monitoring data from the Yunlin region in Taiwan as a representative case, highlighting its applicability to real time monitoring and intelligent emission control. The proposed FFTGNet model offers two key contributions. First, it introduces a Fourier-guided mechanism into the temporal branch, enabling the model to simultaneously capture periodic temporal patterns and short-term fluctuations, thereby addressing the limitations of traditional time-domain approaches in recognizing periodic patterns. Second, on the spatial dimension, FFTGNet employs a graph convolutional structure to enhance the modeling of inter station dependencies, improving the stability and accuracy of multi station information integration in nonstationary pollution fields. This design enables the model to effectively learn spatiotemporal patterns of pollutant distributions under nonstationary and multiscale data conditions, thereby improving its generalization ability and demonstrating its potential for applications in intelligent environmental monitoring and real time early warning systems.

2 Related Work

2.1 Mechanisms and Monitoring Data for Air Pollutant Forecasting

Air pollutant concentration forecasting requires understanding the underlying physical and chemical mechanisms as well as the quality of available monitoring data. Variations in pollutant concentrations are primarily driven by verifiable generation and transformation processes and are constrained by the detection capabilities of monitoring networks. This section reviews existing forecasting approaches and the characteristics of meteorological data, providing the foundation for the design of subsequent methods and experiments. Ozone is a typical secondary pollutant, primarily generated through photochemical reactions between volatile organic compounds (VOCs) and nitrogen oxides (NOx) under solar radiation. Due to the highly heterogeneous reactivities of different VOC species, their marginal contributions to O_3 formation vary across chemical and meteorological regimes. Under NOx limited conditions, practical control strategies typically prioritize the identification and suppression of VOCs with high incremental reactivity [4]. Effective O_3 prediction and control depend on identifying local sources of highly reactive VOCs and their temporal variability [4,5]. For source attribution, Positive Matrix Factorization (PMF) infers latent source profiles and relative contributions from observations and is particularly useful when explicit source signatures are lacking or pollutant mixtures are complex [6,7]. PMF has been widely applied in Taiwan's petrochemical corridors and has been repeatedly validated for stability and utility [8–11]. In addition, PMF is often combined with Maximum Incremental Reactivity (MIR) to translate source factors into O_3 formation potentials for control prioritization [4]. Such source level features can also be incorporated as exogenous inputs in data driven models to disentangle emission vs. meteorological influences, thereby improving robustness and policy relevance [12,13].

Because highly reactive VOCs are rapidly consumed by daytime photochemistry, practitioners often estimate their initial (pre reaction) concentrations prior to source apportionment to avoid underestimating specific sources, consistent with reactivity driven O_3 formation mechanisms [4,6,7]. In summary, PMF provides source based, localized information that complements critical steps in O_3 formation pathways and offers physically meaningful features for forecasting and control. Unlike O_3, which is governed by rapid photochemistry, fine PM_{10} is influenced by both primary emissions and multiple secondary formation pathways, whereas CO primarily reflects combustion sources and background exchange. These differences necessitate forecasting systems capable of simultaneously addressing rapid photochemistry, slower secondary aerosol chemistry, and stable combustion dynamics, while ensuring cross scale temporal sensitivity and accurate source attribution. Building on these mechanisms, air quality monitoring networks provide the observational backbone for model development and validation. Regulatory stations supply hourly data on O_3, NO, NO_2, VOCs, and meteorological variables. However, inter station datasets often contain noise, gaps, and inconsistencies. In Taiwan's industrial corridors, long term VOC records have been applied to construct source profiles and support policy development [8]. Multi pollutant forecasting typically targets hourly O_3 and CO. Systematic documentation of pollutant and meteorological data characteristics, including variability related to local meteorology, land-surface forcing, and ozone variability, establishes a foundation for feature selection, evaluation metrics, and modeling protocols, thereby supporting cross-station models and weather normalization [14–16]. In summary, understanding the physical and chemical mechanisms of pollutant formation, together with the characteristics and limitations of monitoring data, provides a scientific foundation for forecasting. Nevertheless, conventional receptor modeling and mechanistic analyses alone remain insufficient to capture the nonlinear dynamics and cross scale variability inherent in air pollution. To address these limitations, recent studies have increasingly adopted artificial intelligence (AI) and machine learning (ML) methods, which offer enhanced capabilities for integrating heterogeneous data, learning complex dependencies, and improving predictive accuracy. The following section reviews existing research on AI based approaches for air pollution forecasting.

2.2 Machine Learning Approaches for Air Pollution Monitoring and Prediction

In recent years, the application of Artificial Intelligence (AI) and Machine Learning (ML) in air pollution prediction and monitoring has undergone rapid development. Traditional deterministic chemical transport models and linear regression methods often encounter limitations in highly heterogeneous spatiotemporal environments. These approaches are constrained by the sparse distribution of monitoring networks, the complexity of atmospheric processes, and the nonlinear interactions among meteorological factors, pollutant emissions, and geographical conditions. To address these challenges, researchers have increasingly adopted ML and Deep Learning (DL) architectures, which are capable of capturing nonlinear dependencies, integrating multimodal data sources, and demonstrating stronger generalization capabilities across both spatial and temporal dimensions. Recent studies have validated the applicability and potential of machine learning techniques across various pollutants and monitoring contexts, particularly in regions with limited monitoring infrastructure. Gradient boosting methods, such as CatBoost and XGBoost, have been utilized to predict fine particulate matter (PM2.5) concentrations, while classification models have been incorporated to estimate the Air Quality Index (AQI) in urban areas of South Africa [17]. This framework addresses prediction challenges for both monitored and unmonitored cities and further extends applicability to regions without monitoring stations. By integrating heterogeneous data sources, including meteorological variables and satellite derived features, ensemble based models have demonstrated superior performance compared with conventional baseline approaches. Beyond PM2.5, recent studies have increasingly emphasized PM1.0, which has gained recognition as an important indicator for health risk assessment. Due to the extreme sparsity of monitoring data, data driven approaches have been explored to supplement and extend PM1.0 observations. Random Forest models have been employed in South Korea to predict PM1.0 concentrations nationwide using limited monitoring stations, thereby demonstrating the utility of tree based methods for data augmentation and subsequent epidemiological analysis [18]. Furthermore, a two stage framework combining linear mixed effects models with bagged tree ensembles has been developed to estimate hourly PM1.0 levels across eastern and central China. By integrating satellite aerosol products, geographic predictors, and meteorological variables, this approach has shown that machine learning can effectively generate continuous, high resolution PM1.0 maps in regions lacking direct monitoring [19].

In addition to direct predictions of particulate matter with different aerodynamic diameters, recent studies have explored the integration of Land Use Regression (LUR) with machine learning to enhance the resolution and accuracy of exposure modeling. A case study conducted in a Canadian port city compared Ordinary Least Squares based LUR (OLS LUR) with Random Forest augmented LUR (RF LUR) for estimating trace elements of PM1.0. The findings demonstrated that machine learning based LUR provided superior predictive accuracy and alleviated the overestimation issues commonly observed in traditional regression approaches [20]. This hybrid framework highlights the capability of machine learning to incorporate nonlinear relationships and complex feature interactions, thereby complementing conventional statistical techniques, particularly in environments with diverse emission sources. Beyond conventional statistical methods and monitoring data augmentation models, recent research has emphasized the use of deep learning to integrate multimodal data sources for modeling complex pollution dynamics. Long Short Term Memory (LSTM) frameworks have been applied to combine color features extracted from street view images with multiple particulate matter indicators, including PM1.0, PM2.5, for air pollution prediction [21]. This approach demonstrates that visual semantic information embedded in urban morphology can serve as a proxy for emission and transport patterns, thereby substantially enhancing both the accuracy and interpretability of predictive models.

Recent studies have also used UAV-based microscale measurements to characterize the spatial and vertical distributions of particulate matter and gaseous pollutants around built environments. Such measurements can reveal the influence of building-induced airflow, leeward-side accumulation, and local turbulence on pollutant distributions, providing complementary observations for interpreting microscale dispersion patterns [22]. Beyond particulate matter, ozone prediction poses additional challenges due to the nonlinear interactions among chemical reactions, meteorological conditions, and precursor emissions. To address these complexities, spatiotemporal deep learning architectures have been increasingly adopted. A recent approach, termed O3ConvNet, introduces a convolution based dynamic framework that integrates modern temporal convolution modules with adjacency matrices constructed from geographic distance and Dynamic Time Warping (DTW) similarity. This architecture demonstrates flexibility in handling datasets with varying monitoring densities and data quality, while exhibiting strong potential for cross city generalization [23]. In ozone prediction, recent approaches have adopted Deep Ensemble Machine Learning (DEML) combined with Bayesian fusion to generate high resolution spatiotemporal maps. One such framework integrates base learners including Random Forest (RF), Extreme Gradient Boosting (XGB), and Gradient Boosting Machine (GBM) for ensemble fusion, resulting in high resolution O_3 maps across New South Wales, Australia [24]. A comparative summary of representative forecasting models and their characteristics is provided in Table 1. Building upon convolutional and ensemble strategies, subsequent studies have placed increasing emphasis on incorporating geographic and temporal dependencies while enhancing model robustness and interpretability. A representative example is Geo STO3Net, which leverages a ResNet based spatial encoder, a Transformer based temporal encoder, and a Deep Neural Network (DNN) decoder to explicitly capture complex spatiotemporal interactions. This architecture has demonstrated substantial improvements in the accuracy of O_3 prediction [25].

Advancements in model architecture have also focused on improving the robustness and efficiency of deep learning for pollutant prediction. Stabilized Long Short Term Memory networks incorporate memory regularization and modified cell designs to enhance information flow control, thereby improving convergence speed and model stability [26]. Another approach integrates k means clustering with Artificial Neural Networks (ANNs) to facilitate trend classification and reduce training costs [27]. Comparative studies further indicate that simpler nonlinear machine learning models can achieve performance comparable to deep architectures under certain scenarios [28]. These findings suggest that selecting model complexity according to application requirements may offer greater practicality than consistently relying on deeper networks. With respect to interpretability, recent deep learning frameworks have incorporated feature and temporal selection mechanisms to balance predictive performance with model transparency. For example, the introduction of a Time Selection Layer enables embedded selection of both features and time intervals, leading to improved accuracy and enhanced interpretability [29]. Beyond spatiotemporal modeling of ozone, related studies have also emphasized source attribution and classification tasks. In particular, methods have been developed to attribute sources of "consumed" volatile organic compounds (VOCs), which have already participated in the formation of ozone and secondary organic aerosols, thereby providing new insights into precursor attribution [30]. In addition, atmospheric physics-based simulations have been employed to complement observational source analysis. For example, a TAPM-based analysis in the Taichung Basin showed that large stationary sources, topography, synoptic weather types, and regional wind fields shape $PM_{2.5}$ transport and high-pollution events [31]. In classification-oriented pollution tasks, hybrid architectures have also demonstrated superior performance. A recent system integrating Convolutional Neural Networks (CNNs) with Extreme Gradient Boosting (XGBoost) has been applied to classify multiple pollutants based on the Air Quality Index (AQI) [32]. In classification oriented pollution tasks, hybrid architectures have also demonstrated superior performance. A recent system integrating Convolutional Neural Networks (CNNs) with Extreme Gradient Boosting (XGBoost) has been applied to classify multiple pollutants based on the Air Quality Index (AQI). Compared with traditional classifiers such as Decision Trees and Support Vector Machines (SVMs), this approach achieved significantly better performance, highlighting the effectiveness of combining deep learning and machine learning for multi class air pollution classification problems [32].

Table 1: Comparison of forecasting methods for air pollutant prediction.

Method	Typical Models	Strengths	Limitations	Ref.
Statistical Models	Autoregressive Integrated Moving Average (ARIMA)	Simple, interpretable, baseline trends	Limited nonlinear modeling, short memory	[1]
Physical Models	The Air Pollution Model (TAPM)	Mechanistic, policy relevance	High computational cost, sensitive to inputs	[31]
Machine Learning	Random Forest (RF), Extreme Gradient Boosting (XGBoost)	Handles nonlinearities, feature selection	Needs sufficient data, less interpretable	[17–19]
Deep Learning	Long Short Term Memory (LSTM), Convolutional Neural Network–LSTM (CNN LSTM)	Strong spatiotemporal learning, multimodal integration	Data hungry, black box, training cost	[21,25,27]
Hybrid Models	Positive Matrix Factorization with Machine Learning, Land Use Regression with Random Forest	Combines physical interpretability and predictive power	Complex design, requires multi source data	[5,8,11,12]

In summary, while prior work has made substantial progress in modeling multi pollutant interactions and spatiotemporal dependencies, key limitations remain from a modeling perspective. These include insufficient generalization across heterogeneous regions, limited robustness to sparse or incomplete monitoring data, and inadequate modeling of nonlinear and multi scale temporal structures. In particular, most deep learning frameworks operate solely in the time domain, overlooking the frequency domain characteristics inherent in pollution dynamics. This technical gap motivates the development of architectures that explicitly integrate temporal, spatial, and spectral information for enhanced forecasting performance.

3 The Proposed Method

This study proposes FFTGNet (Fast Fourier Transform Temporal Graph Network), a deep learning architecture designed to capture both multi scale temporal periodicity and spatial dependencies of air pollutant concentrations as illustrated in Fig. 1. FFTGNet consists of three core modules: (1) a frequency guided temporal alignment module, which employs Fast Fourier Transform (FFT) [33] and Period Folding Transform to align input sequences along dominant cycles; (2) a frequency aware spatiotemporal backbone composed of interleaved TemporalGLU [34] and Chebyshev Graph Convolution (ChebGCN) [35] blocks; and (3) a static graph support module that encodes fixed inter station relationships via a precomputed adjacency matrix for graph convolution [36]. The following subsections detail the design and functionality of each module.

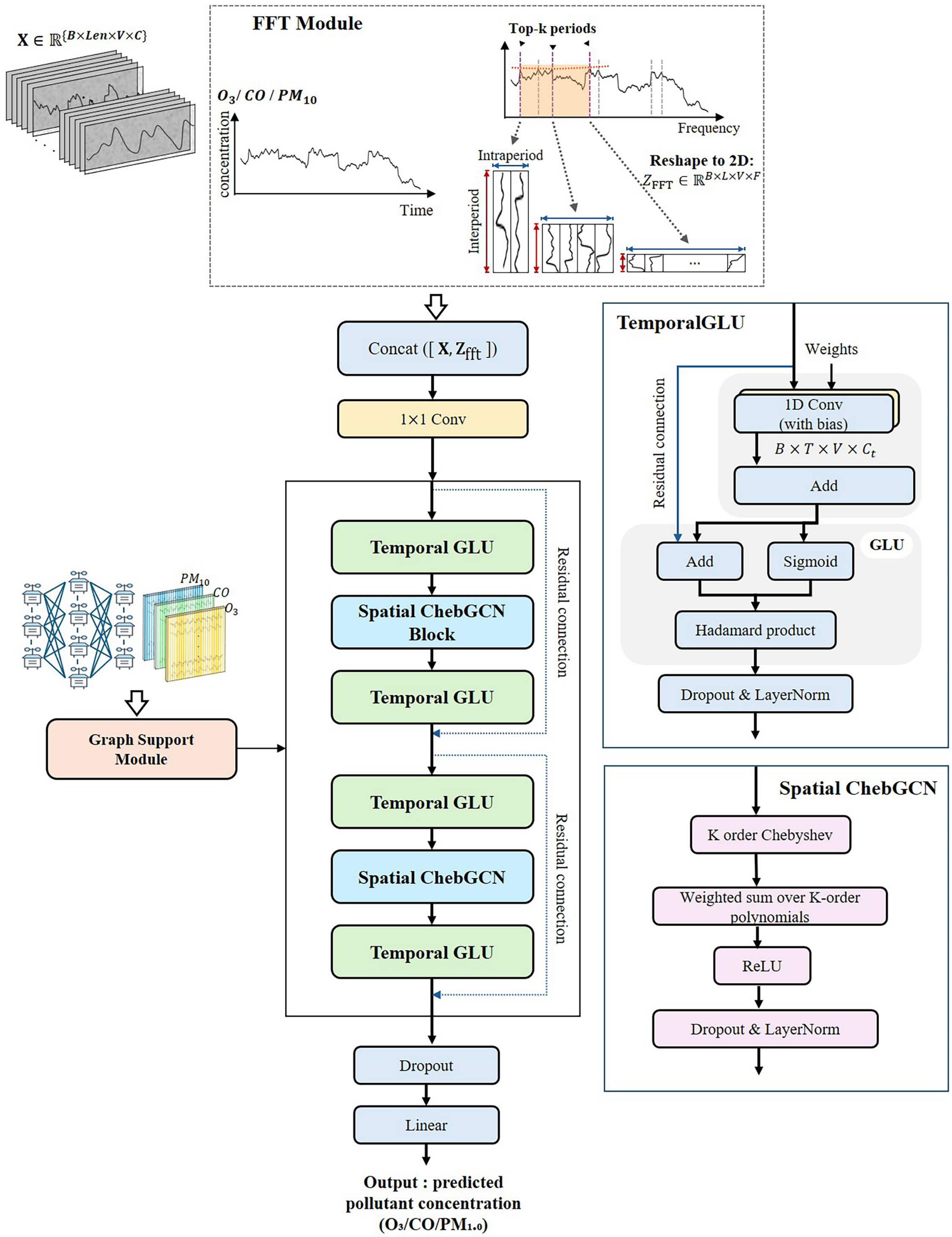


Figure 1: Overall architecture of the proposed FFTGNet model.

As shown in Fig. 1, FFTGNet first applies the Fast Fourier Transform (FFT) to extract dominant periodicities and realigns input sequences using a Period Folding Transform.

The processed sequences are then passed to an interleaved backbone combining TemporalGLU and Chebyshev Graph Convolution (ChebGCN) modules to capture temporal and spatial dependencies.

3.1 Data Representation and Preprocessing

This study utilizes hourly observational data from the Photochemical Assessment Monitoring Stations (PAMS) operated by the Ministry of Environment, Taiwan. Unlike conventional air quality monitoring stations, PAMS not only measure standard pollutants such as ozone, nitrogen oxides (NO_x), CO, and particulate matter (PM_{10}), but also provide measurements for 54 species of volatile organic compounds (VOCs), enabling comprehensive characterization of ozone precursors. Because ozone forms nonlinearly from VOCs and nitrogen oxides under meteorological conditions, PAMS data provide richer precursors for modeling both formation mechanisms and spatiotemporal variability. Ten PAMS stations in Yunlin, Taiwan (Fig. 2), located near the Formosa Mailiao Petrochemical Complex, were selected as a representative industrial corridor.

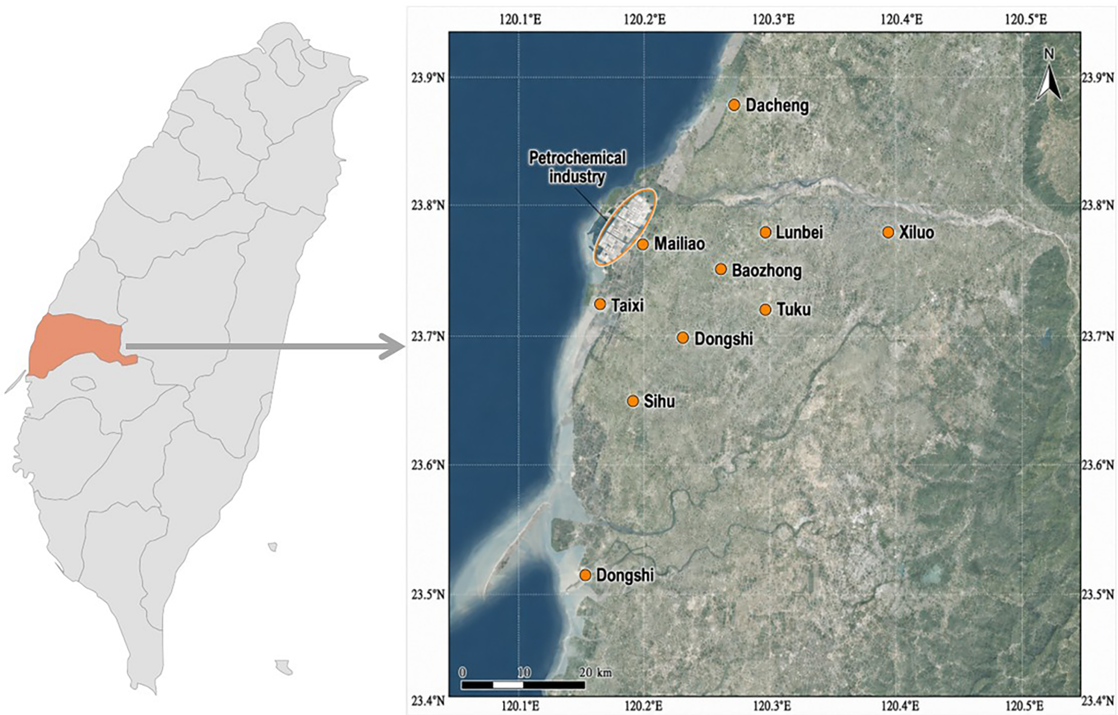


Figure 2: Geographical distribution of ten photochemical monitoring stations. Abbreviations: DC (Douliu City), ML (Mailiao), TX (Tuku), SH (Shuilin), DSI (Dounan Senior Industrial), LB (Lunbei), BZ (Baozhong), TK (Taixi), DS (Dounan Station), and XL (Xiluo).

To ensure consistency and independence across samples, the dataset undergoes a three step preprocessing procedure: (1) Temporal alignment: All station records are aligned to hourly timestamps based on the national standard time axis. (2) Outlier and missing value handling: Maintenance periods are removed. Short gaps (less than 3 h) are interpolated linearly. Outliers are detected using the Median Absolute Deviation (MAD) method, where samples deviating beyond a predefined threshold are truncated. For values below the Method Detection Limit (MDL), a value of MDL/2 is used, following standard environmental monitoring practices to reduce bias introduced by low concentration data. (3) Normalization: All features are standardized using z score normalization based on the mean and standard deviation of the training set, applied separately to each station to prevent data leakage.

After preprocessing, the dataset was organized as a three-dimensional tensor $S \in \mathbb{R}^{T\times V\times C}$, where T denotes the total number of timesteps, V the number of monitoring stations, and C the number of features. Training samples were generated using a sliding window of length L. For each sample, $X_i = S[i{:}\, i+L, :, \mathcal{E}]$, $Y_i = S[i+L, :, \text{target}]$ where $\mathcal{E}$ represents the set of exogenous variables (e.g., meteorological and precursor features) and target $\in \{O_3, CO, PM_{10}\}$. To incorporate chemically interpretable source information, Positive Matrix Factorization (PMF) was applied to the VOC matrix. PMF, a variant of non-negative matrix factorization, decomposes the data into latent source profiles and their temporal contributions, serving as an unsupervised feature extractor for emission sources. Based on reconstruction error and factor interpretability, six representative source components were identified: AAM (Aromatic and Aliphatic Mixtures), CM (Combustion Mixtures), IC (Industrial Coatings), PP (Petroleum Processing), SU (Solvent Usage), and VE (Vehicular Emission). These latent source features, together with meteorological variables and pollutant histories, were integrated as model inputs to enhance the network's ability to capture the temporal evolution of emission sources and pollutant interactions. These PMF-derived source features help the model interpret the composition and temporal variability of emission sources. The processed data, combining pollutant, meteorological, and source information, are then provided as inputs to the FFT Guided Temporal Graph Network (FFTGNet) described in the following section.

3.2 Model Architecture: Frequency Spatiotemporal Graph Network (FFTGNet)

The proposed FFTGNet comprises two modules: a frequency-domain analysis branch and a spatiotemporal graph-convolutional backbone.

3.2.1 Guided Period Extraction and Folding Module

(1) FFT Based Dominant Period Estimation

To stabilize per epoch period estimation, we average the current-batch target series across stations to obtain $z \in \mathbb{R}^L$.

$$Z_k = \sum_{t=0}^{L-1} z_t e^{-j2\pi tk/L} \tag{1}$$

The Fast Fourier Transform (FFT) converts this temporal signal into the frequency domain, as defined in Eq. (1), where Z_k denotes the complex coefficient of the k-th frequency component.

The magnitude spectrum $A_k = |Z_k|$ reflects the energy of each frequency. The frequency with maximum spectral magnitude is selected as the dominant component:

$$\omega^{(e)} = \arg \max_{k\in 1,\ldots,\lfloor L/2\rfloor} A_k \tag{2}$$

The estimate is smoothed across epochs via exponential moving average (EMA) and mapped to an integer period:

$$p^{(e)} = \left\lfloor \frac{L}{\max(1, \tilde{\omega}^{(e)})} \right\rfloor \tag{3}$$

Batch averaging mitigates station-level noise, while EMA prevents oscillatory updates.

Eq. (1) defines the FFT transformation from the temporal to frequency domain, while Eq. (2) identifies the dominant frequency with maximum spectral energy. Eq. (3) shows that the smoothed dominant frequency is converted into an integer period $p^{(e)}$, ensuring numerical stability and consistent temporal rhythm estimation. Batch averaging mitigates station-level noise, and EMA prevents oscillatory updates across epochs.

(2) Period Folding and Feature Integration

After estimating the dominant period $p^{(e)}$, a Period Folding Transform (PFT) reorganizes the original one-dimensional sequence into a two-dimensional tensor that conforms to the detected periodic structure. The transformation folds the long sequence into multiple periodic segments, allowing the model to capture both fine-grained dependencies within each period and broader variations across periods. To ensure divisibility along the time axis, the sequence length L is adjusted to the nearest integer multiple of the estimated period $p^{(e)}$. The adjusted temporal length L' and the corresponding number of folds M are computed as:

$$L' = \left\lfloor \frac{L}{p^{(e)}} \right\rfloor * p^{(e)} \tag{4}$$

$$M = \frac{L'}{p^{(e)}} \tag{5}$$

where L denotes the original sequence length and M represents the number of complete folds. After this transformation, the reorganized data tensor is expressed as $X' \in \mathbb{R}^{B\times p^{(e)}\times M\times V\times C}$ where B is the batch size, V is the number of nodes (or monitoring stations), and C is the number of input channels. The second and third dimensions $(p^{(e)}, M)$ correspond to the intra-period length and the number of folds, respectively. Subsequently, a depthwise two-dimensional convolution is applied on the $(p^{(e)}, M)$ plane for each channel to jointly learn intra-period and inter-period dependencies: $Z_{\mathrm{FFT}} \in \mathbb{R}^{B\times L\times V\times F}$ where F denotes the number of convolution filters. The periodic features Z_{FFT} are then unfolded back to the temporal axis, concatenated with the original input, and compressed by a 1×1 convolution to restore the original dimensionality: $\tilde{X} \in \mathbb{R}^{B\times L\times V\times C}$. Finally, $\tilde{X}$ is fed into the first TemporalGLU module. This design ensures that temporal features are modulated by the current dominant period $p^{(e)}$ prior to temporal modeling, allowing the network to adapt dynamically to evolving periodicity while maintaining temporal continuity.

The use of FFT is motivated by the observation that air pollutant concentrations often exhibit quasi-periodic patterns (e.g., diurnal cycles driven by photochemical reactions and human activities). Instead of assuming strict stationarity, we aim to capture dominant periodic tendencies embedded in the data. By transforming the temporal signal into the frequency domain, the FFT highlights frequency components with high spectral energy, which correspond to recurrent temporal structures. These dominant frequencies serve as structural priors for reorganizing the time series. Based on this observation, we extract the dominant frequency and convert it into an integer period $p^{(e)}$, which is then used to guide the subsequent period folding transformation. This allows the model to align temporal patterns across cycles while still preserving variability under nonstationary conditions.

Compared with localized time-frequency methods such as Short-Time Fourier Transform (STFT) and Wavelet Transform, FFT mainly provides a global dominant periodic prior rather than detailed local frequency decomposition. STFT and Wavelet Transform are more suitable for capturing abrupt local frequency changes and multi-scale transient events. However, they also require additional hyperparameter selection, including window size, scale range, and mother wavelet design, and generally introduce higher computational complexity. Considering that the hourly pollutant observations in this study exhibit clear quasi-periodic structures associated with photochemical and emission-activity cycles, FFT provides an efficient mechanism for extracting dominant periodic tendencies prior to spatiotemporal graph learning. It should be emphasized that the FFT operation in this module is not intended to fully reconstruct the pollutant signal or to assume strict stationarity over the entire observation period. Instead, FFT is used only to estimate a dominant quasi-periodic prior from the input sequence. This prior provides a coarse but informative temporal alignment cue for the Period Folding Transform. Therefore, the novelty of the proposed module does not lie in the standalone use of FFT, but in using the FFT-derived dominant period to reorganize the temporal structure before spatiotemporal graph learning. Unlike conventional FFT-based feature extraction, where spectral coefficients are directly used as additional features, the proposed design uses frequency information to guide period folding and explicitly separate intra-period and inter-period dependencies. Moreover, the folded representation does not replace the original temporal sequence. It is concatenated with the original input and compressed through a 1×1 convolution before being passed to TemporalGLU and ChebGCN. This design allows the subsequent neural modules to learn residual short-term fluctuations, nonlinear deviations, and inter-station spatial dependencies that cannot be fully represented by the dominant period alone. Thus, FFT serves as a frequency-guided temporal prior rather than the sole mechanism for modeling nonstationary pollutant dynamics. Nevertheless, abrupt environmental events, such as dust storms, temporary industrial shutdowns, or extreme weather fluctuations, may introduce local phase shifts, frequency changes, or transient spikes. In such cases, the estimated dominant period may become less representative of local dynamics, and the folding operation may cause local temporal misalignment or partial distortion. Strictly speaking, aliasing is mainly associated with insufficient sampling frequency; in the proposed FFT-based Period Folding mechanism, the more relevant risk is local temporal misalignment when abrupt events disrupt the dominant periodic structure. Localized time-frequency methods, such as Short-Time Fourier Transform (STFT) and Wavelet Transform, are more suitable for capturing local frequency changes and multi-scale transient patterns. However, they also require additional design choices, including window size, scale range, and mother wavelet selection, and may increase computational complexity. Considering that the hourly sampled pollutant data in this study exhibit clear diurnal and emission-activity-related quasi-periodic patterns, FFT is adopted as an efficient mechanism for extracting dominant periodic priors, while STFT, Wavelet Transform, and adaptive frequency modules are regarded as future extensions for improving robustness under abrupt environmental events.

3.2.2 Frequency-Aware Temporal Graph Backbone

After obtaining the frequency-guided representations, the model enters the frequency aware temporal graph backbone, which interleaves TemporalGLU and Chebyshev Graph Convolution (ChebGCN) modules to jointly model temporal dynamics and spatial dependencies across monitoring stations. TemporalGLU focuses on short and long term temporal relationships, while ChebGCN captures spatial correlations using a fixed graph support. By alternating between these modules, the backbone learns coupled spatiotemporal patterns and produces frequency-aware representations that characterize both pollutant periodicity and inter-station propagation. The backbone consists of three key components: TemporalGLU, Graph Support, and ChebGCN, described in the following sections.

(1) TemporalGLU Module

TemporalGLU is adopted because pollutant time series contain both smooth recurrent trends and irregular short-term fluctuations. For example, O_3 usually exhibits a clear diurnal pattern associated with solar radiation and photochemical reactions, whereas CO may contain abrupt local emission spikes and PM_{10} may show short-term accumulation under stagnant meteorological conditions. A simple temporal convolution may treat these signals uniformly, making it difficult to distinguish informative temporal patterns from noise or transient disturbances. The gated mechanism in GLU provides an adaptive feature selection function: the convolutional branch extracts candidate temporal features, while the sigmoid gate controls how much information should be retained at each time step. This helps preserve dominant temporal trends, suppress noisy fluctuations, and improve robustness to abrupt pollutant variations. The TemporalGLU module models temporal dynamics by combining one-dimensional temporal convolution with a Gated Linear Unit (GLU) to adaptively regulate feature flow according to temporal context. Given an input feature tensor $Z \in \mathbb{R}^{B\times L\times V\times C}$, the gating operation is defined as

$$\mathrm{GLU}(X) = (W_z * X + b_z) \odot \sigma(W_g * X + b_g) \tag{6}$$

where W_z and W_g are convolution kernels, b_z and b_g are bias terms, $\sigma(\cdot)$ is the sigmoid activation function, and $\odot$ represents element wise multiplication. Eq. (6) shows that the Gated Linear Unit (GLU) dynamically modulates temporal feature flow, where the sigmoid gate $\sigma(W_g * Z + b_g)$ determines the proportion of information retained at each time step. The first term performs feature transformation, while the second acts as a dynamic gating mask that determines the amount of information preserved at each time step. A residual shortcut connection further stabilizes training and improves temporal feature propagation. The simplified output is expressed as:

$$Y_t = \phi\left(\tilde{X}_t\right) \odot \sigma\left(\psi\left(\tilde{X}_t\right)\right) + \tilde{X}_t \tag{7}$$

where $\phi(\cdot)$ and $\psi(\cdot)$ denote two learnable temporal convolution mappings, $\sigma(\cdot)$ is the sigmoid activation, and $\odot$ represents element-wise multiplication. Eq. (7) shows that the gated convolutional branch dynamically controls the amount of new information added to the residual stream, enabling adaptive fusion of short-term fluctuations and long-term trends.

(2) Graph Support and Spatial Graph Construction

To perform spatial convolution, the ChebGCN module requires a predefined graph structure encoding inter-station relationships. A static correlation graph is built from the training data and remains fixed during inference. Node similarity is computed using the Pearson correlation coefficient between pollutant time series:

$$\rho_{uv} = \frac{\sum_t (X_t^{(u)} - \overline{X}^{(u)})(X_t^{(v)} - \overline{X}^{(v)})}{\sqrt{\sum_t (X_t^{(u)} - \overline{X}^{(u)})^2}\sqrt{\sum_t (X_t^{(v)} - \overline{X}^{(v)})^2}} \tag{8}$$

Only the top-K strongest correlations for each node are retained to form the weighted adjacency matrix A. Formally, the adjacency matrix A is defined as:

$$A_{uv} = \begin{cases} \rho_{uv}, & \text{if } u \in \mathcal{N}^K(v) \text{ or } v \in \mathcal{N}^K(u) \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

The normalized and rescaled Laplacian $\tilde{L}$ is then derived following standard graph convolution practices. Eq. (8) defines the Pearson correlation used to construct the spatial graph, ensuring that highly correlated stations remain directly connected. In this study, each PAMS monitoring station is treated as one graph node, and the edge weight between two stations is computed from the Pearson correlation coefficient of their pollutant concentration time series using the training set only. Constructing the graph from the training set prevents information leakage from the validation and testing periods. For each node, only the top-K strongest correlations are retained to form a sparse adjacency matrix, where K is fixed across all experiments.

This top-K correlation filtering removes weak or noisy station-to-station connections and provides a stable graph support for subsequent spatial convolution. Since Pearson correlation is symmetric, the spatial dependency is modeled as an undirected graph. After top-K filtering, the adjacency matrix is symmetrized before graph normalization. The normalized and rescaled Laplacian derived from this symmetric adjacency matrix is then used as the graph support for ChebGCN. The static Pearson-correlation graph is designed to represent stable inter-station dependency rather than instantaneous wind-driven transport. Stations affected by similar emission sources, regional transport pathways, or shared meteorological regimes often exhibit persistent statistical dependencies over long observation periods. This graph therefore provides a computationally efficient structural prior for modeling spatial dependency among monitoring stations.

(3) ChebGCN Module

ChebGCN is used to model spatial pollutant interactions because pollutant transport in an industrial monitoring network is not limited to immediate neighboring stations. Under regional transport, land–sea breeze circulation, and shared emission influences, a station may be affected by multi-hop dependencies from non-adjacent monitoring sites. Chebyshev polynomial graph convolution enables each node to aggregate information from multi-hop neighborhoods through a K-order graph filter without stacking many graph convolution layers. This design is suitable for capturing non-local inter-station dependencies while avoiding excessive model depth. In the proposed framework, ChebGCN therefore complements TemporalGLU by propagating temporally filtered features across the monitoring network, allowing FFTGNet to jointly model local temporal dynamics and regional spatial transport. To capture spatial correlations among different monitoring stations, this study introduces a graph convolution module based on Chebyshev polynomial approximation (ChebGCN) following the temporal convolution stage. This module utilizes the scaled graph Laplacian $\tilde{L}$ to perform multi hop neighborhood feature diffusion, enabling the model to learn spatial interactions in pollutant propagation. At each time step t, the node feature matrix $U_t \in \mathbb{R}^{V\times C}$ is convolved with a Chebyshev polynomial based graph filter of order K, defined as:

$$H_t = \sum_{k=0}^{K-1} T_k(\tilde{L}) \cdot U_t \cdot \Theta_k \tag{10}$$

where $T_k(\cdot)$ denotes the Chebyshev polynomial of degree K. Eq. (9) shows that the Chebyshev graph convolution performs K-order polynomial filtering on the scaled Laplacian $\tilde{L}$, allowing each node to aggregate information from multi-hop neighborhoods and capture non-local spatial dependencies.

4 Experiments and Results

4.1 Experimental Environment and Setting

To validate the effectiveness of the proposed FFTGNet (Fast Fourier Transform Temporal Graph Network) model for spatiotemporal forecasting of multiple air pollutants, a series of experiments are conducted. This section outlines the overall experimental design and environment settings, including data sources, preprocessing procedures, Positive Matrix Factorization (PMF) feature decomposition, model training pipeline, and evaluation metrics. The experiments are conducted on real world environmental monitoring data to assess the stability and accuracy of the proposed model across various pollutants and temporal resolutions through multi step and cross station forecasting tasks. In this study, Yunlin and its surrounding regions are selected as the research area, primarily due to their proximity to the Mailiao petrochemical industrial zone. This area emits large quantities of volatile organic compounds (VOCs) and nitrogen oxides (NOx), which are key precursors in ozone formation.

This region is influenced by land sea breeze circulation and strong solar radiation, often resulting in high ozone concentrations observed during the afternoon. In addition, the availability of complete multi station PAMS (Photochemical Assessment Monitoring Stations) data allows for effective analysis of the interactions between industrial emissions and meteorological factors in shaping ozone formation dynamics. The experimental dataset covers ten PAMS (Photochemical Assessment Monitoring Stations) located in Yunlin and its surrounding areas, with hourly measurements collected from 2020 to 2024. The monitored variables include 54 species of volatile organic compounds (VOCs), nitrogen oxides (NOx), ozone, CO, fine particulate matter ($PM_{1.0}$), and various meteorological parameters. The data are chronologically divided into 60% for training, 20% for validation, and 20% for testing. All features are standardized using z score normalization based on the statistics of the training set. A sliding window of 24 h is applied to construct input samples, which are used to predict pollutant concentrations for the subsequent hour.

During the preprocessing stage, Positive Matrix Factorization (PMF) is applied to decompose the VOCs data into six representative source related feature groups, namely Aromatic and Aliphatic Mixtures (AAM), Combustion Mixtures (CM), Industrial Coatings (IC), Petroleum Processing (PP), Solvent Usage (SU), and Vehicular Emission (VE). These source specific feature vectors are incorporated into the input of FFTGNet to provide auxiliary information regarding pollutant sources. The model is trained using the Smooth L1 loss function (also known as the Huber loss), which balances training stability and robustness to outliers. The Smooth L1 loss is adopted to ensure numerical stability while maintaining robustness against outliers, which are common in real world pollutant concentration data due to occasional spikes or abrupt events. The optimizer is AdamW with a learning rate of 10^{-3} and a weight decay of 10^{-4}. A cosine annealing schedule is employed to dynamically adjust the learning rate during training. At the end of each training epoch, the FFT module re estimates the primary period, which is then smoothed using an exponential moving average (EMA) update. All training is conducted on an NVIDIA RTX 4090 GPU with a batch size of 32 for 200 epochs. Early stopping is applied based on the root mean squared error (RMSE) on the validation set. During training, the performance metrics on both the training and validation sets are monitored to track convergence and prevent overfitting. Upon completion of training, the final model is evaluated on a separate test set to assess its generalization performance. Three primary evaluation metrics are monitored: Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). These metrics respectively quantify the model's goodness of fit, the magnitude of prediction errors, and the average deviation. The formal definitions of these metrics are provided as following.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{12}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\, y_i - \hat{y}_i \,| \tag{13}$$

In addition, five baseline models are implemented for comparison, including Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), a hybrid CNN–LSTM model, and Spatiotemporal Graph Convolutional Network (STGCN) [37]. For the recently published spatiotemporal baselines (MGCGRU-SAN [38] and AirNN-Adapted [39]), adapted implementations were developed under the same experimental setting. All models were trained and evaluated using the same training/validation/testing split, input features, forecasting horizons, evaluation metrics, batch size, maximum training epochs, and early-stopping strategy to ensure fairness and reproducibility. Model-specific architectural hyperparameters were preserved according to their original formulations whenever possible.

4.2 Model Performance Comparison across Pollutants and Forecast Horizons

Evaluation metrics include R^2, RMSE, and MAE. All results were computed using the same data split and feature set to ensure consistency. As shown in Table 2, a consistent trend was observed across all models: as the forecasting horizon increased, R^2 decreased while RMSE and MAE increased, indicating the cumulative effect of prediction errors over longer forecasting steps.

Table 2: Comparison with recent SOTA and baseline models for O_3 forecasting.

Prediction Horizon	**1 h**			**2 h**			**3 h**			**6 h**		
Model	**R^2**	**RMSE**	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**	**RMSE**	**MAE**
ANN	0.6860	8.3990	6.1770	0.6000	9.4780	7.0100	0.5520	10.0340	7.5290	0.4567	11.0519	8.4118
CNN	0.7790	7.0430	4.7710	0.7460	7.5620	5.6620	0.6930	8.3080	6.3100	0.5744	9.7847	7.4388
CNN–LSTM	0.8470	5.8620	4.2340	0.7720	7.1650	5.2990	0.6990	8.2190	6.1210	0.5632	9.9090	7.4819
LSTM	0.8551	5.7057	4.1121	0.8097	6.5377	4.7556	0.7599	7.3437	5.3759	0.6549	8.8050	6.6298
GRU	0.9034	4.6803	3.4837	0.7748	7.1458	5.3732	0.6701	8.6498	6.6039	0.3032	12.5736	9.4209
ST-GCN	0.9223	4.1975	3.0463	0.8032	6.6809	4.8669	0.6764	8.5665	6.2966	0.3247	12.3772	9.1641
MGCGRU-SAN [38]	0.9334	3.8847	2.8078	0.8431	5.9655	4.4417	0.7809	7.0498	5.2812	0.6650	8.7183	6.5959
AirNN-Adapted [39]	0.9094	4.5320	3.4197	0.7825	7.0229	5.3253	0.7324	7.7902	5.9476	0.5538	10.0613	7.6991
FFTGNet	0.9509	3.3369	2.3875	0.8757	5.3099	3.9157	0.8116	6.5362	4.9044	0.6797	8.5239	6.5079

Note: RMSE, MAE’s units is ppb for O_3, mg/m^3 for PM_{10}, ppm for CO.

To provide a stronger comparison with recent spatiotemporal forecasting methods, we further include two recent graph-based air-quality forecasting models: MGCGRU-SAN [38] and AirNN-Adapted [39]. MGCGRU-SAN integrates mixed graph convolutional GRU and self-attention mechanisms to capture multi-station temporal dependencies, while AirNN introduces an uncertainty-aware graph convolution recurrent framework for air-quality forecasting. Since these models were originally designed for different air-quality forecasting settings, we implemented adapted versions under the same data partition, input sequence length, forecasting horizons, and evaluation metrics to ensure a fair comparison. Tables 2–4 compare FFTGNet with conventional baselines and recent spatiotemporal forecasting models for O_3, CO, and PM_{10}, respectively. Overall, FFTGNet achieves the most consistent performance across pollutants and forecasting horizons. As the prediction horizon increases from 1 to 6 h, all models show a general degradation trend, with lower R^2 and higher RMSE/MAE. This trend indicates the increasing uncertainty and accumulated forecasting errors associated with longer prediction horizons.

For O_3 forecasting, Table 2 shows that FFTGNet achieves the best overall performance across the 1, 2, 3, and 6 h horizons. Compared with conventional temporal models and recent spatiotemporal baselines, FFTGNet maintains higher explanatory ability and lower prediction errors, particularly at longer horizons. This result is consistent with the strong diurnal behavior of O_3, where the proposed FFT-guided period folding mechanism helps align recurrent temporal patterns before spatiotemporal graph learning.

For CO forecasting, Table 3 shows that FFTGNet also provides the best overall performance across the evaluated horizons. CO concentrations are often influenced by localized combustion, traffic-related emissions, and short-term spikes. The superior performance of FFTGNet suggests that the combination of TemporalGLU and graph-based spatial propagation is effective for modeling both abrupt temporal fluctuations and inter-station dependencies.

Table 3: Comparison with recent SOTA and baseline models for CO forecasting.

Prediction Horizon	1 h			2 h			3 h			6 h		
Model	R^2	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**
ANN	0.5100	0.0660	0.0460	0.4010	0.0720	0.0520	0.3290	0.0770	0.0550	0.1958	0.0839	0.0625
CNN	0.6890	0.0520	0.0370	0.5590	0.0620	0.0460	0.4690	0.0680	0.0500	0.3226	0.0775	0.0569
CNN–LSTM	0.6610	0.0550	0.0380	0.5470	0.0630	0.0450	0.4690	0.0680	0.0500	0.4057	0.0721	0.0531
LSTM	0.6519	0.0558	0.0385	0.5723	0.0619	0.0437	0.5118	0.0661	0.0473	0.3876	0.0740	0.0550
GRU	0.8280	0.0396	0.0277	0.6580	0.0558	0.0398	0.5394	0.0648	0.0475	0.3117	0.0792	0.0595
ST-GCN	0.8238	0.0401	0.0264	0.6820	0.0538	0.0361	0.5287	0.0655	0.0452	0.4285	0.0721	0.0527
MGCGRU-SAN [38]	0.8106	0.0416	0.0291	0.6805	0.0540	0.0383	0.5822	0.0617	0.0445	0.4448	0.0711	0.0521
AirNN-Adapted [39]	0.8070	0.0420	0.0311	0.6594	0.0557	0.0422	0.5704	0.0626	0.0476	0.4172	0.0729	0.0567
FFTGNet	0.8574	0.0361	0.0236	0.7372	0.0489	0.0331	0.6601	0.0557	0.0387	0.5157	0.0664	0.0491

Table 4: Comparison with recent SOTA and baseline models for PM_{10} forecasting.

Prediction Horizon	1 h			2 h			3 h			6 h		
Model	R^2	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**
ANN	0.3580	24.3980	12.4220	0.2840	25.7660	13.3050	0.2540	26.3080	14.0200	0.1722	27.7114	15.3490
CNN	0.5900	19.4960	9.3980	0.4550	22.4870	11.7200	0.4160	23.2770	11.8780	0.2194	26.9096	14.4461
CNN–LSTM	0.6000	19.2570	9.4910	0.5040	21.4480	10.9450	0.4100	23.3870	12.2980	0.2920	25.6280	14.2991
LSTM	0.5737	19.8652	9.9612	0.5085	21.3300	11.0871	0.4448	22.6671	12.0371	0.3374	24.7571	13.8524
GRU	0.8120	10.1748	6.7813	0.6818	13.2378	8.8287	0.5950	14.9353	9.9151	0.3723	18.5947	12.5981
ST-GCN	0.8285	9.7164	6.1891	0.7034	12.7798	8.2445	0.6038	14.7722	9.6298	0.3946	18.2613	12.0303
MGCGRU-SAN [38]	0.8236	9.8559	6.4377	0.7138	12.5547	8.3570	0.6243	14.3838	9.4483	0.4799	16.9270	11.3688
AirNN-Adapted [39]	0.8191	9.9809	6.6972	0.7096	12.6463	8.6260	0.6023	14.7990	9.8969	0.4155	17.9436	11.8601
FFTGNet	0.8556	8.9168	5.7005	0.7594	11.5114	7.5065	0.6752	13.3738	8.7594	0.5065	16.4880	11.4009

For PM_{10} forecasting, Table 4 indicates that FFTGNet achieves the highest R^2 and the lowest RMSE across all forecasting horizons. In terms of MAE, FFTGNet obtains the best results at the 1, 2, and 3 h horizons, while MGCGRU-SAN shows a slightly lower MAE at the 6 h horizon. This suggests that FFTGNet remains highly competitive for particulate matter forecasting, especially because PM_{10} is affected by both meteorological accumulation and regional transport.

Figs. 3–5 further visualize the comparative trends of R^2, RMSE, and MAE. Fig. 3 shows that FFTGNet maintains strong explanatory capability across pollutants. Fig. 4 indicates that FFTGNet generally yields lower RMSE, reflecting reduced large-error deviations. Fig. 5 shows that FFTGNet also maintains stable MAE performance, suggesting robust average prediction accuracy. Since RMSE and MAE have different physical units and concentration scales for O_3, CO, and PM_{10}, error magnitudes should be interpreted within each pollutant rather than directly compared across pollutants. These results demonstrate that the performance gain of FFTGNet is not limited to a single pollutant or a single forecasting horizon. The improvement can be attributed to the integrated design of FFT-guided period folding, TemporalGLU, and ChebGCN. The FFT-guided module provides frequency-aware temporal alignment, TemporalGLU captures nonlinear temporal fluctuations, and ChebGCN models inter-station spatial dependencies. Together, these components enable

FFTGNet to capture periodic temporal behavior, localized emission fluctuations, and regional pollutant transport within a unified forecasting framework.

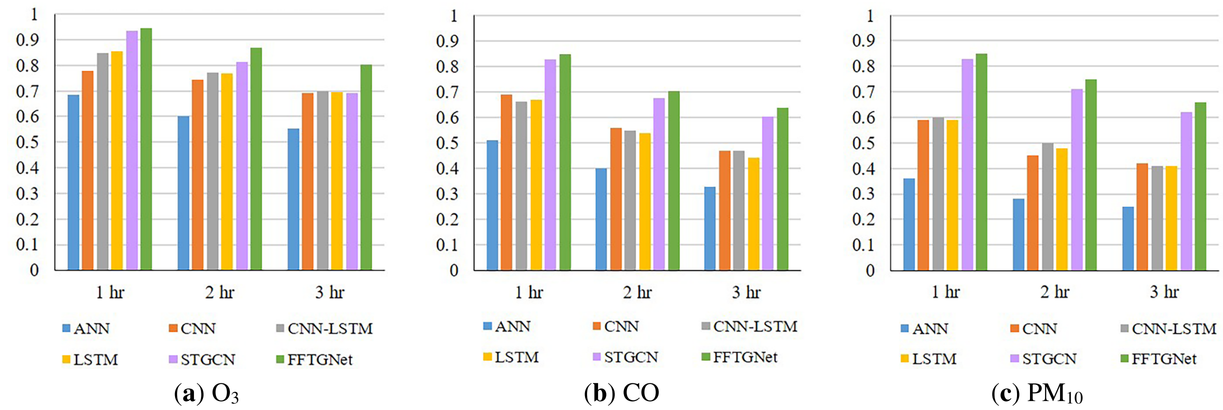


Figure 3: R^2 comparison across pollutants and forecast horizons: (**a**) O_3, (**b**) CO, and (**c**) PM_{10}.

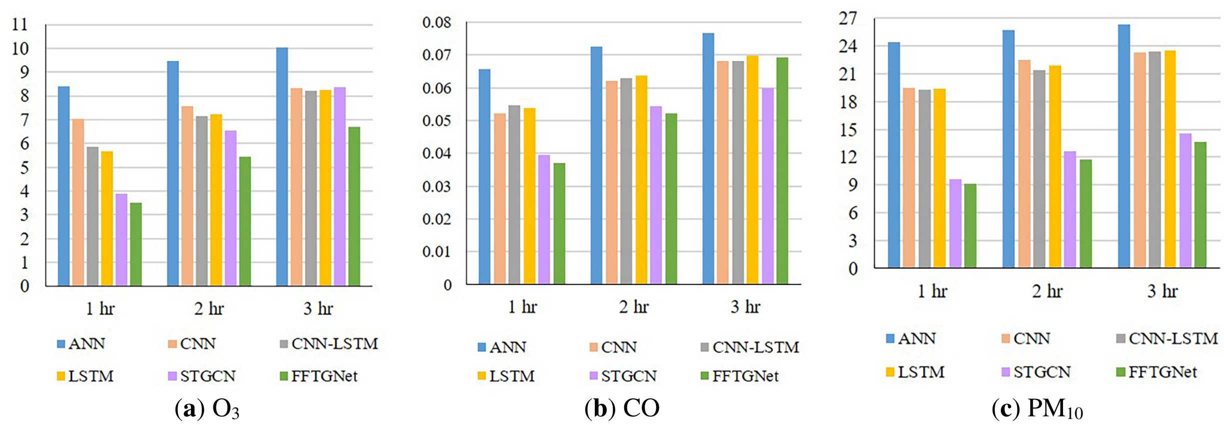


Figure 4: RMSE comparison across pollutants and forecast horizons: (**a**) O_3, (**b**) CO, and (**c**) PM_{10}.

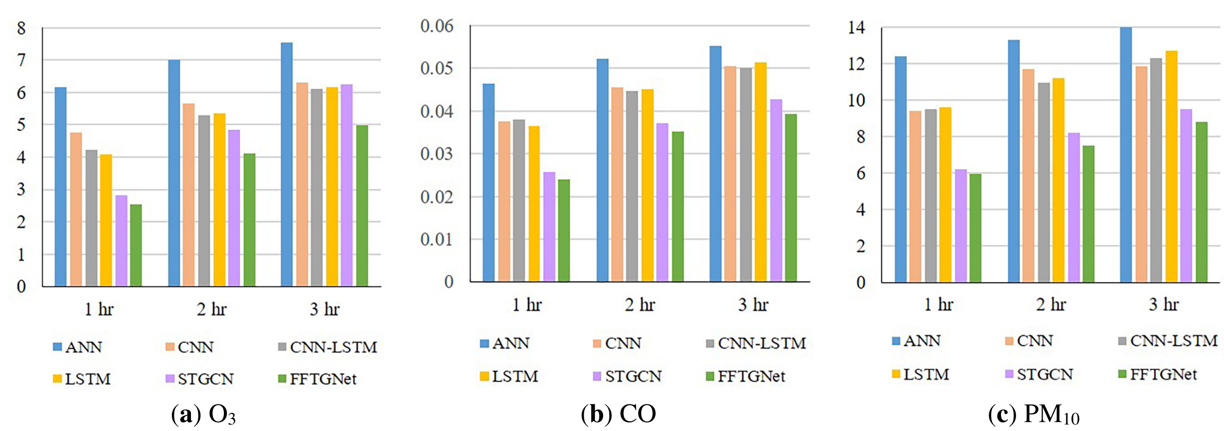
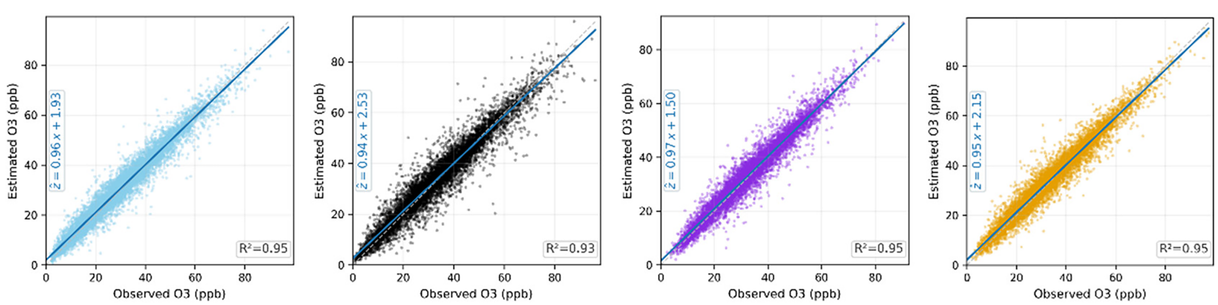


Figure 5: MAE comparison across pollutants and forecast horizons: (**a**) O_3, (**b**) CO, and (**c**) PM_{10}.

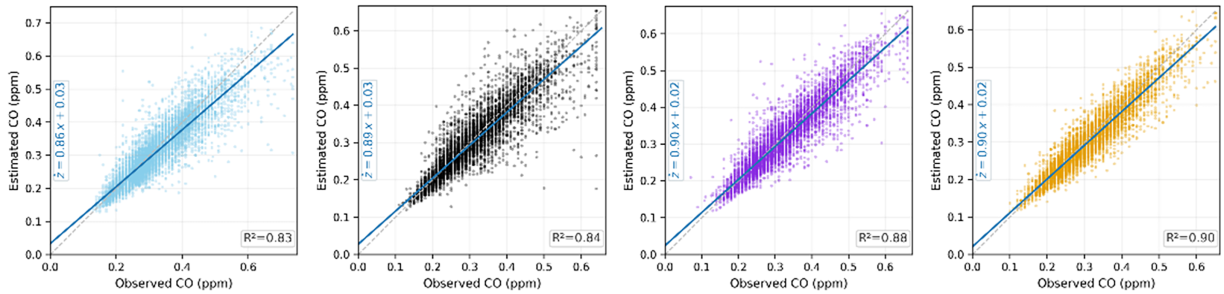
The MAE results reflect the stability of each model in terms of average prediction error as shown in Fig. 5. Compared with RMSE, MAE grows more moderately across horizons, reflecting its lower sensitivity to extreme deviations. This is consistent with the nature of MAE, which is less sensitive to outliers and better captures the general average deviation. In the comparison across pollutants (as shown in Fig. 5), FFTGNet achieves the lowest or joint lowest MAE for O_3, indicating its ability to effectively control average deviation under periodic conditions. For CO, FFTGNet consistently maintains the lowest MAE across the 1, 2, and 3 h forecasting horizons. Even in the 3 h forecast where STGCN occasionally outperforms in RMSE, FFTGNet still yields lower average deviation, suggesting that its dynamic gating mechanism helps suppress the impact of short term spikes. For PM_{10}, the MAE results show that FFTGNet and STGCN perform comparably overall, with FFTGNet achieving slightly lower values in most scenarios. Overall, FFTGNet demonstrates consistent superiority in MAE, indicating its strong capability to maintain low average prediction errors and robust performance across different pollutants and forecasting horizons. Considering the combined results of R^2, RMSE, and MAE, FFTGNet consistently delivers stable and leading performance across different pollutants and forecasting horizons. These outcomes indicate the model's strong adaptability to periodic patterns, localized behaviors, and cross station pollutant transport. To ensure a fair interpretation of error magnitudes, it is important to recognize that the absolute scales of RMSE and MAE differ substantially across pollutants due to their distinct measurement units and typical concentration ranges. For example, O_3 and PM_{10} are usually measured in μg/m^3, whereas CO is expressed in ppm or mg/m^3. These scale differences do not indicate inconsistencies in model performance but simply reflect the physical nature of the data. Therefore, comparisons of RMSE and MAE should be made within the same pollutant category, while cross pollutant performance should be evaluated primarily through normalized or relative metrics. In other words, the metric scaling difference does not imply model inconsistency, but rather arises from differences in measurement definitions and concentration variability.

To further illustrate the prediction accuracy and spatial consistency of FFTGNet, the Fig. 6 presents the one hour ahead prediction–observation comparisons for PM_{10}, O_3, and CO across representative monitoring stations. Each subplot corresponds to a specific site, showing the normalized (0–1) target vs. predicted values with fitted regression lines and correlation coefficient. The results demonstrate that FFTGNet achieves stable and high correlations across all pollutants and stations. Ozone exhibits the strongest linearity, reflecting its pronounced diurnal periodicity captured by the frequency guided mechanism. Carbon monoxide shows slightly lower correlations due to localized emission spikes, while PM_{10} attains intermediate performance, representing the combined effects of meteorological variability and spatial transport. These visual comparisons confirm that FFTGNet maintains consistent prediction quality across both pollutants and spatially distributed monitoring sites.

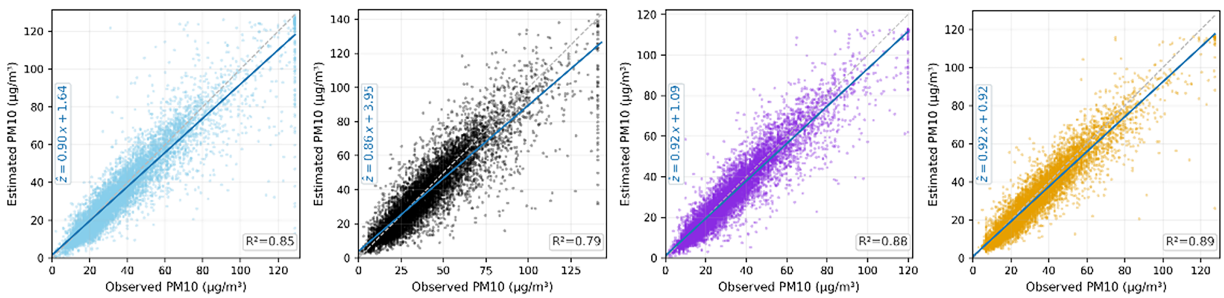


(**a**) O_3 prediction–observation comparison

Figure 6: (Continued)



(**b**) CO prediction–observation comparison



(**c**) PM_{10} prediction–observation comparison

Figure 6: Prediction–observation comparisons for 1-h-ahead forecasts using FFTGNet across four representative monitoring stations: (**a**) O_3, (**b**) CO, and (**c**) PM_{10}. Each subplot shows observed vs. predicted concentrations with the fitted regression line and coefficient of determination (R^2).

4.3 Extended-Horizon, Ablation, and Practical Discussion

4.3.1 Extended-Horizon Forecasting Analysis

It has been extended the forecasting horizons from the original 1–3 h setting to 6, 12, and 24 h to provide a more rigorous evaluation of extended-horizon forecasting stability. Table 5 reports the performance of FFTGNet for O_3, CO, and PM_{10} across 1, 2, 3, 6, 12, and 24 h horizons. As expected, the forecasting difficulty increases as the prediction horizon becomes longer. For O_3, the R^2 decreases from 0.9509 at 1 h to 0.6797 at 6 h, 0.5838 at 12 h, and 0.5576 at 24 h, while RMSE gradually increases from 3.3369 to 10.0174. Similar degradation trends are observed for CO and PM_{10}. For CO, R^2 decreases from 0.8574 at 1 h to 0.3168 at 24 h. For PM_{10}, R^2 decreases from 0.8556 to 0.3505 over the same horizon range. These results confirm that long-horizon air-quality forecasting is substantially more challenging due to accumulated prediction uncertainty, meteorological variability, and nonstationary pollutant transport. Nevertheless, FFTGNet shows a gradual degradation trend rather than an abrupt collapse across increasing horizons. For PM_{10}, R^2 decreases from 0.8556 to 0.3505 over the same horizon range. This suggests that the proposed frequency-guided period folding mechanism may provide a useful temporal alignment prior for mitigating performance degradation in short- to medium-term prediction. In particular, O_3 shows relatively better long-horizon performance than CO and PM_{10}, which is consistent with its stronger diurnal periodicity associated with photochemical reactions and solar-radiation-driven processes. These results also clarify the scope and future applicability of the proposed method. FFTGNet is primarily designed for short- to medium-term and extended-horizon monitoring support in industrial air-quality systems, where recurrent temporal patterns and inter-station dependencies remain informative. Compared with emission-response prediction studies that focus on policy-level assessment of air-quality changes under emission-control scenarios, the present work focuses on monitoring-data-driven multi-pollutant forecasting using PAMS observations. Compared with conventional

ANN-based pollutant prediction models, FFTGNet further integrates FFT-guided temporal restructuring and graph-based inter-station dependency learning. Although the 12 and 24 h results demonstrate the multi-step forecasting applicability of FFTGNet, forecasting at substantially extended horizons remains challenging. Predictions beyond 24 h or toward 72 h may require additional information, such as numerical weather forecasts, emission inventories, dynamic wind-informed graphs, and uncertainty-aware forecasting modules. Therefore, future work will further enhance the proposed framework by incorporating these auxiliary sources and adaptive modeling strategies for more challenging environmental forecasting scenarios.

Table 5: Extended-horizon forecasting performance of FFTGNet from 1 to 24 h.

Pollutant	Metric	1 h	2 h	3 h	6 h	12 h	24 h
O_3	Test R^2	0.9509	0.8757	0.8116	0.6797	0.5838	0.5576
	Test RMSE	3.3369	5.3099	6.5362	8.5239	9.7192	10.0174
	Test MAE	2.3875	3.9157	4.9044	6.5079	7.4234	7.5700
CO	Test R^2	0.8574	0.7372	0.6601	0.5157	0.3620	0.3168
	Test RMSE	0.0361	0.0489	0.0557	0.0664	0.0762	0.0789
	Test MAE	0.0236	0.0331	0.0387	0.0491	0.0563	0.0591
PM_{10}	Test R^2	0.8556	0.7594	0.6752	0.5065	0.3735	0.3505
	Test RMSE	8.9168	11.5114	13.3738	16.4880	18.5808	18.9231
	Test MAE	5.7005	7.5065	8.7594	11.4009	12.7282	13.0339

4.3.2 Ablation Study

Tables 6–8 present the ablation results of FFTGNet for O_3, CO, and PM_{10} forecasting, respectively. Three ablated variants are evaluated: w/o FFT, w/o ChebGCN, and w/o TemporalGLU. The purpose of this analysis is to examine whether the frequency-guided period folding module, the spatial graph convolution module, and the gated temporal modeling module each contribute to the overall forecasting performance. As shown in Table 6, removing any core component leads to performance degradation in O_3 forecasting. The full FFTGNet model consistently achieves the best results across all horizons, with R^2 values of 0.9509, 0.8757, 0.8116, and 0.6797 at the 1, 2, 3, and 6 h horizons, respectively. In particular, at the 6 h horizon, removing FFT reduces R^2 from 0.6797 to 0.3058, while removing TemporalGLU further reduces R^2 to 0.2622. This indicates that frequency-guided temporal alignment and gated temporal modeling are especially important for maintaining O_3 forecasting stability under extended horizons. For CO forecasting,

Table 6: Ablation study of FFTGNet on O_3 forecasting across 1, 2, 3, and 6 h horizons.

Variant	Test R^2				Test RMSE				Test MAE			
Prediction Horizon	**1 h**	**2 h**	**3 h**	**6 h**	**1 h**	**2 h**	**3 h**	**6 h**	**1 h**	**2 h**	**3 h**	**6 h**
Full model	0.9509	0.8757	0.8116	0.6797	3.3369	5.3099	6.5362	8.5239	2.3875	3.9157	4.9044	6.5079
w/o FFT	0.9253	0.8056	0.6776	0.3058	4.1169	6.6405	8.5505	12.5499	2.9615	4.9242	6.4027	9.4011
w/o ChebGCN	0.9308	0.8154	0.6843	0.3168	3.9623	6.4700	8.4615	12.4499	2.7891	4.7292	6.3355	9.4852
w/o TemporalGLU	0.9215	0.7874	0.6569	0.2622	4.2190	6.9441	8.8208	12.9376	3.0396	5.0752	6.6940	9.8771

Table 7: Ablation study of FFTGNet on CO forecasting across 1, 2, 3, and 6 h horizons.

Variant	Test R^2				Test RMSE				Test MAE			
Prediction Horizon	1 h	2 h	3 h	6 h	1 h	2 h	3 h	6 h	1 h	2 h	3 h	6 h
Full model	0.8574	0.7372	0.6601	0.5157	0.0361	0.0489	0.0557	0.0664	0.0236	0.0331	0.0387	0.0491
w/o FFT	0.8324	0.6920	0.6003	0.3959	0.0391	0.0530	0.0604	0.0742	0.0256	0.0361	0.0420	0.0522
w/o ChebGCN	0.8352	0.6938	0.5882	0.4500	0.0388	0.0528	0.0613	0.0708	0.0248	0.0371	0.0426	0.0505
w/o TemporalGLU	0.8295	0.6864	0.5687	0.3917	0.0394	0.0535	0.0627	0.0744	0.0263	0.0364	0.0466	0.0527

Table 8: Ablation study of FFTGNet on PM10 forecasting across 1, 2, 3, and 6 h horizons.

Variant	Test R^2				Test RMSE				Test MAE			
Prediction Horizon	1 h	2 h	3 h	6 h	1 h	2 h	3 h	6 h	1 h	2 h	3 h	6 h
Full model	0.8556	0.7594	0.6752	0.5065	8.9168	11.5114	13.3738	16.4880	5.7005	7.5065	8.7594	11.4009
w/o FFT	0.8324	0.7148	0.6135	0.4009	9.6066	12.5322	14.5903	18.1670	6.1685	8.1436	9.5553	12.0064
w/o ChebGCN	0.8310	0.7156	0.6242	0.4150	9.6451	12.5142	14.3851	17.9517	6.1168	8.2250	9.4669	11.9862
w/o TemporalGLU	0.8325	0.7119	0.6212	0.3924	9.6028	12.5946	14.4424	18.2956	6.1619	8.1688	9.7013	12.0628

Table 7 shows a similar trend. The full model achieves the highest R^2 and the lowest RMSE/MAE across all evaluated horizons. Since CO concentration is often affected by localized emission spikes and short-term fluctuations, the degradation observed after removing TemporalGLU confirms the importance of gated temporal filtering. At the 6 h horizon, the R^2 decreases from 0.5157 in the full model to 0.3959, 0.4500, and 0.3917 when FFT, ChebGCN, and TemporalGLU are removed, respectively. These results suggest that both temporal gating and spatial propagation are beneficial for modeling CO dynamics. For PM_{10} forecasting,

Table 8 further verifies the contribution of each module. The full model achieves the best R^2 across all horizons, reaching 0.8556, 0.7594, 0.6752, and 0.5065 from 1 to 6 h. Removing FFT, ChebGCN, or TemporalGLU consistently increases RMSE and MAE, particularly at the 6 h horizon. This demonstrates that PM_{10} prediction benefits not only from temporal modeling but also from graph-based inter-station dependency learning, since particulate matter is strongly influenced by regional transport and meteorological accumulation. Overall, the ablation study confirms that the three major components of FFTGNet are complementary. The FFT-guided period folding module provides a frequency-aware temporal alignment prior, TemporalGLU adaptively filters nonlinear temporal fluctuations and short-term spikes, and ChebGCN captures inter-station spatial dependencies. The consistent degradation observed in Tables 6–8 after removing each module verifies that the performance improvement of FFTGNet is not caused by a single component alone, but by the integrated design of frequency-guided temporal alignment, gated temporal modeling, and graph-based spatial propagation.

4.3.3 Computational Complexity and Deployment Feasibility

As shown in Table 9, FFTGNet requires higher computational cost than lightweight baselines such as ANN, GRU, and LSTM. This is expected because the proposed model integrates FFT-guided period folding, TemporalGLU, and ChebGCN to jointly capture frequency-aware temporal patterns and inter-station spatial dependencies. Compared with ST-GCN, FFTGNet increases the number of parameters and FLOPs, but its practical inference latency remains at 71.70 ms per sample, which is still far below the hourly sampling interval of the monitoring data. Therefore, although FFTGNet is not the lightest model, it remains feasible for server-side near-real-time air-quality forecasting. The training convergence results also show that FFTGNet reaches its best validation performance at epoch 16, requiring approximately 37.54 min to reach the best epoch under the experimental hardware setting. This training cost is higher than that of simpler baselines but remains acceptable because model training is performed offline. More importantly, the FFT-based dominant period re-estimation is conducted at the training epoch level and smoothed by EMA. During deployment, the learned dominant period can be fixed or updated offline periodically; therefore, the FFT period update does not need to be performed for every online inference sample and does not become a real-time inference bottleneck.

Table 9: Computational complexity and deployment feasibility comparison.

Model	Params M	FLOPs M	Inference ms/Sample	Best Epoch	Avg. Epoch Time s	Time to Best Min
ANN	3.965	7.914	0.17	17	13.51	3.81
CNN	0.871	291.595	0.36	8	56.04	7.48
CNN–LSTM	0.147	31.869	0.49	19	22.74	7.20
LSTM	0.243	80.812	1.28	10	51.75	8.63
GRU	0.033	0.009	0.19	19	0.90	0.28
ST-GCN	0.098	333.743	41.44	12	123.14	23.49
MGCGRU-SAN	0.027	3.634	6.46	17	7.59	2.15
AirNN-Adapted	0.006	12.491	105.26	17	49.67	14.07
FFTGNet	1.372	4624.700	71.70	16	140.71	37.54

FLOPs were estimated using the same profiler and fixed input shape for all models. Since some recurrent or graph operations may not be fully captured by FLOP-counting tools, inference latency is also reported as a practical runtime indicator.

5 Conclusion

This study proposes FFTGNet (Fast Fourier Transform Temporal Graph Network), a frequency-aware spatiotemporal graph neural network for multi-pollutant forecasting in industrial air-quality systems. The proposed framework integrates an FFT-guided dominant-period estimation and period-folding module with a temporal-to-spatial graph backbone composed of TemporalGLU and Chebyshev graph convolution. Rather than using FFT as a standalone spectral feature extractor, FFTGNet uses the estimated dominant quasi-periodic pattern as a temporal restructuring prior to form intra-period and inter-period representations. The folded representation is further fused with the original temporal sequence, enabling TemporalGLU to adaptively filter nonlinear temporal fluctuations and short-term pollutant spikes, while ChebGCN propagates information across inter-station spatial dependencies. This integrated design allows the model to jointly capture periodic temporal behavior, localized pollutant variations, and spatial transport patterns under nonstationary industrial air-quality conditions. Experiments using five years of hourly observations from ten Photochemical Assessment Monitoring Stations in Yunlin, Taiwan, covering O_3, CO, PM_{10}, meteorological variables, and PMF-derived VOC source features, demonstrate that FFTGNet provides the most consistent forecasting performance across pollutants and prediction horizons. Compared with conventional baselines and recent spatiotemporal models, including MGCGRU-SAN and AirNN-Adapted, FFTGNet achieves R^2 values of 0.9509, 0.8574, and 0.8556 for 1 h forecasting of O_3, CO, and PM_{10}, respectively. For O_3 forecasting, the model further maintains R^2 values of 0.8757, 0.8116, and 0.6797 at the 2, 3, and 6 h horizons, while consistently reducing RMSE and MAE compared with recent baselines. For CO, FFTGNet effectively captures localized emission fluctuations and short-term spikes, whereas for PM_{10} it achieves the highest R^2 and lowest RMSE across the evaluated 1–6 h horizons, with only a slight exception in MAE at the 6 h horizon.

These results indicate that the proposed frequency-aware spatiotemporal design improves both predictive accuracy and robustness across heterogeneous pollutant dynamics.

Extended-horizon evaluation further shows that FFTGNet exhibits smooth and progressive performance degradation rather than abrupt collapse when the forecasting horizon is extended to 12 and 24 h. Although prediction uncertainty inevitably increases at longer horizons, FFTGNet shows a gradual degradation pattern rather than an abrupt failure. However, the 24 h results indicate that the predictive capability becomes limited for CO and PM_{10}, with R^2 values of 0.3168 and 0.3505, respectively. In contrast, O_3 retains relatively better long-horizon performance, which may be attributed to its stronger diurnal photochemical behavior and better alignment with the FFT-guided period-folding mechanism. Therefore, FFTGNet is more appropriate for short- to medium-term forecasting, while its use for longer-horizon prediction should be interpreted cautiously, especially for pollutants dominated by localized emissions or meteorological variability. The ablation results further confirm the complementary roles of the three major components. Removing the FFT-guided period-folding module, TemporalGLU, or ChebGCN consistently degrades forecasting performance across O_3, CO, and PM_{10}, demonstrating that the improvement of FFTGNet is not caused by a single module alone, but by the integration of frequency-guided temporal alignment, gated temporal modeling, and graph-based spatial propagation. Computational analysis also indicates that although FFTGNet is more complex than lightweight baselines, its inference latency remains 71.70 ms per sample, which is far below the hourly sampling interval of the monitoring data and therefore remains feasible for server-side near-real-time forecasting. Overall, this study demonstrates that coupling frequency-domain temporal priors with spatiotemporal graph learning provides an effective and practical approach for modeling complex multi-pollutant dynamics in industrial air-quality monitoring networks.

Limitations: Despite the promising results, several limitations should be acknowledged. First, the proposed FFT-guided period folding mechanism mainly captures dominant quasi-periodic patterns and may become less representative under abrupt environmental events, such as dust storms, extreme weather fluctuations, or sudden emission changes. Second, the current static Pearson-correlation graph cannot fully model time-varying and directional pollutant transport caused by seasonal wind shifts or upwind/downwind reversals. Third, although FFTGNet demonstrates stable performance for 12 and 24 h prediction, its extension to substantially longer lead times, such as t + 72, remains challenging, particularly for ozone (O_3), whose concentration is strongly influenced by nonlinear photochemical reactions, solar radiation, precursor emissions, atmospheric transport, and meteorological evolution. Reliable 72 h forecasting would generally require additional information such as numerical weather forecasts, emission inventories, and dynamic wind-informed graph modeling. Therefore, future work will investigate adaptive time-frequency analysis, dynamic graph learning, and uncertainty-aware forecasting frameworks to improve robustness under highly nonstationary environmental conditions.

Acknowledgement: Not applicable.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Chia-Hui Liu; methodology, Chia-Hui Liu; formal analysis, Chia-Hui Liu; data curation, Chia-Hui Liu and Chen-Chuan Cheng; writing—original draft preparation, Chia-Hui Liu; writing—review and editing, Chia-Hui Liu and Chen-Chuan Cheng; validation, Chia-Hui Liu and Chen-Chuan Cheng; supervision, Chia-Hui Liu; project administration, Chia-Hui Liu. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The datasets used in this study are publicly available from the Taiwan Ministry of Environment (MOENV) Open Data Platform and the Air Quality Monitoring Network. The data can be accessed through the MOENV Open Data Platform at https://data.moenv.gov.tw/.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Garbagna L, Babu Saheer L, Maktab Dar Oghaz M. AI-driven approaches for air pollution modelling: a comprehensive systematic review. Environ Pollut. 2025;373(24):125937. doi:10.1016/j.envpol.2025.125937.
2. Guo Q, He Z, Wang Z. Assessing the effectiveness of long short-term memory and artificial neural network in predicting daily ozone concentrations in Liaocheng City. Sci Rep. 2025;15(1):6798. doi:10.1038/s41598-025-91329-w.
3. Agbehadji IE, Obagbuwa IC. Systematic review of machine learning and deep learning techniques for spatiotemporal air quality prediction. Atmosphere. 2024;15(11):1352. doi:10.3390/atmos15111352.
4. Zhang Y, Xue L, Carter WPL, Pei C, Chen T, Mu J, et al. Development of ozone reactivity scales for volatile organic compounds in a Chinese megacity. Atmos Chem Phys. 2021;21(14):11053–68. doi:10.5194/acp-21-11053-2021.
5. Chang KH. Modeling approach for emission reduction of O_3 precursors in southern Taiwan. Atmos Environ. 2008;42(28):6733–42. doi:10.1016/j.atmosenv.2008.05.037.
6. Reff A, Eberly SI, Bhave PV. Receptor modeling of ambient particulate matter data using positive matrix factorization: review of existing methods. J Air Waste Manag Assoc. 2007;57(2):146–54. doi:10.1080/10473289.2007.10465319.
7. Norris GA, Duvall R, Brown SG, Bai S. EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide. Washington, DC, USA: U.S. Environmental Protection Agency; 2014.
8. Liu PG, Yao YC, Tsai JH, Hsu YC, Chang LP, Chang KH. Source impacts by volatile organic compounds in an industrial city of southern Taiwan. Sci Total Environ. 2008;398(1–3):154–63. doi:10.1016/j.scitotenv.2008.02.053.
9. Hsu CY, Chiang HC, Shie RH, Ku CH, Lin TY, Chen MJ, et al. Ambient VOCs in residential areas near a large-scale petrochemical complex: spatiotemporal variation, source apportionment and health risk. Environ Pollut. 2018;240:95–104. doi:10.1016/j.envpol.2018.04.076.
10. Chen CH, Chuang YC, Hsieh CC, Lee CS. VOC characteristics and source apportionment at a PAMS site near an industrial complex in central Taiwan. Atmos Pollut Res. 2019;10(4):1060–74. doi:10.1016/j.apr.2019.01.014.
11. Huang YS, Hsieh CC. VOC characteristics and sources at nine photochemical assessment monitoring stations in western Taiwan. Atmos Environ. 2020;240(4):117741. doi:10.1016/j.atmosenv.2020.117741.
12. Yuan CS, Cheng WH, Huang HY. Spatiotemporal distribution characteristics and potential sources of VOCs at an industrial harbor city in southern Taiwan: three-year VOCs monitoring data analysis. J Environ Manage. 2022;303:114259. doi:10.1016/j.jenvman.2021.114259.
13. Cheng Y, Huang XF, Peng Y, Tang MX, Zhu B, Xia SY, et al. A novel machine learning method for evaluating the impact of emission sources on ozone formation. Environ Pollut. 2023;316(11):120685. doi:10.1016/j.envpol.2022.120685.
14. Elangasinghe MA, Singhal N, Dirks KN, Salmond JA. Development of an ANN–based air pollution forecasting system with explicit knowledge through sensitivity analysis. Atmos Pollut Res. 2014;5(4):696–708. doi:10.5094/apr.2014.079.
15. Chang JH, Griffith SM, Lin NH. Impacts of land-surface forcing on local meteorology and ozone concentrations in a heavily industrialized coastal urban area. Urban Clim. 2022;45:101257. doi:10.1016/j.uclim.2022.101257.
16. Lin GY, Lee YM, Tsai CJ, Lin CY. Spatial-temporal characterization of air pollutants using a hybrid deep learning/Kriging model incorporated with a weather normalization technique. Atmos Environ. 2022;289(25):119304. doi:10.1016/j.atmosenv.2022.119304.

17. Morapedi TD, Obagbuwa IC. Air pollution particulate matter (PM2.5) prediction in South African cities using machine learning techniques. Front Artif Intell. 2023;6:1230087. doi:10.3389/frai.2023.1230087.
18. Hwang H, Kim CH, Park JS, Park S, Kim JB, Lee JY. Augmentation of $PM_{1.0}$ measurements based on machine learning model and environmental factors. J Environ Sci. 2025;156(51):91–101. doi:10.1016/j.jes.2024.06.029.
19. Wang W, Mao F, Zou B, Guo J, Wu L, Pan Z, et al. Two-stage model for estimating the spatiotemporal distribution of hourly PM1.0 concentrations over central and East China. Sci Total Environ. 2019;675(2):658–66. doi:10.1016/j.scitotenv.2019.04.134.
20. Zhang JJY, Sun L, Rainham D, Dummer TJB, Wheeler AJ, Anastasopolos A, et al. Predicting intraurban airborne $PM_{1.0}$-trace elements in a port city: land use regression by ordinary least squares and a machine learning algorithm. Sci Total Environ. 2022;806(Pt 1):150149. doi:10.1016/j.scitotenv.2021.150149.
21. Liu X, Zhang X, Wang R, Liu Y, Hadiatullah H, Xu Y, et al. High-precision microscale particulate matter prediction in diverse environments using a long short-term memory neural network and street view imagery. Environ Sci Technol. 2024;58(8):3869–82. doi:10.1021/acs.est.3c06511.
22. Cichowicz R, Dobrzański M. Spatial analysis (measurements at heights of 10 and 20 m above ground level) of the concentrations of particulate matter (PM10, PM2. 5, and PM1.0) and gaseous pollutants (H_2S) on the university campus: a case study. Atmosphere. 2021;12(1):62. doi:10.3390/atmos12010062.
23. Li A, Li J, Shen Z. An efficient modern convolution-based dynamic spatiotemporal deep learning architecture for ozone prediction. Environ Model Softw. 2025;188(1):106424. doi:10.1016/j.envsoft.2025.106424.
24. Hanigan IC, Yu W, Yuen C, Gopi K, Knibbs LD, Cowie CT, et al. Deep ensemble machine learning with Bayesian blending improved accuracy and precision of modelled ground-level ozone for region with sparse monitoring: Australia, 2005–2018. Environ Model Softw. 2025;187(9):106378. doi:10.1016/j.envsoft.2025.106378.
25. Chen B, Zheng Q, Sun W, Yang G, Feng T, Wang Y. Geo-STO3Net: a deep neural network integrating geographical spatiotemporal information for surface ozone estimation. IEEE Trans Geosci Remote Sens. 2024;62:4102214. doi:10.1109/TGRS.2024.3358397.
26. Kafi F, Yousefi E, Ehteram M, Ashrafi K. Stabilized long short term memory (SLSTM) model: a new variant of the LSTM model for predicting ozone concentration data. Earth Sci Inform. 2025;18(3):311. doi:10.1007/s12145-025-01798-w.
27. Park J. Efficient ozone concentration trend prediction using ANN and K-means clustering. Earth Sci Inform. 2025;18(1):163. doi:10.1007/s12145-024-01676-x.
28. Räss S, Leuenberger MC. Analysis and prediction of atmospheric ozone concentrations using machine learning. Front Big Data. 2024;7:1469809. doi:10.3389/fdata.2024.1469809.
29. Jiménez-Navarro MJ, Martínez-Ballesteros M, Martínez-Álvarez F, Asencio-Cortés G. Explaining deep learning models for ozone pollution prediction via embedded feature selection. Appl Soft Comput. 2024;157(3):111504. doi:10.1016/j.asoc.2024.111504.
30. Cui Y, Liu B, Yang Y, Kang S, Wang F, Xu M, et al. Primary and oxidative source analyses of consumed VOCs in the atmosphere. J Hazard Mater. 2024;476:134894. doi:10.1016/j.jhazmat.2024.134894.
31. Chen HW, Chen CY, Chang TW, Lin GY. Spatial-temporal evaluation of $PM_{2.5}$ concentration for health risk reduction strategy development in a basin with different weather patterns. Atmos Pollut Res. 2023;14(10):101884. doi:10.1016/j.apr.2023.101884.
32. Prakash S, Sangeetha K. Systems classification of air pollutants using adam optimized CNN with XGBoost feature selection. Analog Integr Circuits Signal Process. 2025;122(3):35. doi:10.1007/s10470-025-02299-y.
33. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. Math Comput. 1965;19(90):297–301. doi:10.1090/S0025-5718-1965-0178586-1.
34. Dauphin YN, Fan A, Auli M, Grangier D. Language modeling with gated convolutional networks. In: Proceedings of the 34th International Conference on Machine Learning; 2017 Aug 6–11; Sydney, Australia. p. 933–41.
35. Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering. Adv Neural Inf Process Syst. 2016;29:3844–52. doi:10.5555/3157382.3157527.
36. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: Proceedings of the 5th International Conference on Learning Representations (ICLR); 2017 Apr 24–26; Toulon, France.

37. Yu B, Yin H, Zhu Z. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence; 2018 Jul 13–19; Stockholm, Sweden. p. 3634–40. doi:10.24963/ijcai.2018/505.
38. Zhao G, Yang X, Shi J, He H, Wang Q. A $PM_{2.5}$ spatiotemporal prediction model based on mixed graph convolutional GRU and self-attention network. Environ Pollut. 2025;368(5):125748. doi:10.1016/j.envpol.2025.125748.
39. Dong M, Sun Y, Jin Y, Song C, Zhang X, Luo X. Uncertainty graph convolution recurrent neural network for air quality forecasting. Adv Eng Inform. 2024;62(35):102651. doi:10.1016/j.aei.2024.102651.