Proof

Gorgan University of Agricultural Sciences and Natural Resources

Environmental Resources Research (ERR)

Print ISSN: 2783-4832
Online ISSN: 2783-4670

Daily PM2.5 Prediction and Pollution Episode Detection Using MLP Neural Networks Across Urban Monitoring Sites with Varied Land Uses

Article Info

Article type:
Research Article

Article history:
Received:
Accepted:

Corresponding author:

Keywords:
$PM_{2.5}$ concentration forecasting
artificial neural networks
air pollution
urban monitoring station
meteorological data

Abstract

This research presents a data-driven approach for forecasting daily $PM_{2.5}$ concentrations using a multilayer perceptron (MLP) neural network across three urban monitoring sites in Mashhad, Iran—Sajjad, Torogh, and Vila—each reflecting a unique land-use profile. The study utilized daily datasets collected from 2018 to 2023, and a dedicated MLP model was trained for each station. Various training algorithms were assessed to identify the most suitable configuration, with model complexity fine-tuned by adjusting the number of neurons in the hidden layer. Key input features included meteorological variables from the preceding day (such as wind speed, ambient temperature, precipitation, solar radiation, and relative humidity), the previous day's $PM_{2.5}$ concentration, and calendar-based temporal factors. To improve the network's predictive capability and prevent overfitting, data normalization and early stopping strategies were applied. The best predictive performance was recorded at the Sajjad station, where the model achieved an R^2 value of 0.79 and an MAE of 6.77 $\mu g/m^3$. While the Torogh station yielded moderate predictive accuracy, the Vila station exhibited weaker performance. The models demonstrated strong capability in identifying pollution episodes, with true positive rates between 66% and 74%, and a minimum false alarm rate of 0.18 at the Sajjad station. Spatial disparities in model performance were attributed to localized environmental and climatic factors, including terrain variation and emission source intensity. Overall, the findings confirm the potential of MLP-based models as practical tools for daily air quality prediction and support their integration into urban pollution alert systems.

Cite this article: ----------. 2026. Daily PM2.5 Prediction and Pollution Episode Detection Using MLP Neural Networks Across Urban Monitoring Sites with Varied Land Uses. *Environmental Resources Research*, 14(1), 91-99.

Publisher: Gorgan University of Agricultural Sciences and Natural Resources

Introduction

One of the critical factors affecting public health is air quality, which is largely influenced by the concentration of particulate matter. Among these, fine particulate matter with a diameter of less than 2.5 micrometers ($PM_{2.5}$) has been widely recognized for its direct association with adverse health effects (Biswal et al., 2022). Elevated levels of $PM_{2.5}$ have been linked to changes in lifestyle, reduced life expectancy, and increased mortality rates. Numerous studies have demonstrated a statistically significant correlation between atmospheric $PM_{2.5}$ concentrations and negative health outcomes (Yang et al., 2023). Consequently, various air quality standards for $PM_{2.5}$ have been established based on hourly, daily, and annual exposure limits. According to the latest guidelines of the World Health Organization, the recommended annual and 24-hour average limits for $PM_{2.5}$ are 5 and 15 $\mu g/m^3$, respectively (WHO, 2021). In Iran, based on the national air quality standard approved in 2016, these thresholds are set at 12 and 35 $\mu g/m^3$, respectively (DOE, 2016).

Meteorological factors such as wind speed and direction, temperature, precipitation, and planetary boundary layer height significantly influence the variability of $PM_{2.5}$ concentrations in the atmosphere. Severe air pollution episodes typically occur in regions where geographical features and atmospheric stability inhibit natural ventilation, leading to the accumulation of pollutants like $PM_{2.5}$ (Feng et al., 2019). These particles, primarily emitted from combustion sources such as urban traffic and industrial activities, remain suspended in the atmosphere for extended periods due to their ultrafine size and can rapidly reach critical concentration levels. Under such conditions, immediate interventions—such as traffic restrictions, school closures, or the suspension of polluting activities—are often required. Accordingly, both national and international air pollution control regulations mandate the implementation of automated operational procedures to prevent pollutant concentrations from exceeding predefined alert thresholds (DOE, 2016; WHO, 2021).

Mashhad, one of Iran's major and densely populated metropolitan cities, exhibits both urban and semi-industrial characteristics due to its religious, touristic, and industrial significance. A substantial portion of $PM_{2.5}$ emissions in Mashhad originates from mobile sources, including aging vehicles, diesel buses, and motorcycles. While most urban areas rely on natural gas for heating, suburban districts still utilize liquid fuels, especially in colder months, contributing to localized $PM_{2.5}$ increases. Additionally, natural sources like windblown dust from nearby deserts also elevate ambient $PM_{2.5}$ levels (Shahsavani et al., 2020).

Given the health impacts associated with $PM_{2.5}$, this pollutant has become a major public concern in the city of Mashhad. Forecasting $PM_{2.5}$ concentrations prior to pollution episodes can facilitate more effective interventions to protect public health. A wide range of operational early warning systems—based on statistical and hybrid modeling approaches—have been developed to enable proactive and real-time responses to air pollution events. In this context, predictive models have been increasingly utilized as supportive tools for air quality management in various regions around the world.

Artificial neural networks (ANNs) have been widely applied to forecast the concentrations of various pollutants over different time scales, yielding promising results (Etemad-Shahidi et al., 2010; Zoqi et al., 2016). In air quality prediction studies, methods such as ANNs, multiple linear regression (MLR), and stepwise regression (SWR) are among the most commonly used approaches (Shams et al., 2023). Due to the complex and nonlinear relationships between meteorological parameters and pollution levels, ANNs have demonstrated superior performance compared to traditional statistical models (Cakir and Sita, 2020). Since their initial application in modeling atmospheric pollutant concentrations (Boznar et al., 1993), ANNs have been regarded as a reliable method in this field.

Although forecasting particulate matter concentrations is more complex than

modeling gaseous pollutants—due to the intricate processes involved in aerosol formation, transport, and removal (Sokhi et al., 2021) —neural networks have demonstrated high accuracy owing to their ability to identify and model nonlinear relationships (Su et al., 2025). Feedforward neural networks with error backpropagation (FFNNs) are among the most commonly used neural network architectures for predicting pollutants such as $PM_{2.5}$, PM_{10}, O_3, SO_2, and CO, due to their capacity to model complex nonlinear interactions (Elbayoumi et al., 2015). In one study, several machine learning methods were used to predict $PM_{2.5}$ exceedance events, and FFNNs exhibited superior performance (Suri et al., 2023). In other studies, various machine learning approaches—including FFNNs, pruned neural networks (PNNs), and lazy learning (LL) techniques—have been applied for $PM_{2.5}$ concentration prediction, with FFNNs consistently outperforming the alternatives (Yang and Chen, 2021). Another study analyzed multiple methods for forecasting daily average $PM_{2.5}$ concentrations. The results of two types of multilayer perceptron (MLP) networks—an important subclass of FFNNs—and a radial basis function (RBF) network were compared with two classical models, and the MLP model demonstrated superior predictive performance (Ganesh et al., 2018).

In the present study, a model was developed to predict daily $PM_{2.5}$ concentrations based on air pollution data collected from three monitoring stations—Sajjad, Vila, and Torogh—in the city of Mashhad. These stations represent diverse urban conditions: the Sajjad station is located in a densely populated, high-traffic area; the Vila station is situated in a region with moderate population density; and the Torogh station is positioned on the southeastern outskirts of the city, where air quality is influenced by industrial activities and heavy-duty vehicle traffic.

An MLP neural network was employed to predict the daily average $PM_{2.5}$ concentrations at the selected monitoring stations. Model inputs included the previous day's average $PM_{2.5}$ concentration as well as the previous day's meteorological parameters, such as mean wind speed and direction, precipitation, solar radiation, temperature, and relative humidity. Additionally, to account for variations in traffic patterns across different days of the week and throughout the year—due to Mashhad's religious and touristic nature—variables such as the day of the week (1 to 7) and the month of the year (1 to 12) were incorporated into the model as input features. The primary objective of this study is to protect at-risk populations by providing accurate and timely information on air quality.

Materials and Methods

Study Area

This study was conducted in the metropolitan city of Mashhad, located in northeastern Iran. With a population exceeding 3.5 million, Mashhad is the second most populous city in the country after Tehran. Due to its religious and touristic significance, the city attracts over 20 million domestic and international pilgrims and tourists annually. These characteristics have led to high urban traffic density, increased fossil fuel consumption, and the expansion of commercial and service activities—major sources of $PM_{2.5}$ emissions.

Topographically, Mashhad is situated on a plain at an average elevation of approximately 980 m above sea level. It is surrounded by the Hezar-Masjed Mountains to the north (with elevations exceeding 2800 m) and the Binalood Range to the southwest (reaching approximately 3211 m). This complex terrain gives rise to specific climatic phenomena such as mountain-to-plain breezes, nocturnal temperature inversions during cold nights, and convective airflows in warmer seasons. These conditions can contribute to elevated pollutant concentrations or the persistence of pollutants near the surface, particularly during winter when thermal inversion events are more intense.

The average annual temperature in Mashhad is approximately 12.4 °C, and the mean annual precipitation is reported to be around 248.6 mm. The prevailing winds in this region typically blow from the west and northwest, with an average speed of about

3 m/s (Daneshvar et al., 2025). On average, approximately 86 days per year in Mashhad are classified as having unhealthy air quality for sensitive groups—including children, the elderly, and individuals with respiratory illnesses. This condition is primarily attributed to combustion-related sources such as heavy traffic congestion, domestic heating, industrial activities, and regional climatic factors (Mohammadi et al., 2022).

In this study, data from three air quality monitoring stations were used, each located in areas with distinct environmental characteristics and pollution sources. The Sajjad station was selected to represent urban background pollution associated with $PM_{2.5}$. This station is situated in the city center at an elevation of approximately 1000 m above sea level, near Falasteen Square, and within less than 100 m of major roads such as Sajjad and Ahmadabad Boulevards. This area is among the busiest traffic zones in Mashhad and is significantly influenced by emissions from both light and heavy vehicles.

The Torogh station, located in the southeastern part of Mashhad at an elevation of approximately 1020 m above sea level, lies about 15 km from the city center and near the Torogh Industrial Zone. This station reflects the combined impact of pollution from high diesel vehicle traffic, industrial activities, and emissions from heavy-duty transportation sources.

The Vila station is situated in a mixed residential-commercial area with moderate traffic density in the southwestern part of the city, at an elevation of approximately 970 m. It is located about 5 km from the city center and is relatively distant from direct sources of industrial emissions. Therefore, this station is considered representative of areas with lower background pollution levels.

Design and Training of MLP Neural Networks

In this study, three MLP neural networks were developed to predict the daily average concentration of $PM_{2.5}$ at three air quality monitoring stations in Mashhad: Sajjad, Torogh, and Vila. Each network was independently designed and trained for a specific station—ANN1 for Sajjad, ANN2 for Torogh, and ANN3 for Vila.

The initial dataset comprised several years of hourly pollutant concentrations and meteorological parameters. These data were preprocessed using the Python programming language, with the aid of the pandas library for structured data manipulation and NumPy for numerical computations. During preprocessing, the data were time-indexed by date and aggregated into daily values (24-hour averages) to enable the prediction of next-day $PM_{2.5}$ concentrations. In instances where one or more input variables were missing for a given day, the corresponding row was removed from the dataset to avoid errors associated with imputation of missing values. Ultimately, approximately 70–80% of the daily data for each station remained complete and was deemed suitable for model development.

The general architecture of each network consisted of three main layers: an input layer, a hidden layer, and an output layer. The input layer incorporated the modeling variables, including meteorological and air quality parameters. The hidden layer comprised several neurons that computed the weighted sum of the inputs and passed the results through a nonlinear tangent sigmoid activation function. The final output was produced in the output layer by combining the weighted outputs of the hidden layer neurons using a linear activation function.

The use of the tangent sigmoid function in the hidden layer enables the network to capture nonlinear mappings and model complex relationships, while the linear activation function in the output layer ensures the generation of continuous outputs without additional nonlinear transformation. This combination of activation functions provides an efficient framework for various prediction problems (Zoqi et al., 2010).

To identify the most appropriate training algorithm for each station, initial neural networks with 10 neurons in the hidden layer were designed and trained using several algorithms. The performance of each algorithm was evaluated on the training, validation, and testing datasets using

statistical performance metrics. The algorithm yielding the best overall performance was selected for further modeling.

Once the optimal training algorithm was identified, the process of determining the optimal number of neurons in the hidden layer for each network began. Networks were initially trained with 10 neurons, and the number of neurons was then gradually increased to assess different network configurations. For each configuration, statistical performance indicators were recorded and analyzed for the training, validation, and testing sets. This process continued until the most efficient network structure—based on performance metrics—was identified for each monitoring station.

To prevent overfitting, network structures that exhibited decreased training error but increased validation or testing error were excluded. Network training was performed using the backpropagation algorithm, aiming to minimize the discrepancy between predicted and actual outputs. This process involved a forward pass for signal propagation and a backward pass for updating the connection weights through error backpropagation.

The modeling process was conducted using MATLAB software. To mitigate overfitting, the dataset was split into 70% for training, 15% for validation, and 15% for testing. The early stopping method was employed to further reduce the risk of overfitting. Additionally, to enhance the stability and reliability of the results, each network was independently trained three times.

Data Preparation

For each of the three monitoring stations in Mashhad (Sajjad, Vila, and Torogh), separate daily datasets were compiled for the period from 2018 to 2023. These datasets included both meteorological parameters and $PM_{2.5}$ concentration levels. Meteorological data—comprising air temperature (°C), relative humidity (%), average wind speed (m/s), wind direction (degrees), and precipitation (mm)—were obtained from the Khorasan Razavi Meteorological Organization.

$PM_{2.5}$ concentration data were acquired from the Khorasan Razavi Department of Environment. These data were continuously recorded on an hourly basis using online beta attenuation monitoring instruments. In this method, an airstream passes over a filter, and the decrease in beta radiation intensity—due to the accumulation of particulate matter—is used to calculate the mass concentration of $PM_{2.5}$. Although beta attenuation monitors are recognized as semi-reference methods (Shukla and Aggarwal, 2022), to enhance measurement accuracy, the recorded data were corrected using adjustment coefficients derived from comparisons with gravimetric reference methods when necessary. These correction coefficients were calculated separately for each station and on a seasonal basis and were applied only when the correlation coefficient between beta and gravimetric data exceeded 0.8. Data coverage over the study period (2018–2023) was approximately 80% for each station. To maintain analytical integrity, no extrapolation or interpolation was performed to replace missing data. In cases where input variables were missing for a particular day, the corresponding daily record was entirely excluded from the modeling dataset to ensure consistency and reliability during training and validation phases.

The annual average $PM_{2.5}$ concentration during the study period ranged from 35 to 38 μg/m³ at the Sajjad station, 26 to 32 μg/m³ at the Torogh station, and 20 to 28 μg/m³ at the Vila station. These variations reflect the relative intensity of local emission sources, including traffic, industrial activity, and domestic heating, in different parts of the city.

To provide an overview of the statistical characteristics of the dataset used in this study, descriptive indicators—including minimum, maximum, mean, and standard deviation—were calculated for the main parameters ($PM_{2.5}$ concentration, temperature, humidity, wind speed, and precipitation). The results of this statistical analysis are presented in Table 1. These indicators help identify data variability and offer insights into the distribution of variables over the study period.

These values indicate considerable variability in both pollutant concentrations and meteorological conditions, which is crucial for understanding the dynamics of $PM_{2.5}$ behavior and its interaction with environmental parameters in the urban context of Mashhad.

Table 1. Descriptive statistics of meteorological variables and $PM_{2.5}$ concentrations over the six-year period (2018–2023)

Variable	Range	Minimum	Maximum	Mean	Standard Deviation
$PM_{2.5}$ concentration at Sajjad Station (μg/m^3)	112.08	1.01	113.09	36.54	18.81
$PM_{2.5}$ concentration at Torogh Station (μg/m^3)	151.23	3.62	154.85	29.97	26.43
$PM_{2.5}$ concentration at Vila Station (μg/m^3)	140.00	1.12	141.12	22.30	23.03
Average precipitation (mm)	49.80	0.00	49.80	0.56	2.55
Maximum temperature (°C)	52.90	-9.70	43.20	23.35	10.29
Average temperature (°C)	48.70	-14.50	34.20	16.81	9.18
Maximum wind speed (m/s)	18.00	2.00	20.00	6.99	2.59
Average wind speed (m/s)	9.00	0.25	9.25	3.04	1.07
Average relative humidity (%)	88.25	11.75	100.00	49.04	21.68

As shown in Table 1, the variables under investigation exhibit different scales and ranges. For example, $PM_{2.5}$ concentrations at the Sajjad station range from 1.01 to 113.09 μg/m^3, while the average wind speed varies from 0.25 to 9.25 m/s. These wide disparities in variable ranges highlight the necessity of standardization. Without it, variables with larger numerical ranges—such as $PM_{2.5}$—may lead to unbalanced learning, reduced convergence speed during model training, and numerical instability (Zoqi and Saeedi, 2011).

To address this issue, this study employed a standardization approach in which the input data were transformed to have a mean of zero and a standard deviation of one (Equation 1). This process promotes a more stable and uniform gradient flow, reduces numerical instabilities caused by large input values, and accelerates convergence in gradient-based optimization algorithms. It also enhances the effectiveness of activation functions. Additionally, centering the features around the mean and bringing them onto a comparable scale improves model interpretability and contributes to more efficient learning and better generalization to unseen data (Zoqi, 2024).

Furthermore, to incorporate temporal categorical features such as the day of the week and the month of the year, one-hot encoding was applied. In this method, each category (i.e., each weekday or month) is represented as a distinct binary vector, thereby preventing the erroneous imposition of ordinal or numerical relationships between categories. The use of one-hot encoding eliminates numerical bias and enables the model to capture temporal patterns more accurately, ultimately improving its predictive performance.

$$X_{scaled} = \frac{X - \mu}{\sigma} \quad (1)$$

where X_{scaled} is the standardized value, X is the original value, μ is the mean of the values, and σ is the standard deviation.

Input Variables

The selection of variables examined in this study was based on a review of relevant literature and the availability of existing data. To assess the appropriateness of the selected variables and evaluate the relationships among them, Pearson correlation analysis was employed. This method is a fundamental and widely used technique in statistical analysis and data mining for measuring the strength and direction of the linear relationship between two continuous variables (Zoqi, 2024). The correlation coefficient, calculated according to Equation (2), quantifies the linear dependence between each pair of variables within the range of [–1, +1], where +1 indicates a strong positive relationship, –1 signifies a strong negative relationship, and 0 denotes the absence of a significant linear association. The aim of this process was to identify the most influential

variables affecting the dependent variable and to eliminate irrelevant or redundant features.

$$r_{XY} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \quad (2)$$

In Equation (2), X and Y represent a pair of variables, while $\bar{X}$ and $\bar{Y}$ denote the mean values of these variables, respectively.

In this study, a Pearson correlation matrix was constructed—a symmetric structure in which each cell represents the correlation coefficient between a pair of variables. Analyzing this matrix enables the identification of highly correlated variables—those with an absolute correlation coefficient greater than 0.8. Such high correlations indicate multicollinearity and redundancy of information; in these cases, the associated variables are excluded to prevent adverse effects on model stability and accuracy. Additionally, variables showing a weak correlation with the output variable (absolute value less than 0.1) are removed, as they have minimal explanatory power for the dependent variable and do not significantly contribute to model performance (Cheng et al., 2022). A careful application of Pearson correlation analysis enhances the precision of variable selection, improves model generalizability, and reduces unnecessary complexity.

Results and Discussion

Correlation Analysis between Input Variables and $PM_{2.5}$ Concentration

Prior to designing the neural network model, the Pearson correlation matrix was calculated to assess the relationships between $PM_{2.5}$ concentration and other measured parameters at each station. As shown in Figure 1, some variables exhibit significant correlations with $PM_{2.5}$ concentrations, while others show negligible influence on this pollutant. Among the examined parameters, relative humidity demonstrated a moderately strong negative correlation with $PM_{2.5}$ (correlation coefficients ranging from -0.38 to -0.45 across different stations), indicating a decrease in particle concentration with increasing humidity. This inverse relationship can be attributed to enhanced wet deposition and reduced atmospheric residence time of particles (Emekwuru and Ejohwomu, 2023).

Air temperature showed a moderate positive correlation with $PM_{2.5}$ (correlation coefficients between 0.30 and 0.36 across stations), which may result from atmospheric stability during warmer periods of the day, reduced humidity, and limited dispersion of pollutants. Moreover, a negative correlation was observed between temperature and relative humidity, as rising temperatures increase the air's capacity to hold water vapor, thereby reducing relative humidity (Emekwuru and Ejohwomu, 2023). Under such conditions, higher temperatures not only elevate anthropogenic activity and pollutant emissions but also reduce the effectiveness of natural removal processes such as wet deposition, while facilitating the resuspension of particulate matter. Additionally, local atmospheric stability in the lower boundary layer can further inhibit pollutant dispersion (Nyayapathi et al., 2025). Overall, hot and dry conditions promote increased generation, resuspension, and persistence of $PM_{2.5}$, ultimately leading to elevated concentrations. Conversely, during winter, the dominant meteorological conditions—such as high relative humidity, stratiform cloud cover, and light precipitation—contribute to lower temperatures and enhanced wet deposition, thereby reducing $PM_{2.5}$ levels (Emekwuru and Ejohwomu, 2023).

However, the influence of temperature on $PM_{2.5}$ concentration extends beyond variations in relative humidity or atmospheric instability. One of the key factors in this context is the inversion layer height, which is directly affected by the thermal structure of the atmosphere (Liu et al., 2022). Climatic assessments of Mashhad indicate that the base height of temperature inversion layers varies between 500 and 1000 m from June to September, and between 100 and 500 m from October to May. Notably, in March, the mean minimum inversion height reaches approximately 120 m (Mansouri Daneshvar et al., 2024). This period coincides with the highest frequency of critical air pollution episodes, as strong, low-altitude inversions trap pollutants within the lower atmospheric layers, leading to significant increases in their concentrations. Therefore, temperature plays a more complex role in regulating $PM_{2.5}$

levels by influencing atmospheric stratification and stability, which cannot be fully captured through simple correlation analyses.

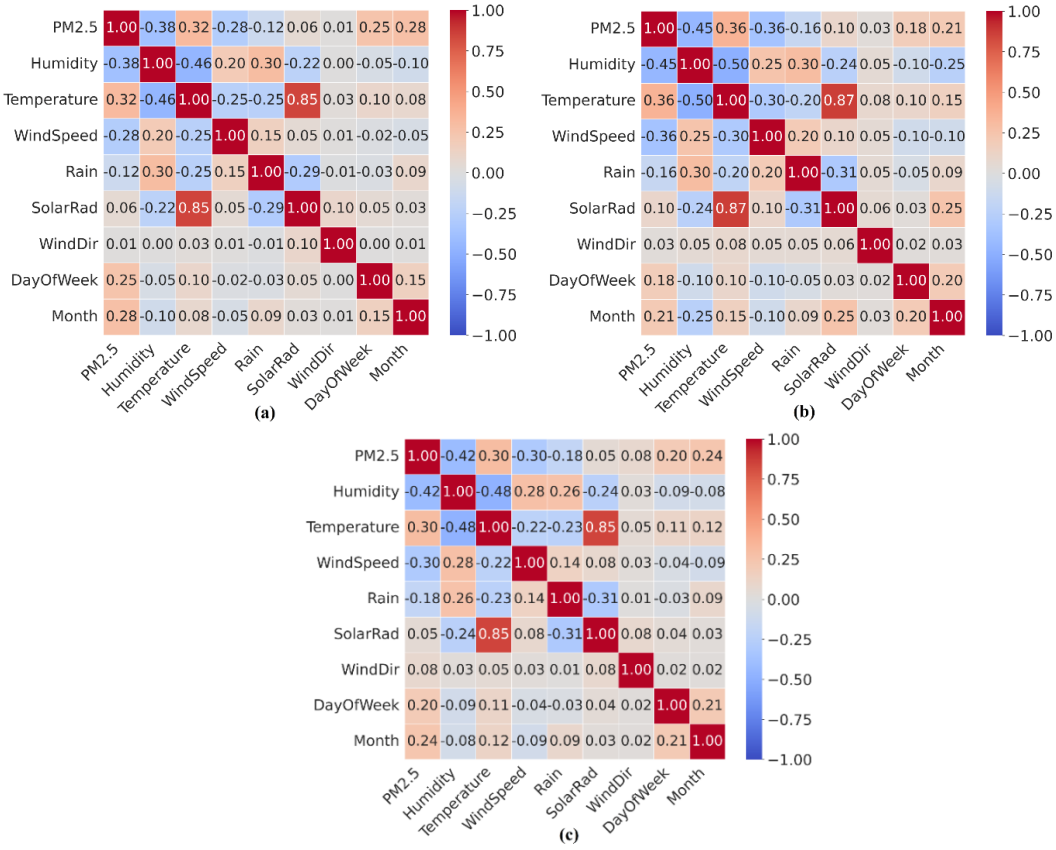


Figure 1. Spearman correlation matrix between pollutant and meteorological variables at stations: (a) Sajjad, (b) Torogh, and (c) Vila

The analysis of the Pearson correlation matrix revealed a weak negative correlation between precipitation and $PM_{2.5}$ concentration across the monitored stations, with coefficients ranging from -0.12 to -0.18. Although this correlation is statistically weak, its influence should not be entirely disregarded. Precipitation contributes to the temporary reduction of airborne particulate matter through the wet deposition process, which removes particles from the atmosphere and transports them to the ground. In the semi-arid climate of Mashhad, where the average annual precipitation is reported to be approximately 220–260 mm (Daneshvar et al., 2025), not only is the number of rainy days limited, but the intensity and persistence of rainfall are also generally insufficient to cause a significant and sustained reduction in particle concentrations (Jaafari et al., 2018).Therefore, on short-term timescales (daily or weekly), precipitation cannot be considered a major determinant in explaining $PM_{2.5}$ concentration variability, and its impact is mainly confined to intense or exceptional rainfall events. These findings are consistent with previous studies, which have demonstrated that the effect of precipitation in arid and semi-arid regions is significantly weaker than in humid climates due to lower rainfall frequency and intensity (Marzouni et al., 2016).

Based on the correlation matrix results, the mean wind speed showed a moderate negative correlation with $PM_{2.5}$ concentration, with correlation coefficients ranging from -0.28 to -0.36 across different stations. This suggests that higher wind speeds are generally associated with lower $PM_{2.5}$ concentrations. Wind plays an essential role in ventilating polluted air through

processes such as dilution, horizontal and vertical dispersion, and the transport of pollutants across the lower atmospheric layers, particularly under conditions with strong regional airflow. However, in this study, the most frequent mean wind speed was observed in the range of 2.0 to 3.88 m/s, indicating that for most days of the year, Mashhad experiences relatively low wind speeds, and strong wind events are rare. Consequently, the moderate level of observed correlation can be attributed to these typically low wind speeds. If stronger winds were more common, the impact of wind on reducing $PM_{2.5}$ concentrations would likely be more pronounced.

The analysis revealed that although westerly winds are predominant in the region, their low speeds, along with local obstructions such as terrain irregularities and dense urban structures, significantly limit their effect on the dispersion of $PM_{2.5}$ particles. As a result, wind direction showed no statistically significant correlation with $PM_{2.5}$ concentrations. Therefore, to avoid incorporating low-impact variables, reduce statistical noise, and enhance model accuracy, wind direction was excluded from the ANN input parameters.

Although solar radiation can contribute to the formation of secondary $PM_{2.5}$ through photochemical reactions under certain conditions, in Mashhad's dry and semi-industrial climate, $PM_{2.5}$ concentrations are predominantly influenced by primary sources, including vehicular traffic and wind-induced resuspension of surface dust. Moreover, due to the strong correlation between solar radiation and air temperature, and the absence of a meaningful relationship between solar radiation and the model output, this variable was also excluded from the model inputs to reduce multicollinearity and simplify the model structure.

The analysis of temporal variables revealed a weak but positive correlation between $PM_{2.5}$ concentrations and both the day of the week and month of the year. Spearman's rank correlation coefficients for the relationship between day of the week and $PM_{2.5}$ ranged from 0.18 to 0.25, and for month of the year, from 0.21 to 0.28 across the three studied stations. Although these correlations are statistically weak, they reflect the influence of temporal patterns in human activity, such as traffic volume and commercial operations, on particulate matter distribution. Additionally, time-related variables may indirectly affect pollutant behavior by influencing meteorological conditions like temperature and atmospheric stability (Zhao et al., 2018).

Further examination of individual stations showed that this correlation was strongest at the Sajjad station, compared to the Villa and Torogh stations. This variation is primarily attributed to differences in the type and intensity of human activities surrounding each monitoring site. The Sajjad district, as a major administrative-commercial hub in Mashhad, frequently experiences non-local traffic flows. Traffic patterns in this area fluctuate significantly during weekends, public holidays, and religious or social events, which in turn influence pollutant concentrations. These temporal fluctuations lead to greater sensitivity of $PM_{2.5}$ levels to time-related variables at the Sajjad station.

In contrast, the Villa station is more influenced by local activities and relatively stable traffic conditions, while the Torogh station is situated near industrial zones and heavy transportation corridors, which are less affected by temporal variations such as holidays or special events. Consequently, temporal variability in air pollution patterns is less pronounced at these two stations compared to Sajjad.

These findings are consistent with previous studies in other metropolitan areas, which have shown that administrative-commercial zones and urban attraction centers are more susceptible to fluctuations in $PM_{2.5}$ concentrations due to unstable traffic volumes (Zhao et al., 2018).

Neural Network Structure

To identify the optimal training algorithm for neural network modeling, an MLP was independently designed for each station using a fixed architecture comprising a single hidden layer with 10 neurons. The hyperbolic tangent sigmoid function was applied as the

activation function in the hidden layer, while a linear activation function was used in the output layer. Five commonly used training algorithms available in MATLAB—Levenberg–Marquardt (LM), Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG), Resilient Backpropagation (RP), and Gradient Descent (GD)—were evaluated. Among these, the LM algorithm consistently outperformed the others across all three stations and was consequently selected as the optimal training algorithm. This algorithm was employed in subsequent stages to refine the network architecture and enhance prediction accuracy.

In this study, an independent neural network was developed for each monitoring station. The selected LM algorithm was used to train an MLP with a single hidden layer of variable size. Starting from 10 neurons, the number of neurons in the hidden layer was gradually increased to assess changes in the model's performance metrics, including the coefficient of determination (R^2) and root mean square error (RMSE), across the training, testing, and validation datasets. The performance results for near-optimal configurations are summarized in Table 2.

Table 2. Optimization of the hidden layer neuron count for the three monitoring stations

Station	Neuron Count	R^2 (Train)	RMSE (Train)	R^2 (Test)	RMSE (Test)
Sajjad	15	0.802	6.12	0.784	7.12
	16 (Optimal)	0.813	5.88	0.793	6.77
	17	0.817	5.71	0.786	6.93
Torogh	10	0.786	6.91	0.769	7.89
	11 (Optimal)	0.800	6.76	0.780	7.65
	12	0.808	6.59	0.774	7.82
Vila	18	0.706	8.81	0.635	9.81
	19 (Optimal)	0.721	8.66	0.647	9.56
	20	0.729	8.53	0.639	9.73

The results presented in Table 2 indicate that increasing the number of neurons in the hidden layer consistently enhances model accuracy on the training dataset. However, this trend does not uniformly extend to the test and validation datasets. For instance, at the Sajjad station, increasing the number of neurons from 15 to 16 improved the correlation coefficient (R^2) on the test data from 0.784 to 0.793, while the RMSE decreased from 7.12 to 6.77 $\mu g/m^3$. Nevertheless, further increasing the neuron count to 17 yielded only a marginal improvement in training performance, accompanied by a decline in test set accuracy, indicating the onset of overfitting. This behavior was also observed at the Torogh and Villa stations, suggesting a consistent pattern regarding the impact of increasing network complexity. These results underscore the importance of avoiding unnecessary increases in model parameters, which may elevate the risk of overfitting.

It is worth noting that, during the model training process, the early stopping technique was employed to prevent overfitting. In this study, early stopping was implemented by halting training when the validation error failed to improve over six consecutive epochs. This approach effectively mitigates the risk of the model becoming overly fitted to the training data. However, despite the use of early stopping, an excessive increase in the number of neurons in the hidden layer can still lead to overfitting due to the model's elevated capacity (Zoqi and Ghavidel, 2009). In such cases, the network becomes excessively complex, and its generalization ability diminishes. As a result, instead of learning generalizable patterns, the model begins to memorize specific fluctuations in the training dataset, leading to reduced performance on the test and validation sets.

Based on the results presented in Table 2, the optimal neural network architectures for each monitoring station were identified as follows: ANN1 for the Sajjad station with 16 neurons,

ANN2 for the Torogh station with 11 neurons, and ANN3 for the Vila station with 19 neurons in a single hidden layer. All models utilized the hyperbolic tangent sigmoid activation function in the hidden layer and a linear activation function in the output layer.

The variation in the optimal number of neurons across the stations can be attributed to the distinct statistical characteristics of each dataset. Factors such as data variance and coefficient of variation, the presence of nonlinear patterns of varying complexity, and the signal-to-noise ratio at each station significantly influence the model capacity required for effective learning. Specifically, stations like Vila, which exhibit more severe fluctuations or more complex temporal patterns in $PM_{2.5}$ concentrations, require a larger number of neurons in the hidden layer to adequately capture the intrinsic structure of the data. This finding underscores that the optimal neural network architecture cannot be uniform or universally applied across all datasets; rather, it must be individually determined based on the statistical properties and inherent complexity of the data at each location. In this context, the trade-off between bias and variance is particularly important. Increasing the number of neurons enhances model capacity, which may reduce bias but increase variance, potentially leading to overfitting. Conversely, too few neurons may result in high bias and underfitting. Therefore, determining the optimal number of neurons should strike a balance between these two factors, ensuring that the model maintains sufficient generalization capability while accurately capturing complex data patterns (Geman et al., 1992).

Neural Network Performance

The performance of the neural network at the three air quality monitoring stations in Mashhad—Sajjad, Torogh, and Vila—is illustrated in Figure 2. This figure presents the statistical performance indicators for the training, validation, and testing phases. Based on these results, the model exhibited strong learning capability during the training phase and maintained satisfactory performance in both the validation and testing stages. These findings demonstrate the model's adequate generalization ability in predicting $PM_{2.5}$ concentrations for unseen data.

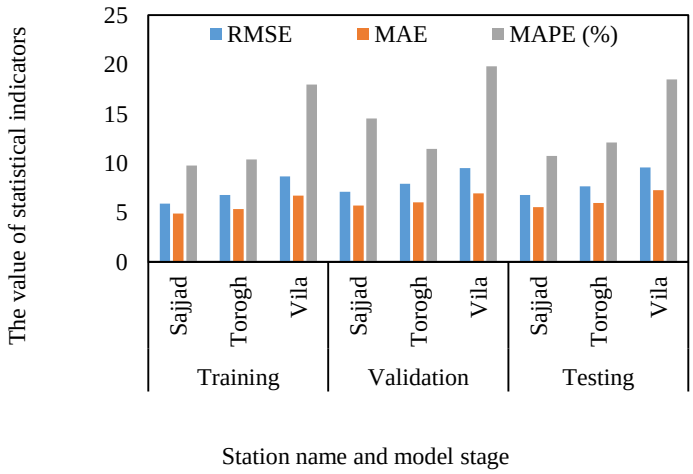


Figure 2. Comparison of MLP model performance indicators during the training, validation, and testing phases at the three monitoring stations

The Sajjad station, located in one of the most traffic-congested urban areas of Mashhad, demonstrated the highest predictive performance. This outcome can be attributed to the relatively stable pattern of traffic-related pollution, spatial and temporal consistency in pollutant dispersion, and the absence of diverse or complex emission sources. The model's statistical performance during the testing phase at this station included an R^2 of 0.79, an RMSE of 6.77 $\mu g/m^3$, a Mean Absolute Error (MAE) of 5.54 $\mu g/m^3$, and a Mean Absolute Percentage Error (MAPE) of approximately 10.73%,

indicating good agreement between observed and predicted concentrations.

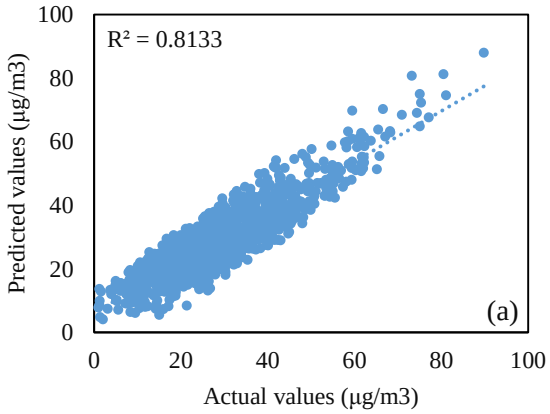
At the Torogh station, despite its peripheral location, high pollution levels were recorded—primarily due to industrial activities, the movement of heavy-duty diesel vehicles, soil erosion, and transboundary pollutant transport from neighboring regions. In this case, the model achieved a testing-phase R^2 of 0.78, an RMSE of 7.65 µg/m^3, an MAE of 5.94 µg/m^3, and a MAPE of approximately 12.09%. Although the predictive accuracy at Torogh was slightly lower than at Sajjad, the model's performance remains within an acceptable range (Gulati et al., 2023).

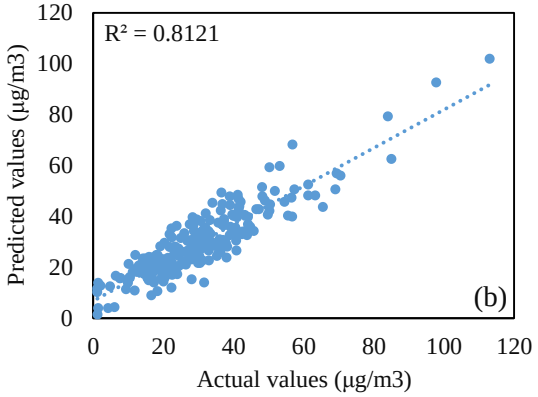
The Vila station, located in a moderately urbanized area near the Binalood mountain range, exhibited the weakest performance in predicting daily $PM_{2.5}$ concentrations. Nevertheless, even at this station, the model demonstrated an acceptable ability to capture the general data patterns. During the testing phase, the statistical indicators included an R^2 of 0.65, an RMSE of 9.56 µg/m^3, an MAE of 7.24 µg/m^3, and a MAPE of approximately 18.49%.

The reduced accuracy at this station can be attributed to the high spatiotemporal variability of multiple pollution sources—such as dispersed residential, commercial, and traffic-related activities—which, unlike the Sajjad and Torogh stations, lack spatial and temporal coherence. Furthermore, the station's geographical location near mountainous terrain influences the formation of meteorological phenomena such as temperature inversions, regional airflows, and specific wind patterns. These factors lead to reduced atmospheric stability and nonlinear pollutant transport, thereby complicating accurate prediction. The interplay of these conditions has diminished the model's capacity to precisely simulate pollutant concentration behavior in this area.

Figure 3 illustrates the regression results between observed and predicted $PM_{2.5}$ concentrations during the different phases of the modeling process. As shown, a strong correlation between actual and predicted values is evident at all three monitoring stations.

Based on the results presented in Figures 2 and 3, the neural network model demonstrated higher accuracy in areas with more consistent pollution patterns, such as the Sajjad station, while its performance declined in regions characterized by diverse pollution sources or distinct climatic conditions. For further validation, the findings of this study were compared with those of previous research, as summarized in Table 3, which corroborates the accuracy and effectiveness of the modeling approach employed in the present study.





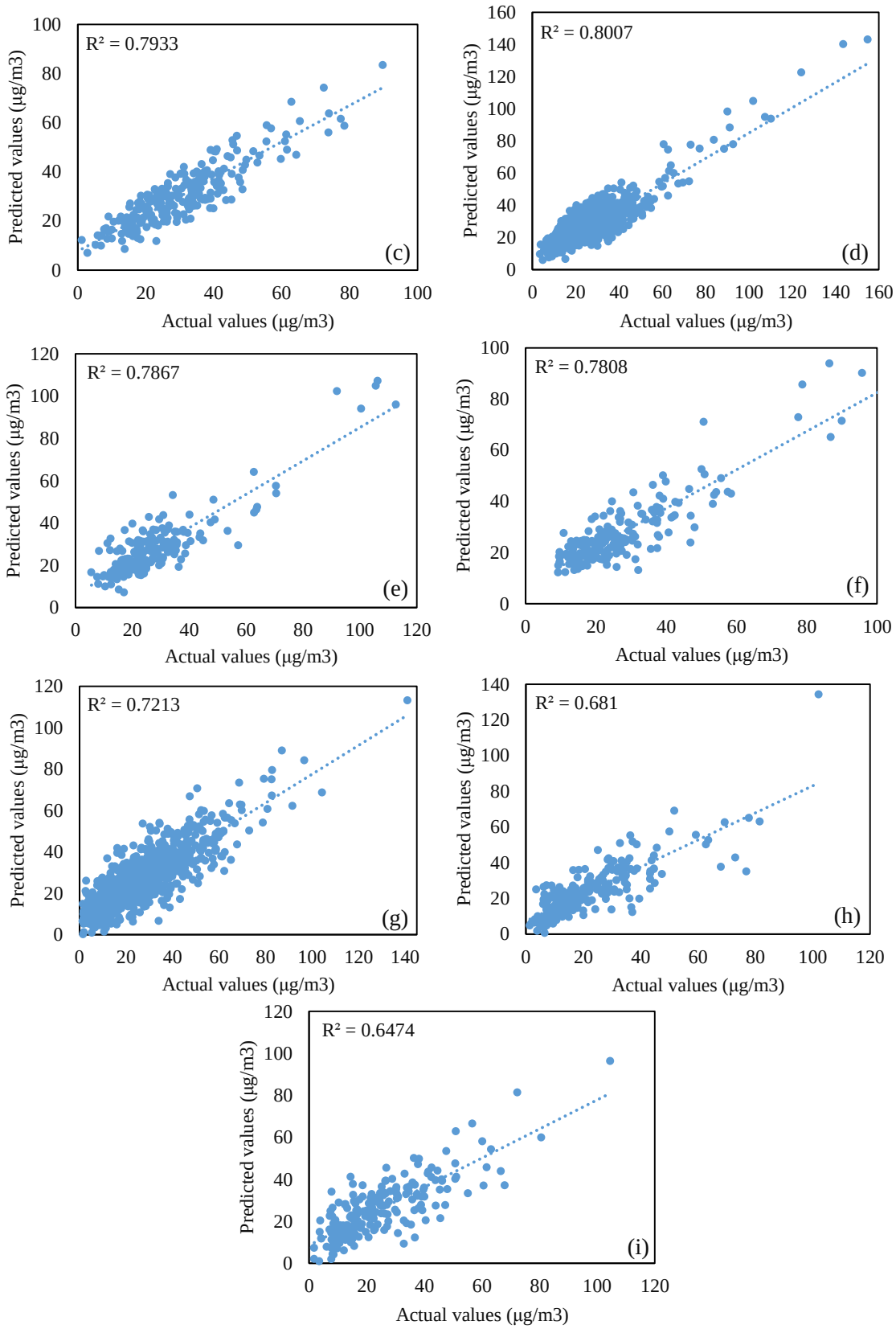


Figure 3. Regression plots comparing the actual and predicted daily $PM_{2.5}$ concentrations using the MLP neural network model at the three monitoring stations: Sajjad Station – (a) training data, (b) validation data, (c) testing data; Torogh Station – (d) training data, (e) validation data, (f) testing data; Vila Station – (g) training data, (h) validation data, and (i) testing data

Table 3. Comparison of $PM_{2.5}$ prediction model performance across different studies

Study Location	Modeling Method	R^2	RMSE (μg/m^3)	Reference
Mashhad, Iran	MLP	0.65 – 0.79	6.77 – 9.56	Present study
Qingdao, China	Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM)	0.83, 0.85	8.35, 14.36	(Bai et al., 2024)
China	Geographically Weighted Gradient Boosting Machine (GW-GBM)	0.71 – 0.76	25.02 – 25.31	(Zhan et al., 2017)
Beijing, China	XGBoost, K-Nearest Neighbors (KNN), Support Vector Regression (SVR), LSTM	0.51 – 0.76	19.58 – 28.54	(Wu et al., 2022)
Rio de Janeiro, Brazil	Holt–Winters (HW) and ANN	–	5.81 – 14.93	(Ventura et al., 2019)
Ahvaz, Iran	MLP	0.74	46.44	(Goudarzi et al., 2021)
Tehran, Iran	3D Convolutional Neural Network with Gated Recurrent Unit (3D CNN-GRU)	0.78	6.44	(Faraji et al., 2022)
Shanghai, China	CNN combined with Gradient Boosting Machine (CNN-GBM)	0.85	10.02	(Luo et al., 2020)
Shanghai, China	Weighted Artificial Neural Network (WANN)	0.87	10.66	(Guo et al., 2023)
Kolkata, India	Multiple Linear Regression (MLR) and ANN	0.51, 0.69	8.55, 12.58	(Bera et al., 2021)
Isfahan, Iran	ANN, SVM, and KNN	0.90, 0.85, 0.82	5.44, 7.24, 8.51	(Mohammadi et al., 2024)
Thailan	Spatio-temporal Deep Learning Model	0.77	7.63	(Sirisumpun et al., 2023)

Model Performance Evaluation in Predicting Daily $PM_{2.5}$ Exceedances

To assess the performance of a prediction system, various evaluation metrics can be employed. One of the most critical indicators is the system's ability to accurately predict exceedances of the daily air quality standard. According to the Iranian National Ambient Air Quality Standard, the permissible daily limit for $PM_{2.5}$ is set at 35 μg/m^3 (DOE, 2016).

To analyze the model's effectiveness in forecasting exceedance events, three key statistical indicators were used: True Positive Rate (TPR), False Alarm Rate (FAR), and F1-score. These metrics were calculated by comparing the predicted values against observed data for instances where daily $PM_{2.5}$ concentrations exceeded the threshold of 35 μg/m^3.

The parameters used to compute these performance indicators include: A, the number of correctly predicted exceedance events; F, the total number of predicted exceedances (both true and false); M, the total number of actual exceedance events; and N, the total number of days analyzed (i.e., total data points). In this context, the ratio of true positive predictions to all predicted exceedances (A/F) represents the precision of the model in identifying exceedance cases.

The TPR, as defined in Equation (3), reflects the model's ability to correctly identify days with actual exceedances and is considered a measure of the model's sensitivity.

$$TPR = \frac{A}{M} \quad (3)$$

The FAR, as defined in Equation (4), indicates the proportion of predicted exceedance events that did not actually exceed the threshold.

$$FAR = \frac{F - A}{F} \quad (4)$$

The F1-score (Equation 5) represents the harmonic mean of precision and sensitivity

(TPR). This metric is particularly useful in imbalanced classification problems, where the number of exceedance and non-exceedance cases are unevenly distributed. The F1-score effectively balances the model's ability to correctly detect exceedances while minimizing false alarms. A value closer to 1 indicates superior model performance. In environmental applications, F1-scores greater than 0.70 are typically considered acceptable.

$$F1\text{-}Score = \frac{2 \cdot Precision \cdot TPR}{Precision + TPR} \qquad (5)$$

Table 4 presents the performance metrics of the ANN model in predicting exceedance events of the daily $PM_{2.5}$ threshold (35 µg/m^3) at three air quality monitoring stations in Mashhad during the period 2018–2023. The results indicate notable spatial differences in model performance. Specifically, the ANN model demonstrated superior predictive accuracy at the Sajjad station, while a marked decline in performance was observed at the Vila station.

Table 4. Performance metrics of the ANN model for predicting daily $PM_{2.5}$ exceedance events (>35 µg/m^3) at air quality monitoring stations in Mashhad

Station	N (Total Days)	M (Observed Exceedances)	F (Predicted Exceedances)	A (True Positives)	TPR	FAR	F1-score
Sajjad	1571	463	421	343	0.74	0.18	0.78
Torogh	1210	225	198	159	0.71	0.19	0.75
Villa	1443	323	295	211	0.66	0.28	0.68

At the Sajjad monitoring station, the model's performance indicators demonstrate considerable success in identifying pollution events characterized by daily $PM_{2.5}$ concentrations exceeding the regulatory threshold. A TPR of 0.74 indicates that 74% of actual exceedance events were correctly identified. Moreover, a FAR of 0.18 reflects the model's ability to minimize false positive predictions. The F1-score, a composite metric balancing precision and recall, was calculated as 0.78, indicating a favorable balance between sensitivity and precision in detecting critical pollution events. The model's superior performance at the Sajjad station can be attributed to the relatively lower density of unpredictable pollution sources—such as construction activities and episodic public events—which often reduce forecasting accuracy. Additionally, the relative consistency of pollution patterns in this area likely enhances the model's generalizability.

The ANN model's performance in predicting $PM_{2.5}$ exceedance events aligns well with findings from similar studies conducted in regions characterized by stable pollution patterns. For instance, F1-scores ranging from 0.77 to 0.88 have been reported in Isfahan (Mohammadi et al., 2024); an F1-score of 0.62 was observed in Jakarta (Marsha and Larkin, 2019); and TPR values between 60% and 75% were documented in the western United States (Moustris et al., 2010). Furthermore, in Athens, the TPR and FAR were reported as approximately 0.78 and 17.9%, respectively (Toharudin et al., 2023).

The Torogh monitoring station, located on the southeastern outskirts of Mashhad and adjacent to industrial zones, exhibited relatively stable performance comparable to that of the Sajjad station. The F1-score and FAR at this station were calculated as 0.75 and 0.19, respectively, indicating acceptable accuracy in predicting pollution events and a low rate of false alarms. The presence of relatively consistent pollution patterns stemming from industrial activities, along with the spatial distance from the city center—which reduces the diversity of pollution sources—are key factors contributing to reduced data noise and improved model performance. The findings of this study are consistent with results reported in other industrial regions. For instance, ANN models achieved F1-scores of approximately 0.76 in the Shah Alam Industrial Area, Malaysia (Arafin et al.,

2024), and F1-scores ranging from 0.74 to 0.80 in the Puli Industrial Zone in Taiwan (Yin et al., 2021).

In contrast, at the Vila station, the F1-score dropped to 0.68, indicating a decline in model performance. This station is situated in the southwestern part of Mashhad, in an area characterized by specific topographic features, including proximity to southern highlands and urban green spaces. Such conditions cause significant variability in wind flow and pollution dispersion patterns. Furthermore, temperature inversion phenomena—especially during early morning hours or colder seasons—can lead to the accumulation of pollutants in the lower atmospheric layers, complicating their dispersion patterns. These dynamics are often overlooked in models that lack vertical atmospheric data. Additionally, the meteorological data used in the modeling process were derived from synoptic stations, which mainly provide generalized representations of regional atmospheric conditions and are insufficient for capturing fine-scale local variations. This mismatch between input data and real-world conditions has likely reduced the model's predictive accuracy at this station.

Therefore, the integration of high-resolution mesoscale meteorological models, such as the Weather Research and Forecasting (WRF) model, can enhance the accuracy of local input data—particularly in regions with complex terrain and heterogeneous climatic conditions—thereby improving prediction performance. Similar studies have also highlighted the adverse effects of topographical complexity and climatic variability on the performance of air pollution forecasting models. For example, research conducted in Thailand (Bera et al., 2021) and India (Sirisumpun et al., 2023) reported that in mountainous or valley regions, ANN models experienced significant performance declines, with mean F1-scores ranging between 0.58 and 0.65. These findings are consistent with the performance observed at the Vila station.

Overall, the analysis of Table 4 reveals that the ANN model performs significantly better at stations located in areas with well-defined emission sources, stable dispersion patterns, and relatively simple topographic conditions. This underscores the importance of spatial adaptability in model design, the necessity of high-resolution meteorological data, and the inclusion of local geographic characteristics in the development of air quality forecasting and early warning systems. Given the model's satisfactory performance at selected stations, it can serve as a foundational component for the implementation of localized air pollution alert systems.

Conclusions

The findings of this study demonstrate that the MLP neural network model, when individually designed and optimized for each specific location, can provide reliable performance in forecasting daily $PM_{2.5}$ concentrations in both urban and industrial environments. Analysis across the three monitoring stations revealed that the model's performance is strongly influenced by topographic features, homogeneity of emission sources, and local climatic stability. Stations characterized by stable dispersion patterns and well-defined pollutant sources yielded significantly higher predictive accuracy, whereas performance declined in areas with variable meteorological conditions or diverse emission sources.

Moreover, it was shown that applying techniques such as data standardization, early stopping during the training phase, and tailoring the network architecture based on the statistical characteristics of each station substantially enhanced the model's precision. A key finding was the pivotal role of Pearson correlation analysis in identifying effective input variables. The results indicated that parameters such as air temperature, relative humidity, and average wind speed exhibited the strongest correlations with $PM_{2.5}$ concentration fluctuations. In contrast, variables such as wind direction and solar radiation had negligible influence on predictive accuracy and were excluded from the model's input dataset. A notable observation was the variation in the optimal number of hidden layer neurons across the different stations, underscoring the

importance of avoiding a uniform network architecture for all geographic locations.

Based on these results, it can be concluded that neural network models—when adapted to site-specific characteristics—offer considerable potential as supportive tools in the development of localized air pollution early warning systems and in aiding urban management in responding to critical pollution events. However, the implementation of such systems requires improved spatial resolution of meteorological datasets, exploration of various temporal forecasting scenarios, and ongoing validation of model performance using updated observational data.

Acknowledgement
The authors would like to thank the Khorasan Razavi Meteorological Organization (Iran) for providing access to meteorological data used in this study. We also gratefully acknowledge the Khorasan Razavi Department of Environment (Iran) for supplying the PM2.5 concentration data. This research was conducted as part of a sabbatical leave under the Faculty–Industry and Society Collaboration Program at the University of Birjand (Iran). The authors sincerely appreciate the support and official approval of the University of Birjand for this research.

Disclosure Statement
No potential conflict of interest was reported by the authors.

Data Availability Statement
All data generated or analyzed during this study are available from the corresponding author upon reasonable request.

References

Arafin, S.K., Ul-Saufie, A.Z., Ghani, N.A.M., Ibrahim, N., 2024. A two-stage feature selection method to enhance prediction of daily pm2. 5 concentration air pollution. Environment and Natural Resources Journal 22(6), 500-509.

Bai, X., Zhang, N., Cao, X., Chen, W., 2024. Prediction of pm2. 5 concentration based on a cnn-lstm neural network algorithm. PeerJ 12, e17811.

Bera, B., Bhattacharjee, S., Sengupta, N., Saha, S., 2021. Pm2. 5 concentration prediction during covid-19 lockdown over kolkata metropolitan city, india using mlr and ann models. Environmental Challenges 4, 100155.

Biswal, B.K., Bolan, N., Zhu, Y.-G., Balasubramanian, R., 2022. Nature-based systems (nbs) for mitigation of stormwater and air pollution in urban areas: A review. Resources, Conservation and Recycling 186, 106578.

Boznar, M., Lesjak, M., Mlakar, P., 1993. A neural network-based method for short-term predictions of ambient so2 concentrations in highly polluted industrial areas of complex terrain. Atmospheric environment. Part B. urban atmosphere 27(2), 221-230.

Cakir, S., Sita, M., 2020. Evaluating the performance of ann in predicting the concentrations of ambient air pollutants in nicosia. Atmospheric Pollution Research 11(12), 2327-2334.

Cheng, J., Sun, J., Yao, K., Xu, M., Cao, Y., 2022. A variable selection method based on mutual information and variance inflation factor. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 268, 120652.

Daneshvar, M.R.M., Rahmati, L., Rasouli, S.J., Aliaqa, A.D., 2025. An analysis of the spatial and temporal correlations between rainfall events and earth surface characteristics, mashhad district in northeastern iran. Theoretical and Applied Climatology 156(2), 1-18.

DOE. 2016. *Ambient air quality standards (approval no. 19322 of the supreme council for environmental protection)*. Monitoring and Assessment Office, Deputy for Human Environment.

Elbayoumi, M., Ramli, N.A., Yusof, N.F.F.M., 2015. Development and comparison of regression models and feedforward backpropagation neural network models to predict seasonal indoor pm2. 5–10 and pm2. 5 concentrations in naturally ventilated schools. Atmospheric Pollution Research 6(6), 1013-1023.

Emekwuru, N., Ejohwomu, O., 2023. Temperature, humidity and air pollution relationships during a period of rainy and dry seasons in lagos, west africa. Climate 11(5), 113.

Etemad-Shahidi, A., Zoghi, M., Saeedi, M., 2010. An alternative data driven approach for prediction of thermal discharge initial dilution using tee diffusers. International Journal of Environmental Science & Technology 7, 29-36.

Faraji, M., Nadi, S., Ghaffarpasand, O., Homayoni, S., Downey, K., 2022. An integrated 3d cnn-gru deep learning method for short-term prediction of pm2. 5 concentration in urban environment. Science of The Total Environment 834, 155324.

Feng, H., Zou, B., Wang, J., Gu, X., 2019. Dominant variables of global air pollution-climate interaction: Geographic insight. Ecological indicators 99, 251-260.

Ganesh, S.S., Arulmozhivarman, P., Tatavarti, V.R., 2018. Prediction of pm 2.5 using an ensemble of artificial neural networks and regression models. Journal of Ambient Intelligence and Humanized Computing, 1-11.

Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. Neural computation 4(1), 1-58.

Goudarzi, G., Hopke, P.K., Yazdani, M., 2021. Forecasting pm2. 5 concentration using artificial neural network and its health effects in ahvaz, iran. Chemosphere 283, 131285.

Gulati, S., Bansal, A., Pal, A., Mittal, N., Sharma, A., Gared, F., 2023. Estimating pm2. 5 utilizing multiple linear regression and ann techniques. Scientific Reports 13(1), 22578.

Guo, Q., He, Z., Wang, Z., 2023. Predicting of daily pm2. 5 concentration employing wavelet artificial neural networks based on meteorological elements in shanghai, china. Toxics 11(1), 51.

Jaafari, J., et al., 2018. Study of pm10, pm2. 5, and pm1 levels in during dust storms and local air pollution events in urban and rural sites in tehran. Human and ecological risk assessment: An International Journal 24(2), 482-493.

Liu, B., et al., 2022. The relationship between atmospheric boundary layer and temperature inversion layer and their aerosol capture capabilities. Atmospheric research 271, 106121.

Luo, Z., Huang, F., Liu, H., 2020. Pm2. 5 concentration estimation using convolutional neural network and gradient boosting machine. Journal of Environmental Sciences 98, 85-93.

Mansouri Daneshvar, M.R., Khatami, R., Ebrahimi, M., 2024. An analysis of urban heat island based on the temperature inversion across the atmospheric boundary layer in mashhad, iran (2018–2022). Spatial Information Research 32(5), 607-622.

Marsha, A., Larkin, N.K., 2019. A statistical model for predicting pm2. 5 for the western united states. Journal of the Air & Waste Management Association 69(10), 1215-1229.

Marzouni, M.B., et al., 2016. A comparison of health impacts assessment for pm10 during two successive years in the ambient air of kermanshah, iran. Atmospheric Pollution Research 7(5), 768-774.

Mohammadi, F., Teiri, H., Hajizadeh, Y., Abdolahnejad, A., Ebrahimi, A., 2024. Prediction of atmospheric pm2. 5 level by machine learning techniques in isfahan, iran. Scientific Reports 14(1), 2109.

Mohammadi, M., Hatami, M., Esmaeli, R., Gohari, S., Mohammadi, M., Khayami, E., 2022. Relationships between ambient air pollution, meteorological parameters and respiratory mortality in mashhad, iran: A time series analysis. Pollution 8(4), 1250-1265.

Moustris, K.P., Ziomas, I.C., Paliatsos, A.G., 2010. 3-day-ahead forecasting of regional pollution index for the pollutants no 2, co, so 2, and o 3 using artificial neural networks in athens, greece. Water, Air, & Soil Pollution 209, 29-43.

Nyayapathi, P.P., Namuduri, S., Kolli, S.K., 2025. A comprehensive review of vertical profiling of ambient air quality-particulate matter and its impacts on climatic & environmental health. Air Quality, Atmosphere & Health, 1-27.

Shahsavani, A., et al., 2020. Short-term effects of particulate matter during desert and non-desert dust days on mortality in iran. Environment international 134, 105299.

Shams, S.R., et al., 2023. Assessing the effectiveness of artificial neural networks (ann) and multiple linear regressions (mlr) in forcasting aqi and pm10 and evaluating health impacts through airq+ (case study: Tehran). Environmental Pollution 338, 122623.

Shukla, K., Aggarwal, S.G., 2022. A technical overview on beta-attenuation method for the monitoring of particulate matter in ambient air. Aerosol and Air Quality Research 22(12), 220195.

Sirisumpun, N., Wongwailikhit, K., Painmanakul, P., Vateekul, P., 2023. Spatio-temporal pm2. 5 forecasting in thailand using encoder-decoder networks. IEEE Access 11, 69601-69613.

Sokhi, R.S., et al., 2021. Advances in air quality research–current and emerging challenges. Atmospheric Chemistry and Physics Discussions 2021, 1-133.

Su, Q., Ruslan, L., Shuyi, C., 2025. Ann modeling for predicting nox and particulate matter emissions from cement kilns. Journal of Cleaner Production, 145707.

Suri, R.S., et al., 2023. Air quality prediction-a study using neural network based approach. Journal of Soft Computing in Civil Engineering 7(1), 93-113.

Toharudin, T., et al., 2023. Boosting algorithm to handle unbalanced classification of pm 2.5 concentration levels by observing meteorological parameters in jakarta-indonesia using adaboost, xgboost, catboost, and lightgbm. IEEE Access 11, 35680-35696.

Ventura, L.M.B., de Oliveira Pinto, F., Soares, L.M., Luna, A.S., Gioda, A., 2019. Forecast of daily pm 2.5 concentrations applying artificial neural networks and holt–winters models. Air Quality, Atmosphere & Health 12, 317-325.

WHO. 2021. *Who global air quality guidelines: Particulate matter (pm2. 5 and pm10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.

Wu, Y., Lin, S., Shi, K., Ye, Z., Fang, Y., 2022. Seasonal prediction of daily pm2. 5 concentrations with interpretable machine learning: A case study of beijing, china. Environmental Science and Pollution Research 29(30), 45821-45836.

Yang, M.-C., Chen, M.C., 2021. Composite neural network: Theory and application to pm2. 5 prediction. IEEE Transactions on Knowledge and Data Engineering 35(2), 1311-1323.

Yang, T., Wang, J., Huang, J., Kelly, F.J., Li, G., 2023. Long-term exposure to multiple ambient air pollutants and association with incident depression and anxiety. JAMA psychiatry 80(4), 305-313.

Yin, P.-Y., Chang, R.-I., Day, R.-F., Lin, Y.-C., Hu, C.-Y., 2021. Improving pm2. 5 concentration forecast with the identification of temperature inversion. Applied Sciences 12(1), 71.

Zhan, Y., et al., 2017. Spatiotemporal prediction of continuous daily pm2. 5 concentrations across china using a spatially explicit machine learning algorithm. Atmospheric Environment 155, 129-139.

Zhao, N., Liu, Y., Vanos, J.K., Cao, G., 2018. Day-of-week and seasonal patterns of pm2. 5 concentrations over the united states: Time-series analyses using the prophet procedure. Atmospheric Environment 192, 116-127.

Zoqi, M.J., 2024. Generalization of artificial neural network for predicting methane production in laboratory-scale anaerobic bioreactor landfills. Global J Environ Sci Manag 10(1), 225-244.

Zoqi, M.J., Ghamgosar, M., Ganji, M., Fallahi, S., 2016. Application of gmdh and genetic algorithm in fraction in biogas from landfill modeling. 18(3), 1-12.

Zoqi, M.J., Ghavidel, A., 2009. Neural network modeling and prediction of methane fraction in biogas from landfill bioreactors. Iranian Journal of Health and Environment 2(2), 140-149.

Zoqi, M.J., Saeedi, M., 2011. Modeling leachate generation using artificial neural networks. Journal of Water and Wastewater 77, 76-84.

Zoqi, M.J., Zoqi, T., Saeedi, M., 2010. Prediction of cod and nh4+-n concentrations in leachate from lab-scale landfill bioreactors using artificial neural networks. Journal of Water and Wastewater; Ab va Fazilab (in persian) 21(2), 52-60.