ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Fuzzy Inference Approach for $PM_{2.5}$ Modelling with High Accuracy and Low Complexity

Rati Wongsathan[1], Win Mya Thway[2] and Wutthichai Puangmanee[3]

ABSTRACT

$PM_{2.5}$ is a silent yet severe pollutant that accumulates in the human body, causing long-term health issues such as lung cancer. This study proposes a novel fuzzy inference system (FIS) for $PM_{2.5}$ forecasting, addressing the nonlinear and dynamic nature of air pollution. Unlike complex, data-intensive black-box models, the proposed FIS is transparent, interpretable, and simple to implement. It uses only two lagged $PM_{2.5}$ change rates and nine fuzzy rules for accurate prediction. The model requires no geographical or emission-source data, which are often costly and region-specific. Fuzzy rules are derived from natural $PM_{2.5}$ rise-and-fall patterns, ensuring logical consistency and minimal inputs. Using data from Chiang Mai, Thailand—one of the most polluted cities—the model was benchmarked against MLR, MLP, LSTM, SVM, and Gradient Boosting. The FIS achieved up to 5% higher accuracy. Although the Diebold-Mariano test found no significant difference, FIS showed comparable robustness with 49% fewer parameters and 56% fewer FLOPs. Optimal performance occurred at three input lags and 27 fuzzy rules, balancing accuracy and complexity. Moreover, the Chiang Mai FIS generalized well to other $PM_{2.5}$-affected cities—Bangkok, Jakarta, and Ho Chi Minh City—without modifications, maintained reliability for both daily and extended to hourly forecasts.

Article information:

Keywords: $PM_{2.5}$, Forecasting Model, Fuzzy Inference System, LSTM, Gradient Boosting

Article history:
Received: June 2, 2025
Revised: October 2, 2025
Accepted: October 9, 2025
Published: October 18, 2025 (Online)

DOI: 10.37936/ecti-cit.2025194.262425

1. INTRODUCTION

Over the past decade, fine particulate matter ($PM_{2.5}$) has become a major environmental and health concern, especially in developing and industrialized urban areas. Despite being preventable, $PM_{2.5}$ remains a heavy burden, mainly form combustion sources. Since 2019, agricultural burning-related $PM_{2.5}$ has caused about US$1.1 trillion in health damages. This is nearly double Thailand's 2019 GDP —nearly twice Thailand's 2019 GDP—and driven health disparities across 39 countries, underscoring the need for targeted mitigation policies [1].

Biomass combustion has intensified $PM_{2.5}$ pollution across Southeast Asia. From 2000–2019, rising hotspot activity correlated with higher respiratory disease mortality [2]. Our 2012–2024 analysis in Thailand confirms this pattern, except in 2022 due to effective fire control. Transboundary haze remains significant, with Sumatra identified as the main haze source in Kuala Lumpur via NAME simulations [3]. In Northern Thailand and Yunnan, WRF-Chem and HYSPLIT analyses attribute dry-season $PM_{2.5}$ to cross-border emissions from Myanmar and Lao PDR [4], [5]. $PM_{2.5}$ forecasting requires integrating physical, chemical, and meteorological factors, making it computationally demanding. Models such as WRF-Chem [6], HYSPLIT [7], and CTMs [8] simulate atmospheric dispersion and chemical transformations but face uncertainties in parameterization, input data, and resolution. Moreover, transboundary smoke and emission-transport mechanisms further complicate $PM_{2.5}$ prediction.

Approaches to $PM_{2.5}$ modelling differ across regions and frameworks. Classical multiple linear regression (MLR) has been applied in India [9], Pakistan [10], and Vietnam [11], while nonlinear general-

[1]The author is with the Department of Electrical Engineering, Faculty of Engineering and Technology, North-Chiang Mai University 169 Hang Dong, Chiang Mai, 50230 Thailand, Email: rati@northcm.ac.th
[2,3]The authors are with the Department of Software Engineering, Faculty of Engineering and Technology, North-Chiang Mai University 169 Hang Dong, Chiang Mai, 50230 Thailand, Email: winmyathway@ncuai.org and wutichai@northcm.ac.th
[1]Corresponding author: rati@northcm.ac.th

ized additive mixed models (GAMM) are used across the Northern Hemisphere [12]. Support vector regression (SVR) is common in China [13]. Neural networks (NNs) enhance prediction accuracy: MLP in India [9] and Pakistan [10], CNN in Vietnam [11], and LSTM or bi-LSTM in Jakarta [14], Liuzhou [15], and Vietnam [11]. Hybrid models [10] and ensemble methods—random forest (RF) [11], [15] and gradient boosting (GB) [16]—show strong performance. However, these black-box, data-driven models require extensive training, risk overfitting, and lack interpretability due to fixed, non-adaptive rules and implicit assumptions. In contrast, adaptive fuzzy inference systems (FIS), though underutilized in $PM_{2.5}$ forecasting, offer a compelling alternative. By using dynamically adjustable IF-THEN rules, FIS can achieve reliable predictions with minimal training data. Expert-defined rules allow post hoc error refinement, while the nonlinear fuzzy basis function (FBF) (ϕ_{fuzz}) provides flexible yet interpretable input–output mapping [17]-[18].

FIS-based models predict air quality index (AQI) by capturing nonlinear pollutant relationships, achieving over 89% accuracy with CO_x, SO_2, NO_x, PM_{10}, and O_3 [19]-[20]. Genetic algorithm (GA)-optimized FIS outperformed XGBoost and light gradient boost machine (LightGBM) for PM_{10} in Mexico [21], while grade-segmented FIS with fuzzy cognitive mapping estimated $PM_{2.5}$ and PM_{10} with lower complexity than ANFIS [22]. Hybrid FIS–tree models predicted indoor $PM_{2.5}$ [23]. Challenges remain in parameter optimization: evolutionary tuning may reduce interpretability, whereas grid-search hyperparameter selection is computationally intensive yet vital for generalizability.

This study presents a FIS $PM_{2.5}$ model with ϕ_{fuzz} and evaluates its performance against benchmark models: MLR with linear basis functions (ϕ_{lin}), SVR with kernel functions (ϕ_{SVR}), NNs using fuzzy-like activations (ϕ_{NN}), LSTM for temporal learning (ϕ_{LSTM}), and GB with decision tree outputs (ϕ_{GB}). The study focuses on Chiang Mai, northern Thailand, affected by local and transboundary biomass burning. This region experienced persistent $PM_{2.5}$ pollution for over two decades, ranking globally first in 2024 with levels exceeding 300 $\mu g/m^3$. Prolonged exposure led to rising hospitalizations and lung cancer cases [24], highlighting the need for prediction. Chiang Mai FIS $PM_{2.5}$ model generalizes to other polluted cities—Bangkok, Jakarta, and Ho Chi Minh City.

This study highlights cost-effective FIS forecasting. Benchmark models are trained on 2014–2018 data, whereas FIS relies on one year of pattern recognition with minimal training. Testing covers 2019–2024. FIS parameters are manually tuned, while benchmarks optimize errors. Performance is evaluated via R^2, MAE, RMSE, MAPE, MBE, MdAE, residual analysis (t-stat, p-values), and Diebold-Mariano Test. Complexity is assessed through parameter counts, inference operations, and asymptotic complexity $O(\cdot)$. Key contributions: (1) FIS achieves competitive $PM_{2.5}$ prediction with minimal inputs and pre-processing, suitable for resource-limited or real-time applications; (2) offers interpretable, rule-based transparency; (3) FIS rules reveal $PM_{2.5}$ patterns in Northern Thailand, showing increases during biomass burning, declines under mitigation, and stabilization during trend reversals; (4) FIS generalizes spatially and temporally, applicable to Bangkok, Jakarta, and Ho Chi Minh City, enabling reliable hourly forecasting.

2. METHODS AND MATERIALS

2.1 Study Area

Chiang Mai (CM), Northern Thailand, was selected for PM. modelling, having been studied extensively [25]–[31]. Surrounded by high mountains, the region traps air pollutants and borders Myanmar and Lao PDR. Intensive agricultural activity, combined with widespread open-field and forest burning starting in December, drives seasonal pollution from January to May. Cross-border fires exacerbate transboundary haze, the "smoke season," visible from dense fire hotspots (Fig. 1). With over 12 million residents, $PM_{2.5}$ poses serious public health risks. While Bangkok, Thailand's capital, experienced episodic $PM_{2.5}$ spikes in January 2025 (65–101 $\mu g/m^3$) from traffic and industry. In contrast, CM suffers sustained seasonal pollution with ∼217 days annually exceeding WHO guidelines. Peak in March–April often surpass 70 $\mu g/m^3$ and can reach 271 $\mu g/m^3$. Accurate $PM_{2.5}$ forecasting is therefore critical for effective long-term mitigation.

2.2 Data and Data-Preprocessing

Developing an effective $PM_{2.5}$ prediction model for Chiang Mai requires integrating diverse datasets, focusing on environmental and atmospheric parameters. Meteorological predictors such as temperature, pressure, gust wind, and relative humidity [25], [28], [31] were obtained from daily bulletins and historical records from the Chiang Mai Meteorological Department. Co-occurring air pollutants—CO, NO_x, and SO_2—were sourced from the Pollution Control Department (PCD) of Thailand. To assess generalization, $PM_{2.5}$ data from other regions were included: daily data from Bangkok and Jakarta, where sources include traffic, industry, and biomass burning, and hourly data from Ho Chi Minh City, Vietnam, where $PM_{2.5}$ mainly arises from construction and motorcycles, influenced by coastal factors [32]. These datasets enabled evaluation of the spatial and temporal performance of the FIS developed for daily $PM_{2.5}$ forecasting in Chiang Mai.

$PM_{2.5}$concentrations, the target variable, are typically measured by ground monitoring stations, but high installation and maintenance costs limit coverage. Portable electrostatic dust monitors (EDMs) achieve over 90% accuracy [33], and low-cost particulate matter sensors (LCPMS) using laser scattering reach up to 98% consistency with official PCD data [34]. These affordable sensors make $PM_{2.5}$ a practical standalone predictor, reducing reliance on expensive or error-prone datasets. Accurate forecasting using only historical $PM_{2.5}$ data—especially in regions with consistent seasonal patterns like Chiang Mai—offers a cost-effective, scalable solution. This study uses data from two official Chiang Mai stations, 35T (City Hall) and 36T (City High School), recording $PM_{2.5}$ from 2012 to 2024, initially as integers, with decimal precision added in 2023 (Fig. 2).

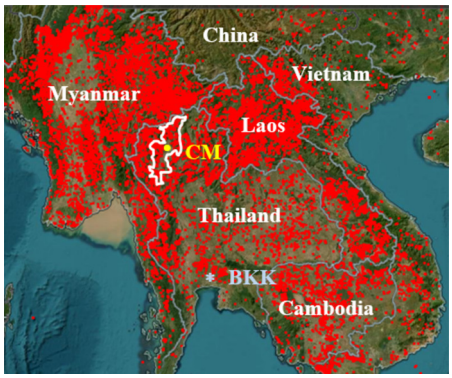


Fig.1: *MODIS AQUA Satellite-Detected Fire Hotspots on April 1, 2024, Indicating Transboundary $PM_{2.5}$ Haze over Chiang Mai and Surrounding Areas.*

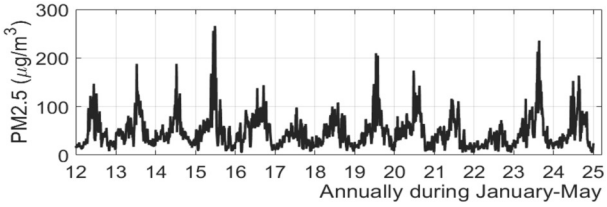


Fig.2: *Daily $PM_{2.5}$ Concentrations in Chiang Mai, Thailand, between 2012 and 2024.*

Daily $PM_{2.5}$ data from January–May 2012–2024 (∼1,967 points) were compiled, with ∼30 missing values filled via cubic interpolation (Appendix A). Typical models split data into training and testing subsets (commonly 70:30) to optimize parameters. In contrast, the proposed FIS requires only ∼10% of selected years, chosen for distinct $PM_{2.5}$ patterns—extremes, sharp peaks, or rapid changes—to construct fuzzy rules (see Section 3), reducing the need for extensive training while preserving predictive performance. Feature selection is unnecessary, as the model uses only past $PM_{2.5}$ values. To maintain interpretability, the number of lag inputs is limited to two, forming a simple three-variable model with the predicted $PM_{2.5}$, suitable for 3D visualization. Adding more variables would increase complexity. Unlike models using multiple predictors or machine learning algorithms—which require data normalization (e.g., Z-score)—the FIS relies solely on $PM_{2.5}$ and rule-based logic. Thus, normalization is unnecessary, and the focus is on capturing logical patterns in $PM_{2.5}$ behaviour through fuzzy rule construction, preserving transparency and simplicity.

2.3 FIS Framework for $PM_{2.5}$ Prediction Modeling

The $PM_{2.5}$ modelling (Fig. 3) employs a Mamdani-type FIS optimized for interpretability via minimal input variables and fuzzy sets, adhering to Zadeh's granularity principle to prevent overfitting. Temporal $PM_{2.5}$ values (e.g., $PM_{2.5}(t–2)$, $PM_{2.5}(t–1)$, $PM_{2.5}(t)$) are converted into rate changes, $\Delta PM_{2.5}(t)$ and $\Delta PM_{2.5}(t–1)$, which serve as fuzzy inputs. The output is the predicted increment $IncPM_{2.5}(t)$, used to compute $PM_{2.5}(t+1) = PM_{2.5}(t) + IncPM_{2.5}(t)$. Thus, the output of FIS model for $PM_{2.5}$ can be represented by the function f_{fuzz}: $[Min_{\Delta PM_{2.5}}, Max\Delta PM_{2.5}]^2 \rightarrow [Min_{IncPM_{2.5}}, Max_{IncPM_{2.5}}]$, which minimizes

$$\sum_{i=4}^{t}(\Delta PM_{2.5_{Actual}}[i] - f_{fuzz}[\Delta PM_{2.5}(i), \Delta PM_{2.5}(i-1)])^2. \quad (1)$$

For the inputs and output, their domains are covered with spanning P and Q membership functions (MFs), respectively, as $\mu_{\Delta PM_{2.58}(t),i}$ and $\mu_{\Delta PM_{2.58}(t-1),j}$, and $\mu_{IncPM_{2.5}(t),k}$, where $i, j = 1, \ldots, P$, and $k = 1, \ldots, Q$. MF values $\mu \in [0, 1]$. Three stages include fuzzification, rule evaluation, and defuzzification (Fig. 3). In fuzzification, MFs convert crisp input values into fuzzy $\mu_{PM_{2.5}(t)}$ and $\mu_{PM_{2.5}(t-1)}$, representing the input strength. Trapezoidal MFs (TrMFs), $\mu_{TrMF}(x; a, b, c, d)$, are used at domain boundaries, while triangular MFs (TMFs), $\mu_{TMF}(x; \alpha, \beta, \gamma)$, are used in central regions. MF parameters ($a, b, c, d, \alpha, \beta$, and γ) are design by trial-and-error with expert, including choice of P and Q.

In the rule evaluation, expert-defined fuzzy values analysis are fed into the rule base to construct logical mappings between input PM. trends and output increments of $PM_{2.5}$. The $P \times P$ fuzzy IF-THEN rules are formulated from expert knowledge as follows, Rule ij-th; If $\Delta PM_{2.5}(t)$ is A_i and $\Delta PM_{2.5}(t–1)$ is Bj THEN $IncPM_{2.5}(t)$ is C_{ij}, where A_i, and B_j, are input linguistic terms (e.g., Negative Big (NB), Zero Rate (Z), or Positive Big (PB)), and C_{ij} is the output linguistic terms (e.g., Decrease (D), Not Change (NC), or Increase (I)) corresponding to the respective MFs (i and j). The interaction of input fuzzy values through the antecedent part of each rule

yields the firing strength w_{ij}, computed using the minimum operator,

$$w_{i,j} = \min\left(\mu_{\Delta PM_{2.5}(t),i}, \mu_{\Delta PM_{2.5}(t-1),j}\right). \quad (2)$$

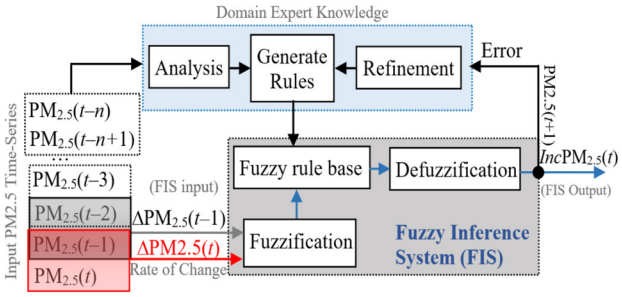


Fig.3: *The Fuzzy Inference System (FIS) for $PM_{2.5}$ Forecasting and Expert-Based Rule Generation.*

In the defuzzification, C_{ij} is represented by a crisp value $\delta PM_{2.5}(t)_{ij}$, and the overall incremental $PM_{2.5}$ output is computed using weighted average (3) with ϕ_{fuzz} given in (4). Thus, the fuzzy output can be expressed linearly in (5). The predicted $PM_{2.5}$ of the FIS model is given in (6).

$$IncPM_{2.5}(t) = \frac{\sum_{i=1}^{P}\sum_{j=1}^{P} w_{ij} \times \delta PM_{2.5}(t)_{ij}}{\sum_{i=1}^{P}\sum_{j=1}^{P} w_{ij}}, \quad (3)$$

where $w_{ij} = \mu_{\Delta PM_{2.5}(t),i} \cdot \mu_{\Delta PM_{2.5}(t-1),j}$ is the fuzzy weight.

$$\phi_{fuzz,ij} = \frac{\mu_{\Delta PM_{2.5}(t),i} \cdot \mu_{\Delta PM_{2.5}(t-1),j}}{\sum_{i=1}^{P}\sum_{j=1}^{Q} \mu_{\Delta PM_{2.5}(t),i} \cdot \mu_{\Delta PM_{2.5}(t-1),j}}. \quad (4)$$

$$IncPM_{2.5}(t) = \mathbf{\Phi}^T \times \delta PM_{2.5}(t)_{ij}, \quad (5)$$

where $\mathbf{\Phi} = [\phi_{fuzz,11}(\boldsymbol{X}), \ldots, \phi_{fuzz,PP}(\boldsymbol{X})]^T$ is the fuzzy basis function vector and $\boldsymbol{X} = [\Delta PM_{2.5}(t), \Delta PM_{2.5}(t\text{–}1)]^T$.

$$PM_{2.5}(t+1) = PM_{2.5}(t) + IncPM_{2.5}(t). \quad (6)$$

3. DESIGNING THE FIS FOR MODELLING $PM_{2.5}$

For FIS design, analysing $\Delta PM_{2.5}$ involves plotting a histogram and calculating key metrics such as range (R), standard deviation (SD), and mean. A single-peaked histogram near zero suggests three partitions—negative (N), zero (Z), and positive (P)—are sufficient. Multiple peaks or broader spreads may require five partitions (e.g., NB, NS, Z, PS, PB). The distribution shape guides fuzzy MF design: symmetrical distributions use equal partitions, asymmetrical distribution use unequal partitions, allocating smaller bins to denser regions. Smooth datasets require 3–5 fuzzy sets, with three for simplicity and five for precision, following Zadeh's granularity principle. K-means clustering can partition $\Delta PM_{2.5}$ into three or five clusters. A low SD/R ratio (< 0.2) indicates smooth trends, while a high SD/R (> 0.2) supports five partitions.

In designing a fuzzy prediction model for $PM_{2.5}$, training data from 2014–2018 and testing data from 2019–2024 are used. Cluster distributions for three and five partitions across 2014, 2016, and 2018 (Fig. 4) highlight differences in granularity, with dense central clustering suggesting finer granularity near the mean and broader partitions for sparse tails, as indicated by a primary peak and a smaller secondary peak to the right of zero.

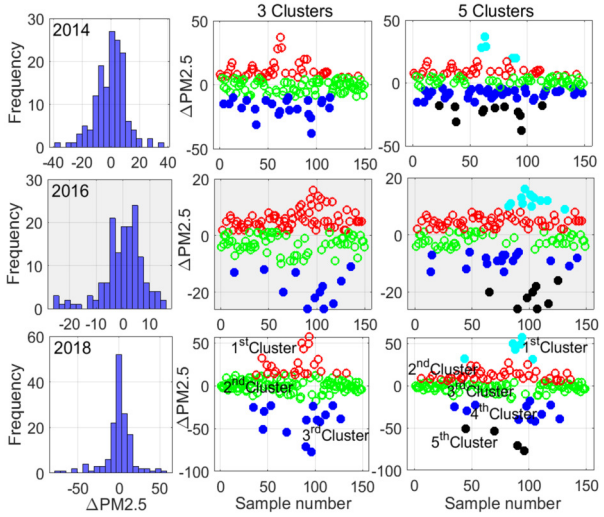


Fig.4: *Comparison of $\Delta PM_{2.5}$ Distributions and Cluster Partitions across 2014–2018, with Three Partitions Effectively Capture Key Trends with Simplicity.*

Variability metrics of $\Delta PM_{2.5}(t)$ indicate low variability, with $SD/R < 0.2$ in 2014 ($R = 75$, $SD = 10.9$, $SD/R = 0.14$), 2016 ($R = 134$, $SD = 18.19$, $SD/R = 0.13$), and 2018 ($R = 42$, $SD = 7.64$, $SD/R = 0.18$), supporting three partitions. Histograms and scatterplots reveal three main clusters—N, Z, and P—justifying the choice, while five partitions may overcomplicate the system. K-means clustering ($k = 3$) confirms clear groupings, balancing simplicity and accuracy for low-variability data. Based on these findings, two FIS models were developed. The natural rise-and-fall fuzzy (NRF-fuzzy) system follows the inherent temporal $PM_{2.5}$ trend, while the optimized NRF-fuzzy (ONRF-fuzzy) system introduces RMSE minimization while preserving interpretability. Together, these models evaluate baseline predictive performance and optimization benefits, ensuring robustness of the proposed fuzzy

framework for $PM_{2.5}$ forecasting.

For $\Delta PM_{2.5}(t)$ and $\Delta PM_{2.5}(t-1)$), Fig. 5 defines three MFs: NB (trapezoidal $[-40,-40,-10,-5]$) for significant decreases; Z (triangular $[-10,0,10]$) for minimal changes, and PB (trapezoidal [5, 10, 40, 40]) for significant increases. Trapezoidal shapes cover extreme ranges (NB, PB) and triangular shapes cover the central range (Z), balancing interpretability and precision. These MFs capture the symmetric distribution of $\Delta PM_{2.5}$ centred around zero, reflecting both variability and incremental $PM_{2.5}$ changes. For the output $IncPM_{2.5}(t)$, the NRF-fuzzy model uses five MFs: VD with TrMF $[-50,-50,-30,-20]$, SD with TMF $[-30,-15,0]$, Z with TMF $[-10,0,10]$, SI with TMF [0, 15, 30], and VI with TrMF [20, 30, 50, 50] (Fig. 5(b)). In contrast, the ONRF-fuzzy model simplifies $IncPM_{2.5}(t)$ to three TMFs: SD for significant decreases $[-50,-5,0]$, NC for minimal changes $[-15,0,15]$, and SI for significant increases [0, 5, 50] (Fig. 5(e)). Next, for designing fuzzy rules, a reasonable and interpretable set of 9 rules of NRF-fuzzy model to predict $PM_{2.5}(t+1)$=$PM_{2.5}(t)$+$IncPM_{2.5}(t)$ are presented in Table 1, for example: Rule 1: If $\Delta PM_{2.5}(t)$ is NB and $\Delta PM_{2.5}(t1)$ is NB, Then IncPM(t) is VD (Sharp decrease continues).

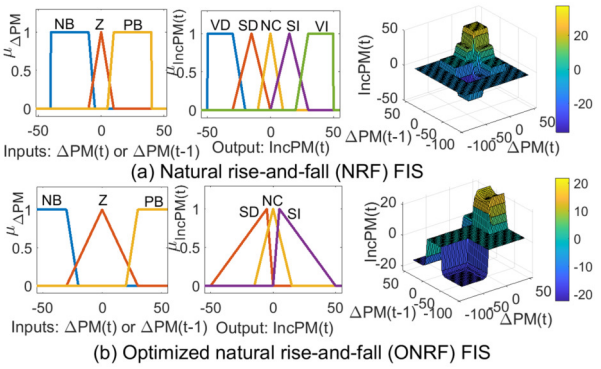


Fig.5: *Input and Output MFs, and 3D Surface Input-Output Plots of (a) NRF and (b) ONRF FIS.*

Table 1: *Comparison of Standard Rules and Trend-Modified Rules for Predicting Increment $PM_{2.5}$.*

	$\Delta PM(t)$ \ $\Delta PM(t-1)$	*NB*	*Z*	*PB*
Standard rules (NRF) for $IncPM_{2.5}(t)$	*NB*	*VD*	*SD*	*NC*
	Z	*SD*	*NC*	*SI*
	PB	*NC*	*SI*	*VI*
	$\Delta PM(t)$ \ $\Delta PM(t-1)$	*NB*	*Z*	*PB*
Trend-modified rules (ONRF) for $IncPM_{2.5}(t)$	*NB*	*SD*	*NC*	*SD*
	Z	*NC*	*NC*	*SD*
	PB	*NC*	*SI*	*NC*

Rules 1 and 9 predict strong decreases or increases (VD or VI) when both inputs align, while counteracting inputs, such as in Rules 3 and 7, predict stability (NC). Rule 5 captures stable conditions when inputs indicate no change, and intermediate rules (e.g., 2, 4, 6, 8) handle gradual transitions with moderate changes.

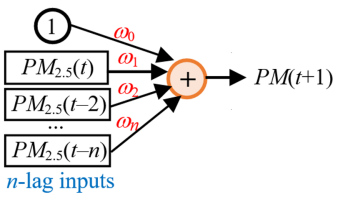
To enhance temporal representation, the NRF-fuzzy model is extended by using 3-input system with 27 rules. Each input variable $\Delta PM_{2.5}(t)$, $\Delta PM_{2.5}(t-1)$, and $\Delta PM_{2.5}(t-2)$ is described by three linguistic terms—NB, Z, and PB—while the output $IncPM_{2.5}(t)$ used five terms: VD, SD, NC, SI, and VI. Rule construction followed the NRF principle: when all inputs aligned (NB–NB–NB, PB–PB–PB), the system produced the strongest outcomes (VD, VI); conflicting signals yielded NC; and partial alignments resulted in moderate changes (SD, SI). For example, Rule 1 states: **If** $\Delta PM_{2.5}(t)$, $\Delta PM_{2.5}(t-1)$, and $\Delta PM_{2.5}(t-2)$ are {NB, NB, and NB}, **Then** $IncPM(t)$ is VD. While the full set of rules is summarized in Table 2.

Table 2: *FIS Rules with Three Input Lags with Three MFs.*

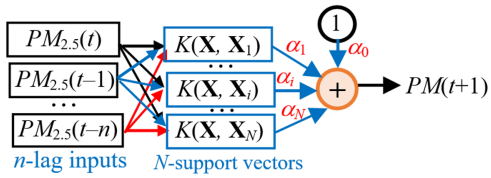
Rule	Lag of $\Delta PM_{2.5}$			*IncPM(t)*	Rule	Lag of $\Delta PM_{2.5}$			*IncPM(t)*
	t	*t–1*	*t–2*			*t*	*t–1*	*t–2*	
2	*NB*	*NB*	*Z*	*SD*	15	*Z*	*Z*	*PB*	*SI*
3	*NB*	*NB*	*PB*	*NC*	16	*Z*	*PB*	*NB*	*NC*
4	*NB*	*Z*	*NB*	*SD*	17	*Z*	*PB*	*Z*	*SI*
5	*NB*	*Z*	*Z*	*SD*	18	*Z*	*PB*	*PB*	*SI*
6	*NB*	*Z*	*PB*	*NC*	19	*PB*	*NB*	*NB*	*NC*
7	*NB*	*PB*	*NB*	*NC*	20	*PB*	*NB*	*Z*	*NC*
8	*NB*	*PB*	*Z*	*NC*	21	*PB*	*NB*	*PB*	*SI*
9	*NB*	*PB*	*PB*	*SD*	22	*PB*	*Z*	*NB*	*NC*
10	*Z*	*NB*	*NB*	*SD*	23	*PB*	*Z*	*Z*	*SI*
11	*Z*	*NB*	*Z*	*SD*	24	*PB*	*Z*	*PB*	*SI*
12	*Z*	*NB*	*PB*	*NC*	25	*PB*	*PB*	*NB*	*SI*
13	*Z*	*Z*	*NB*	*SD*	26	*PB*	*PB*	*Z*	*SI*
14	*Z*	*Z*	*Z*	*NC*	27	*PB*	*PB*	*PB*	*VI*

The main characteristics of the 27-rule base, compared with the previous 9-rule version using two lags, are summarized as follows. Rules 1 and 27 represent the strongest outcomes when all inputs align. Rules 3, 6, 7, 8, 12, 16, 19, 20, and 22 capture counteracting effects, yielding NC (no change). Mixed combinations of NB, Z, and PB produce SD or SI, depending on the dominant trend, while central cases such as Z–Z–Z or Z–Z–PB correspond to stable or small-change conditions.

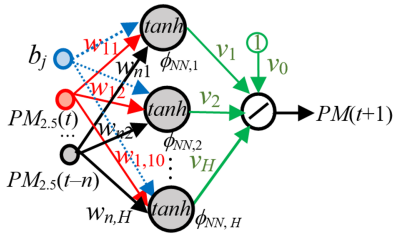
In the benchmark $PM_{2.5}$ prediction models (2012-2018 training data), Figures 6(a)–(e) illustrate the MLR, SVR, NN, LSTM, and GB models, respectively. The corresponding basis functions and output prediction expressions, Eq. (7)–(21), are summarized in Table 3.



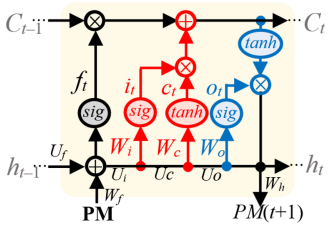
(a) MLR



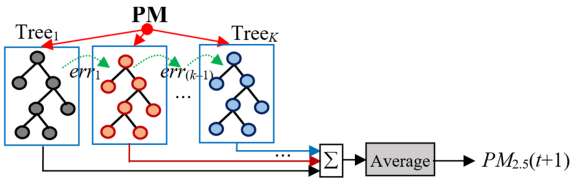
(b) SVR



(c) fully connected (dense) NN



(d) LSTM



(e) GB (using binary decision tree)

Fig.6: *The Benchmark PM2.5 Prediction Models.*

For MLR, the fuzzy basis function ϕ_{fuzz}, Eq. (4) is replaced by linear basis functions ϕ_{lin}, Eq. (7), representing lagged $PM_{2.5}$ features. Using Lasso (L1) regularization, the optimal subset of two lags is $\phi_{lin} = [PM_{2.5}(t), PM_{2.5}(t-5)]$, reflecting immediate local emissions and delayed transboundary haze effects. The coefficients (in Eq. (8)) are ω_0=3.04, ω_1=0.84, ω_5=0.084. For SVR (Eq. (10)), $\phi_{SVR}=K(\boldsymbol{X}, \boldsymbol{X_i})$ (Eq. (9)) with RBF kernels maps input to high-dimensional spaces. Lasso selects the same two lags (t and $t5$). Grid search tuned $C \in [0.001, 1000]$, the epsilon-insensitive loss $\varepsilon \in [0.001, 10]$, and the kernel scale $\sigma \in [0.01, 100]$, yielding C=1000, ε=0.001, and σ=100. The model used 698/1,050 support vectors, indicating high accuracy but potential overfitting due to complexity.

For NN, Eq. (12), ϕ_{fuzz} is replaced by activation functions $g(\cdot)$ applied to the weighted inputs, forming ϕ_{NN}, Eq. (11). Using two lagged $PM_{2.5}$ values, the model was tuned over hidden nodes $H \in [5, 30]$, learning rate $\eta \in [0.001, 0.2]$), and epochs $N_{epochs} \in [100, 1000]$) to balance bias-variance trade-offs. The

Table 3: *Benchmark $PM_{2.5}$ Prediction Models, Including Their Basis Functions and Output Prediction Expressions.*

Model	Basis function $\phi(x)$ and Output $PM_{2.5}(t+1)$
MLR	$\phi_{lin} = [PM2.5(t), PM2.5(t-5)]$, (7) $PM2.5(t+1) = \boldsymbol{\omega}^{\mathrm{T}} \phi_{lin} + \omega_0$. (8)
SVR	$\phi_{SVR} = K(\mathbf{X}, \mathbf{X}_i) = \exp\left(-\frac{\lVert \mathbf{X} - \mathbf{X}_i \rVert^2}{2\sigma^2}\right)$, (9) where input $\mathbf{X} = [PM2.5(t), PM2.5(t-5)]^{\mathrm{T}}$, and N support vectors $\mathbf{X}_i$, $PM2.5(t+1) = \sum_{i=1}^{N} \alpha_i \cdot \phi_{SVR} + \alpha_0$. (10)
NN	$\phi_{NN,j} = g(w_{1j} PM2.5(t) + w_{2j} PM2.5(t-1) + b_j)$,(11) $PM2.5(t+1) = \sum_{j=1}^{H} v_j \cdot \phi_{NN,j} + v_0$. 12)
LSTM	Gates and candidate: $f_t = \mathrm{sig}(W_f \mathbf{PM}_t + U_f h_{t-1} + b_f)$, (13) $i_t = \mathrm{sig}(W_i \mathbf{PM}_t + U_i h_{t-1} + b_i)$, (14) $c_t = \tanh(W_c \mathbf{PM}_t + U_c h_{t-1} + b_c)$, (15) $o_t = \mathrm{sig}(W_o \mathbf{PM}_t + U_o h_{t-1} + b_o)$. (16) Cell and hidden state update: $C_t = f_t \odot C_{t-1} + i_t \odot c_t$, (17) $\phi_{LSTM} = h_t = o_t \odot \tanh(C_t)$. (18) Output: $PM2.5(t) = W_h \cdot h_t + b_h$, (19) where $f, i, c, h, o, C \in \mathbb{R}^H$, weighted matrices $W \in \mathbb{R}^{H \times 2}$, bias vectors $b \in \mathbb{R}^H$, and $\odot$ is element-wise multiplication.
GB	$\phi_{GB,k} = \mathbf{Tree}_k([PM2.5(t), PM2.5(t-1)])$, (20) $PM2.5(t+1) = \frac{1}{K} \sum_{k=1}^{K} \phi_{GB,k}$. (21)

optimal configuration—$g(\cdot)$=tanh, H = 15, η = 0.01, and N_{epochs} = 500—achieved minimal RMSE on validation set.

For LSTM, integrating gating mechanisms and memory updates over time, the input vector **PM**=[$PM_{2.5}(t)$, $PM_{2.5}(t–1)$]T is processed through four gates: forget (f_t) (Eq. (13)), input (i_t) (Eq. (14)), candidate state (c_t) (Eq. (15)), and output (o_t) (Eq. (16)), each computed using weighted sums of the input **PM** and previous hidden state (h_{t-1}). The cell state (C_t) (Eq. (17)) updated by combining the previous stage and new candidate, modulated by the forget and input gates. The ϕ_{fuzz} is replaced by the LSTM output $\phi_{LSTM} = h_t$ (Eq. (18)), derived from the updated cell state and output gate, then passed through a linear layer producing the final prediction $PM_{2.5}(t+1)$ (Eq. (19)). Key hyperparameters: hidden units H, η, dropout (p_{drop}), and N_{epochs}. These tuned within $H \in [50, 200]$, $\eta \in [0.001, 0.01]$, $p_{drop} \in [0.2, 0.5]$, and $N_{epochs} \in [50, 200]$, balancing accuracy and overfitting. The optimal configuration—H = 100, p_{drop} = 0.2, η = 0.001, and Nepochs=100—yielded reduced

RMSE and validation loss. Finally, for GB, ϕ_{fuzz} is replaced by the outputs of sequentially trained decision trees forming ϕ_{GB} (Eq. (20)). Each tree corrects residual errors from the previous one, and their weighted sum provides the final $PM_{2.5}$ prediction (Eq. (21)). Model complexity and convergence depend on the number of trees (K), η, and maximum tree depth ($MaxDepth$). These were tuned over $K \in [50, 200]$, $\eta \in [10^{-3}, 1]$, and varying MaxDepth. The optimal setup, $\eta = 0.1$, $K = 50$, and $MaxDepth = 6$, yielded the lowest RMSE under 5-fold cross-validation.

4. RESULTS AND DISCUSSION

4.1 Direct Comparison of $PM_{2.5}$ Prediction Models

On the 2019–2024 test data, optimized FIS and tuned benchmarks (MLR, SVR, NN, LSTM, GB) were evaluated using R^2, RMSE, MAE, MAPE, MBE, and MdAE, with residuals tested under H: mean error = 0 via t-statistics and p-values (Table 4, Fig. 7). Overall, MLR, FIS, and SVR yielded the highest R^2 (0.76, 0.75, 0.75). For RMSE, MLR = 14.96, SVR = 15.33, FIS = 15.42. MAE favoured SVR, MLR, NN (≈ 9.8). MAPE remained low for SVR = 23.37, NN = 23.76, FIS = 23.76. Minimal bias appeared in FIS = 0.17, GB = 0.30, MLR = 0.42. Median errors were smallest for SVR = 5.76, NN = 5.93, GB = 6.09, confirming their stable, low-error prediction performance across years.

Residual analysis using t-statistics and p-values confirms statistically sound model performance. Under the null hypothesis (H), unbiased prediction requires a t-statistic near zero and p-value > 0.05. The FIS, SVR, and GB models met these conditions, showing negligible, statistically insignificant residuals. This highlights their robustness and adaptability. Overall, the proposed FIS achieved the most consistent accuracy across all metrics, while MLR and NN followed closely, demonstrating strong and competitive predictive capabilities across the 2019–2024 evaluation period.

Figure 8 compares the predictive performance of six models—MLR, SVR, NN, LSTM, FIS, and GB—for $PM_{2.5}$ forecasting. In each panel, red lines denote predicted values, and green circles indicate observed data. All models effectively capture overall trends and seasonal fluctuations in $PM_{2.5}$ concentrations.

While all models capture $PM_{2.5}$ trends, deviations occur at peaks—most notably in SVR and LSTM, which tend to over- or underpredict extremes, reflecting noise sensitivity and possible overfitting. NN and GB exhibit slight delays in high-concentration spikes, while MLR and FIS align closely with observed data, especially during rapid $PM_{2.5}$ fluctuations. FIS delivers the most stable and accurate peak responses, balancing trend-following and residual control. Overall, FIS ranks highest in robustness, followed by MLR,

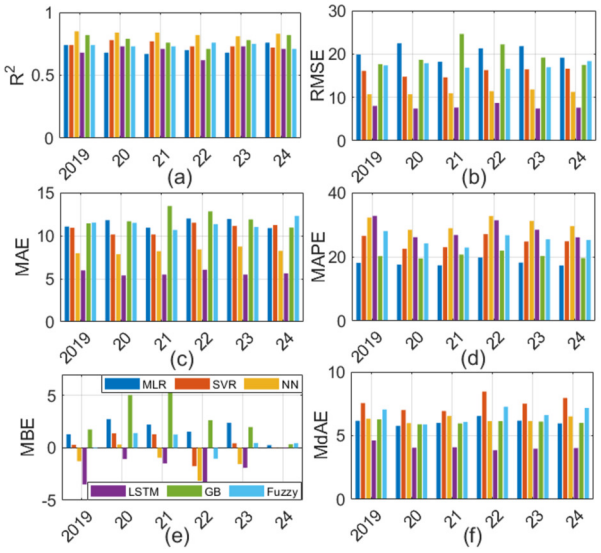


Fig. 7: *Performance Comparison of Various Models (MLR, SVR, NN, LSTM, GB, and FIS) Using R^2, RMSE, MAE, MAPE, MBE, and MdAE across the Years 2019–2024 for the Study Area of Chiang Mai, Thailand.*

Table 4: *Performance Metrics and Residual Analysis for Various Models, Including MLR, SVR, NN, LSTM, GB, and FIS Models, with Results Presented Numerically by Year for the Study Area of Chiang Mai, Thailand.*

Test data	Model	R^2	RMSE	MAE	MAPE	MBE	MdAE	t-stat	P-value
2019	MLR	0.74	19.88	11.06	18.11	1.29	6.16	0.79	0.43
	SVR	0.68	22.48	11.80	17.56	2.74	5.77	1.50	0.14
	NN	0.67	18.23	10.93	17.34	2.23	6.00	1.49	0.13
	LSTM	0.70	21.31	11.99	19.75	1.55	6.54	0.88	0.37
	GB	0.68	21.82	11.94	18.17	2.40	6.17	1.34	0.18
	FIS	0.76	19.16	10.88	17.31	0.26	5.96	0.16	0.87
2020	MLR	0.74	16.12	10.92	26.49	0.29	7.55	0.22	0.82
	SVR	0.78	14.77	10.14	22.50	1.39	7.01	1.15	0.25
	NN	0.77	14.63	10.14	22.98	1.29	6.93	1.08	0.28
	LSTM	0.73	16.29	11.50	27.05	-1.75	8.46	1.32	0.18
	GB	0.73	16.47	11.14	24.77	0.43	7.51	0.32	0.74
	FIS	0.72	16.62	11.24	24.83	0.02	7.96	0.02	0.98
2021	MLR	0.85	10.71	7.96	32.24	-1.28	6.33	1.47	0.14
	SVR	0.84	10.73	7.85	28.40	0.32	5.98	0.36	0.71
	NN	0.84	10.95	8.20	28.92	-0.94	6.54	1.05	0.29
	LSTM	0.82	11.43	8.41	32.72	-3.15	6.14	3.49	0.00
	GB	0.81	11.84	8.74	31.16	-1.55	6.15	1.61	0.11
	FIS	0.83	11.27	8.24	29.56	-0.01	6.51	0.02	0.98
2022	MLR	0.68	8.04	5.97	32.73	-3.49	4.62	5.87	0.00
	SVR	0.73	7.43	5.38	26.11	-1.07	4.05	1.77	0.08
	NN	0.71	7.64	5.48	26.78	-1.49	4.09	2.42	0.01
	LSTM	0.62	8.71	6.05	31.37	-3.80	3.85	3.75	0.00
	GB	0.73	7.43	5.48	28.41	-1.91	3.97	3.24	0.00
	FIS	0.71	7.62	5.63	26.05	-0.04	4.03	0.05	0.95
2023	MLR	0.82	17.66	11.43	20.19	1.75	6.28	1.21	0.23
	SVR	0.79	18.68	11.67	19.50	5.04	5.89	3.41	0.00
	NN	0.76	24.63	13.44	20.67	5.49	5.96	2.78	0.00
	LSTM	0.71	22.21	12.82	21.93	2.64	6.14	1.45	0.15
	GB	0.78	19.19	11.89	20.25	1.99	6.10	1.27	0.20
	FIS	0.82	17.49	10.94	19.58	0.34	6.00	0.23	0.81
2024	MLR	0.74	17.37	11.52	28.03	-0.05	7.05	0.04	0.97
	SVR	0.73	17.89	11.49	24.17	1.41	5.88	0.96	0.33
	NN	0.73	16.85	10.66	22.86	1.27	6.08	0.92	0.36
	LSTM	0.76	16.59	11.34	26.70	-1.04	7.27	0.76	0.45
	GB	0.75	16.94	11.01	25.45	0.46	6.61	0.33	0.74
	FIS	0.71	18.37	12.30	25.24	0.44	7.17	0.29	0.77

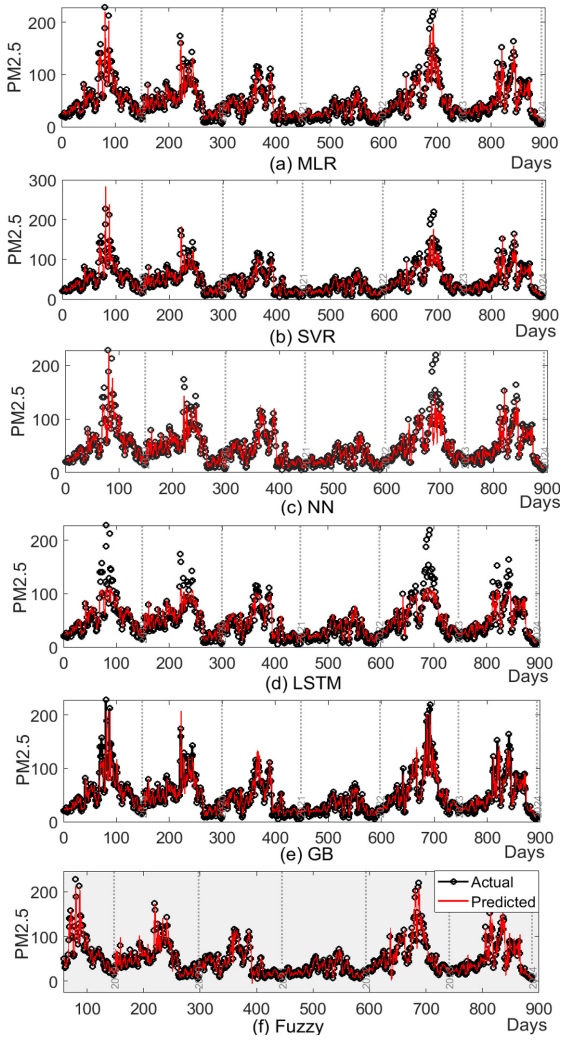


Fig.8: *Comparison of Actual versus Predicted $PM_{2.5}$ from 2019 to 2024 using (a) MLR, (b) SVR, (c) NN, (d) LSTM, (e) GB, and (f) FIS Models in the Study Area, Chiang Mai, Thailand.*

NN, and GB performing moderately, and SVR and LSTM showing higher variability.

4.2 Pairwise $PM_{2.5}$ Model Comparison via DMT

The Diebold-Mariano Test (DMT) [35] statistically compares the forecasting accuracy of two models over short horizons by evaluating differences in their error distributions. In this study, it assesses six $PM_{2.5}$ prediction models. The loss differential d represents the performance gap between models, defined as either squared-error $d = ({e_1}^2 - {e_2}^2)$ or absolute error $d = |e_1| - |e_2|$. depending on the nature of the evaluation. The DM statistic is subsequently computed using the mean and variance of d as

$$DM_{stat} = \frac{\text{mean}(d)}{\sqrt{\text{Var}(d)/N}}, \tag{22}$$

where N represents the sample size.

The DMT evaluates the null hypothesis H_0: $E(d) = 0$. A large absolute DM_{stat} or $p < 0.05$ rejects H0, confirming significant accuracy differences. As shown in Table 5, the FIS shows a slight advantage over others, but high p-values (> 0.05) indicate no statistically significant difference, implying comparable predictive accuracy. The only exception is the GB model, whose negative DM_{stat} and $p < 0.05$ confirm statistical significance over MLR. For all other pairs, $p > 0.05$ suggests no meaningful differences. Overall, the FIS model performs competitively, neither outperforming nor underperforming others.

Table 5: *Diebold-Mariano Test Statistics and p-Values (DM_{stat}/p-Value) for Pairwise Comparisons among MLR, SVR, NN, LSTM, GB, and FIS models.*

Model	MLR	SVR	NN	LSTM	FIS	GB
MLR	-	-1.25/0.21	-1.32/0.18	-1.27/0.20	-0.19/0.84	-2.10/0.03
SVR	1.25/0.21	-	-0.78/0.43	-0.37/0.71	0.94/0.34	-0.34/0.73
NN	1.32/0.18	0.78/0.43	-	0.76/0.45	1.18/0.23	0.61/0.54
LSTM	1.27/0.20	0.37/0.71	-0.76/0.45	-	0.93/0.35	0.17/0.86
FIS	**0.19/0.84**	**-0.94/0.34**	**-1.18/0.24**	**-0.93/0.35**	**-**	**-1.29/0.19**
GB	2.10/0.03	0.34/0.73	-0.61/0.54	-0.17/0.86	1.29/0.19	-

4.3 Comparison of Computational Complexity

Assessing model complexity (Table 6), the MLR with $p = 2$ inputs—$PM_{2.5}(t-1)$ and $PM_{2.5}(t-5)$—and one output requires 3 parameters $(\omega_1, \omega_5, \omega_0)$ and each prediction involves 2 multiplications, 2 additions, and 1 bias addition (Eq. (8)), totalling $\approx 2p$ FLOPs, whereas the SVR, with $N = 96$ support vectors, has 97 parameters (α_i) (Eq. (10)) and for each support vector computes p subtractions, p squarings, $p-1$ additions for the Euclidean norm, and 1 exponential operation yielding$3p+1$ FLOPs per vector and $N \cdot (3p+1)$ FLOPs (Eq. (9)); consequently, MLR offers minimal complexity $O(p)$ and high interpretability, while SVR scales as $O(N \cdot p)$ with greater expressiveness. For a NN with $p = 2$, one hidden layer of $H = 15$ neurons, and one output, the total parameters ($\boldsymbol{w}$'s and $\boldsymbol{b}$'s) are $(p+1) \cdot H + H + 1$, with each hidden neuron performing $2p+1$ FLOPs and the output node $2H$ FLOPs, yielding a total inference cost of $(2p+1) \cdot H + 2H(O(H \cdot p))$. For the LSTM, four gates per cell has $4[(p+1) \cdot H + H^2]$ parameters (Eq. (13)-(16)) and $4H \cdot (p+H) + 10H$ FLOPs $(O(H \cdot (p+H)))$, while dropout ($p_{drop} = 0.2$) affects training only. For the GB model, the total number of parameters across K trees is approximately $K \cdot (2^{MaxDepth} - 1)$ and the FLOPs is $\approx K \cdot Max_{Depth} + K$ or $O(K \cdot Max_{Depth})$. For the proposed FIS with p inputs, each with P MFs ($TrMF$ trapezoidal and TMF triangular), the total parameters are $p \cdot (3P + TrMF) + 3TMF$, and inference requires $3p \cdot P + 3Rule + 1$ FLOPs, where $Rule = P^p$. This scale asymptotically as $O(p \cdot P + P^p)$, capturing both input/output fuzzification and exponential rule growth.

Table 6: *Comparison of $PM_{2.5}$ Model Complexity in terms of Parameter Counts, Floating-Point Operations (FLOPs), and Asymptotic Complexity $O(\cdot)$.*

Model	Parameter Counts (p=2, all models)	Operation Counts (FLOPs)	$O(\cdot)$
MLR	p+1=3	$2p$ =4	$O(p)$
SVR	N+1=97	~$N\cdot(3p+1)$=672	$O(p\cdot N)$
NN	$p\cdot H+2H+1$=61	$2p\cdot H+3H$=105	$O(p\cdot H)$
LSTM	$4[(p+1)\cdot H+H^2]$ =41,200	~$4H\cdot(p+H)+10H$ =41,800	$O(p\cdot H+H^2)$
GB	$K\cdot(2^{MaxDepth}-1)$ =3,150	$K\cdot(Max_{Depth}+1)$ =350	$O(K\cdot Max_{Depth})$
FIS	$p\cdot(3P+TrMF)+3TMF$ =31	$3P\cdot p+3Rule+1$ =46	$O(p\cdot P+Rule)$

4.4 Optimal FIS structure

For the FIS, increasing the number of inputs (p, lagged $\Delta PM_{2.5}$) or the number of MFs (P) raises total parameters, rules, and FLOPs. These factors must be considered to avoid excessive complexity. Table 7 summarizes these values for $p = 2-10$ and $P = 3$, 5, and 7. The parameter count grows linearly with p and P, following $O(pP + P^p)$ with exact expression of $3pP + 3P^p + 1$, while the number of fuzzy rules and inference cost grow exponentially as Rule $=P^p$. Configurations with $P = 3$ remain tractable for moderate input sizes ($p \leq 10$), whereas $P \geq 5$ become demanding for $p \geq 4-6$, and $P = 7$ is infeasible beyond small p. These findings emphasize balancing model granularity (P) and computational tractability (p) when designing FIS models. These trade-offs are summarized in Tables 8, guiding selection of optimal lags and MFs for $PM_{2.5}$ forecasting, with results for 2024 shown in Fig. 9.

Examining input lags shows that increasing p from 2 to 3 improves R^2 from ≈0.82-0.93 → 0.98-0.99 and reduces RMSE, MAE, and MAPE, while further increases to 4 or 5 lags degrade performance, indicating three lags optimally balance pattern capture and overfitting. Increasing MFs from $P = 3$ to 5 at the same lag improves R^2 and reduces errors, but excessive MFs (e.g., $p = 3$, $P = 5$ or $p = 4$, $P = 5$) raise complexity without significant gains, risking overfitting. Comparisons of 2023 and 2024 data confirm stability, with $p = 3$ and $P = 3$ yielding $R^2 \approx 0.98-0.99$, RMSE≈ $4.8 - 4.9$, MAE≈ $1.5 - 1.6$. Figure 9 illustrates these effects: increasing lags from 2 to 3 (Fig. 9(a)-(c)) substantially improves alignment with target $PM_{2.5}$, while further lags (Fig. 9(d),(f)) add little benefit and may slightly overfit. Increasing MFs from 3 to 5 (Fig. 9(b) vs. (a), Fig. 9(e) vs. (c)) enhances flexibility but adds complexity, with diminishing returns for higher rule counts (e.g., 125). Table 8 summarizes these results, highlighting the optimal FIS configuration.

Therefore, the optimal FIS structure for PM. forecasting is $p = 3$ lags and $P = 3$ MFs (27 rules), achieving $R^2 \approx 0.99$ while minimizing overfitting and maintaining strong generalization. Compared with $p = 2$, $P = 3$ ($R^2 \approx 0.82$), it reduces RMSE, MAE, and MAPE, better capturing temporal and nonlinear patterns. This configuration increases parameters by ∼35%, rules ×3, and FLOPs by ∼50%, yet remains computationally efficient.

Table 7: *Parameters and FLOPs by Lag and MFs.*

Lag (p)	MFs (P)	Input param (p^{3P+2})	Output Param. ($3P$)	Total Param.	Rule (P^p)	Flops
2	3	24	9	33	9	46
	5	36	15	51	25	106
	7	48	21	69	49	190
3	3	36	9	45	27	109
	5	54	15	69	125	421
	7	72	21	93	343	1093
5	3	60	9	69	243	734
	5	104	15	119	3125	9401
	7	148	21	169	16807	504,426
8	3	96	9	105	6561	19,695
	5	176	15	191	390625	1,171,876
	7	256	21	277	5,764,801	17,294,406
10	3	120	9	129	59049	177,148
	5	224	15	239	9,765,625	29,296,876
	7	336	21	357	282,475,249	847,426,748

Table 8: *Analysis of FIS Structure Complexity versus Forecasting Error Metrics for $PM_{2.5}$ Forecasts.*

year	FIS Structure			Performance					
	Lags	MFs	Rule	R^2	RMSE	MAE	MAPE	MBE	MdAE
2023	2	3	9	0.82	17.49	10.94	19.58	0.34	6.00
	2	5	25	0.93	10.17	8.50	18.81	-1.83	9.52
	3	3	27	0.99	4.93	1.62	4.69	0.81	0.00
	3	5	125	0.96	7.82	5.87	12.82	-0.98	5.15
	4	3	81	0.89	13.23	10.02	23.04	-2.82	10.22
	4	5	625	0.97	6.38	4.48	10.92	1.05	4.11
	5	3	243	0.95	9.03	5.60	12.91	-1.78	0.00
2024	2	3	9	0.71	18.37	12.30	25.24	0.44	7.17
	2	5	25	0.91	9.898	7.59	22.54	-0.27	7.55
	3	3	27	0.98	4.79	1.53	2.77	-0.71	0.0
	3	5	125	0.95	7.49	5.51	16.62	-0.72	5.36
	4	3	81	0.87	11.87	8.98	30.18	-1.16	9.30
	4	5	625	0.96	6.03	4.43	16.37	1.47	4.43
	5	3	243	0.93	8.74	6.03	22.28	-0.58	4.75

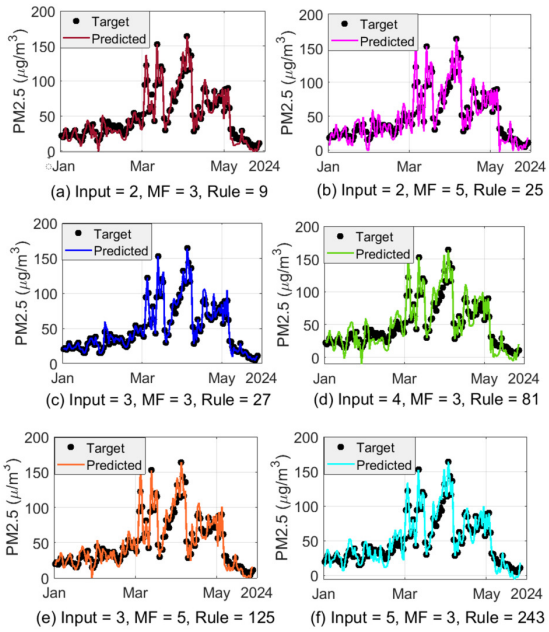


Fig.9: *Chiang Mai $PM_{2.5}$ Forecasts (Jan–May 2024) by FIS; Input 3 & MF 3 (27 Rules) Gives Optimal Accuracy.*

4.5 Comparison of Model Performance

This subsection compares the FIS $PM_{2.5}$ model with prior studies (Table 9). In India [9], FIS achieves comparable R^2 to MLR and NN but with a wider RMSE range; its simplicity offers a practical alternative. In Pakistan [10], FIS outperforms both MLR and NN in RMSE and MAE while maintaining similar MAPE. In Ho Chi Minh City, Vietnam [11], FIS surpasses MLR and LSTM, with higher R^2, lower errors, and reduced computational complexity, demonstrating robust, efficient $PM_{2.5}$ prediction. The FIS model shows competitive R^2 compared with GAMM for the Northern Hemisphere [12], despite slightly higher RMSE and MAE, and does not require detailed atmospheric data, highlighting adaptability. In Jakarta [14], FIS outperforms LSTM with meteorological features in MAE while using fewer inputs. For Liuzhou, China [15], LSTM with wavelet-transformed inputs exceeds FIS in RMSE, MAE, and MAPE, but FIS achieves comparable MAPE without intensive pre-processing. In Central and Bangkok, Thailand [16], Light GB surpasses FIS in RMSE and MAE, yet FIS maintains competitive R^2 with simpler, more interpretable, and computationally efficient design.

4.6 Comparison of Model Performance

To test generalization, the Chiang Mai FIS (nine-rule ORNF system) was applied without retraining to Bangkok, Thailand; Jakarta, Indonesia; and Ho Chi Minh City, Vietnam (Fig. 10). In Bangkok, FIS achieved $R^2 \approx 0.70$, MAE=5.48μg/m^3, RMSE=7.98, and MAPE≈21.6%, capturing ∼70% of daily variance with slightly higher RMSE than MAE due to occasional peaks. Jakarta showed higher absolute errors (MAE=9.17, RMSE=11.06), lower relative error (MAPE15.9%), negative mean bias (MBE ≈-2.12), and moderate $R^2 \approx 0.58$, reflecting higher volatility. In Ho Chi Minh City, daily-trained FIS applied to hourly data yielded MAE=3.96, RMSE=8.85, $R^2 \approx 0.62$, and MAPE≈22.3%, demonstrating adaptability across temporal resolutions. Statistical tests confirmed unbiased predictions across all sites: Bangkok ($t = 0.84$, $p = 0.40$), Jakarta ($t = 0.11$, $p = 0.36$), Ho Chi Minh City ($t = 0.23$, $p = 0.32$), highlighting robustness and reliability for $PM_{2.5}$ forecasting in diverse spatial and temporal contexts

4.7 Limitation and Further Suggestions of This Study

The study's main limitation is its reliance on only $PM_{2.5}$ data with two lags ($PM_{2.5}(t)$ and $PM_{2.5}(t-1)$), excluding other variables such as temperature, PM_{10}, wind speed, ozone, and toxic gas levels, which could enhance forecasting accuracy. To improve the model, incorporating hotspot count or changes as inputs could better capture local environmental factors, such as dust sources from biomass burning and forest fire in Chiang Mai, Thailand. Additionally, optimizing the parameters of MFs and fuzzy rules through techniques like GA or particle swarm optimization

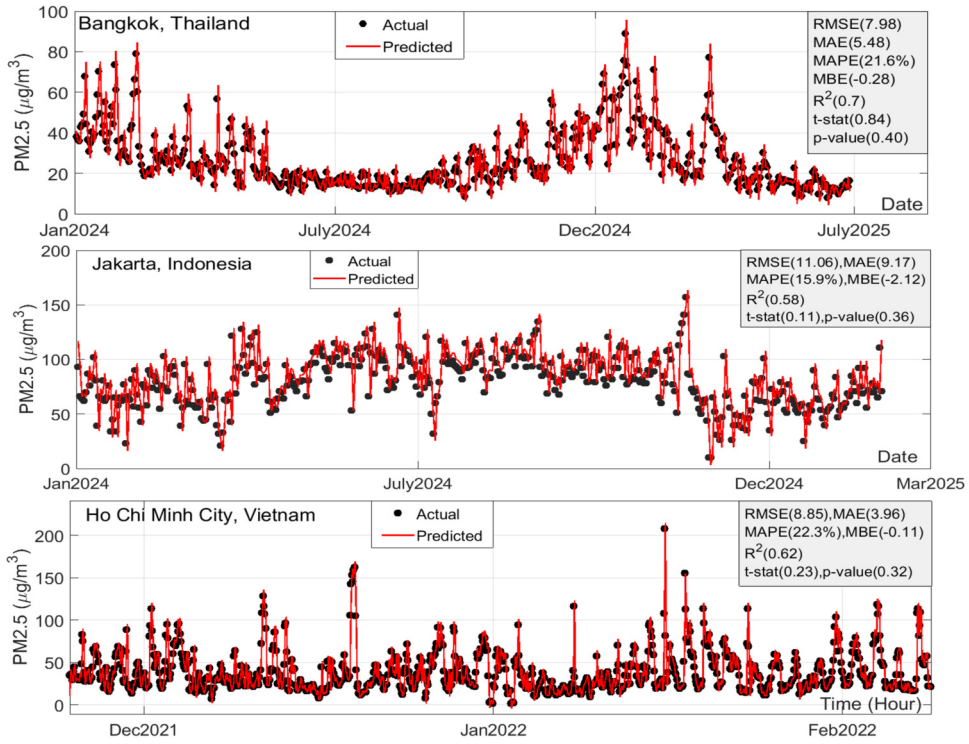


Fig.10: *Generalization Testing of the FIS for Temporal and Spatial $PM_{2.5}$ Prediction in (top) Bangkok, Thailand, (middle) Jakarta, Indonesia, and (bottom) Ho Chi Minh City, Vietnam.*

Table 9: *Summary of Prior $PM_{2.5}$ Prediction Studies.*

Study areas/Ref.	Model	Inputs	Error Metric (Testing)
India [9]	MLR	PM10, NO_2, SO_2	R^2 (0.72), RMSE (11.76)
	NNs		R^2 (0.75-0.8), RMSE (9.5-11)
Pakistan [10]	MLR	Lagged $PM_{2.5}$	RMSE(22.552.3),MAPE(0.17 0.22), MAE (16.8–36.8)
	NN		RMSE(24–58), MAPE(0.18–0.25), MAE (16.7– 40.3)
Ho Chi Minh City, Vietnam[11]	MLR	5 lagged $PM_{2.5}$	R^2(0.6-0.7), RMSE(8.1-8.29), MAE(5.4-5.6),MAPE(30-40)
	LSTM (H=64)		R^2(0.6-0.7), RMSE(8.11-8.27), MAE(5.36-5.5),MAPE(29-37)
Northern Hemisphere [12]	GAMM	near-surface atmospheric visibility	R^2 (0.8), RMSE (15), MAE (8)
Jakarta, Indonesia [14]	LSTM (H_1=20, H_2=10)	Lagged $PM_{2.5}$, meteorology	RMSE (18.5-19.4), MAE (14.9-15.6)
Liuzhou, China [15]	LSTM	Lagged $PM_{2.5}$	RMSE (4.46), MAE (3.81), MAPE (0.17)
Bangkok, Thailand[16]	Light GB	Lagged $PM_{2.5}$, meteorology	R^2 (0.8-0.86), RMSE (5.93–6.68), MAE (3.98–4.43)

(PSO) could improve the model's adaptability to nonlinear patterns and regional data.

Future developments should consider the complexity of real-world interactions, where $\Delta PM(t)$ and $\Delta PM(t-1)$ may not follow linear patterns, especially when key external factors are missing. Adding explanatory variables could better capture these interactions and enhance accuracy. Care is needed to avoid overfitting the fuzzy rule base, which may reduce generalizability. Optimization algorithms for parameter selection and rule generation could refine the model. Expanding this time-series approach to other regions and exploring advanced techniques such as deep learning may further improve predictive accuracy and the overall performance.

5. CONCLUSIONS

This study introduced an innovative FIS model for $PM_{2.5}$ pollution prediction, prioritizing simplicity, interpretability, and independence from geographically specific inputs. Utilizing lagged $PM_{2.5}$values and expertly designed fuzzy rules, the model achieves comparable performance to advanced benchmarks, including NNs, LSTM, SVR, and GB, while avoiding extensive data pre-processing or computational complexity. Its evaluation across error metrics highlights consistent reliability and adaptability, with notable success in volatile regions like Chiang Mai, Thailand. Pairwise comparisons using the DMT confirmed that the FIS model performs comparably to other methods, with no statistically significant differences in predictive accuracy. Among nonlinear models, the FIS also offers the most favourable trade-off between accuracy and computational efficiency, maintaining significantly lower complexity. Moreover, the Chiang Mai FIS model shows strong robustness and generalization across spatial and temporal domains. Applied directly to Bangkok, Jakarta, and Ho Chi Minh City (hourly data), it maintained accurate predictions, demonstrating its practical applicability for diverse $PM_{2.5}$ monitoring and forecasting.

Despite its strengths, the model excludes meteorological and exogenous variables influencing $PM_{2.5}$ levels in SEA countries. Future research should extend to other regions and integrate such factors—particularly hotspot counts—while applying optimization techniques to enhance fuzzy parameters. Expanding studies geographically will further validate the FIS's robustness. This study positions the FIS as a practical, interpretable foundation for advancing air quality forecasting.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to North-Chiang Mai University for providing research funding to support the preparation of this article.

AUTHOR CONTRIBUTIONS

Conceptualization, R. Wongsathan; methodology, R. Wongsathan; software, R. Wongsathan and W. Puangmanee; validation, R. Wongsathan, W.M. Thway, and W. Puangmanee; formal analysis, R. Wongsathan; investigation, R. Wongsathan and W.M. Thway; data curation, R. Wongsathan; writing—original draft preparation, R. Wongsathan; writing—review and editing, R. Wongsathan and W.M. Thway; visualization, R. Wongsathan and W. Puangmanee; supervision, R. Wongsathan; funding acquisition, W.M. Thway. All authors have read and agreed to the published version of the manuscript.

References

[1] H. Yin *et al.*, "Global health costs of ambient PM2.5 from combustion sources: A modelling study supporting air pollution control strategies," *Lancet Planet. Health*, vol. 8, no. 7, pp. e476–88, Jul. 2024.

[2] T. Amnuaylojaroen and N. Parasin, "Perspective on Particulate Matter: From Biomass Burning to the Health Crisis in Mainland Southeast Asia," *Toxics*, vol. 11, no. 7, pp. 1–14, Jul. 2023.

[3] J. Murulitharan, "Assessing PM2.5 pollution and the role of transboundary pollution emissions across Greater Kuala Lumpur, Malaysia from 2018 to 2023," M.S. thesis, Apollo - Univ. Cambridge Repository, Cambridge, U.K., 2025. [Online]. Available: `https://doi.org/10.17863/CAM.117887`

[4] K. Inlaung, C. Chotamonsak, R. Macatangay and V. Surapipith, "Assessment of Transboundary PM2.5 from Biomass Burning in Northern

Thailand Using the WRF-Chem Model," *Toxics*, vol. 12, no. 7, pp. 1–15, Jul. 2024.

[5] Y. Luo, H. Wei and K. Yang, "The impact of biomass burning occurred in the Indo-China Peninsula on PM2.5 and its spatiotemporal characteristics over Yunnan Province," *Sci. Total Environ.*, vol. 908, pp. 168174, 2024.

[6] J. Hong, F. Mao, Q. Min, Z. Pan, W. Wei, T. Zhang, and W. Gong, "Improved $PM_{2.5}$ predictions of WRF-Chem via the integration of Himawari-8 satellite data and ground observations," *Environ. Pollut.*, vol. 263, p. 114451, 2020.

[7] S. Raffuse, S. O'Neill and R. Schmidt, "A model for rapid PM. exposure estimates in wildfire conditions using routinely available data: Rapidfire v0.1.3," *Geosci. Model Dev.*, vol. 17, pp. 381–397, 2024.

[8] B. T. Dinkelacker, P. G. Rivera, J. D. Marshall, P. J. Adams and S. N. Pandis, "High-resolution downscaling of source resolved $PM_{2.5}$ predictions using machine learning models," *Atmos. Environ.*, vol. 310, p. 119967, 2023.

[9] S. Gulati *et al.*, "Estimating PM2.5 utilizing multiple linear regression and ANN techniques," *Sci. Rep.*, vol. 13, p. 22578, 2023.

[10] H. Iftikhar, M. Qureshi, J. Zywiołek, J. L. López-Gonzales and O. Albalawi, "Short-term PM2.5 forecasting using a unique ensemble technique for proactive environmental management initiatives," *Front. Environ. Sci.*, vol. 12, p. 1442644, 2024.

[11] T. N. T. Nguyen, T. D. Trinh, P. C. L. T. Vu and P. T. Bao, "Statistical and machine learning approaches for estimating pollution of fine particulate matter (PM2.5) in Vietnam," *J. Environ. Eng. Landsc. Manag.*, vol. 32, no. 4, pp. 292–304, 2024.

[12] H. Hao, K. Wang, G. Wu, J. Liu, and J. Li, "PM2.5 concentrations based on near-surface visibility in the Northern Hemisphere from 1959 to 2022," *Earth Syst. Sci. Data*, vol. 16, pp. 4051–4076, 2024.

[13] Z. Liu, X. Huang and X. Wang, "PM2.5 prediction based on modified whale optimization algorithm and support vector regression," *Sci. Rep.*, vol. 14, p. 23296, 2024.

[14] T. Istiana, B. Kurniawan, S. Soekirno, A. Wihono, D. E. Nuryanto, B. A. Pertala, and A. Sopaheluwakan, "Fine particulate matter concentration forecasting using long short-term memory network and meteorological inputs," *Glob. J. Environ. Sci. Manag.*, vol. 10, no. 4, pp. 1759–74, 2024.

[15] X. Hu, J. Shi, C. He, and J. Fang, "Combined prediction model of PM2.5 concentration based on wavelet transform and LSTM," *J. Phys.: Conf. Ser.*, vol. 2555, p. 012009, 2023.

[16] S. Sukprasert, H. Shimadera, S. Araki, T. Thongthammachart, P. Thanasutives, L. V. Mui, T. Matsuo, and A. Kondo, "Improving machine learning-based PM2.5 prediction by segregating biomass emission factor from chemical transport model," *E3S Web Conf.*, vol. 530, p. 01004, 2024.

[17] H. M. Kim and J. M. Mendel, "Fuzzy basis functions: comparisons with other basis functions," *IEEE Trans. Fuzzy Syst.*, vol. 3, no. 2, pp. 158–168, May 1995.

[18] E. Egrioglu, E. Bas, and M. Y. Chen, "A fuzzy Gaussian process regression function approach for forecasting problem," *Granul. Comput.*, vol. 9, p. 47, pp. 1–11, 2024.

[19] M. Saleem, N. Shingari, M. S. Farooq, B. Mago, and M. A. Khan, "Real-time air quality monitoring model using fuzzy inference system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 6, pp. 838–846, Jun. 2024.

[20] Z. N. Zulkiflee and A. Idris, "Fuzzy inference system model for air quality index prediction," in *Proc. Sci. Math.*, vol. 23, pp. 133–143, 2024.

[21] F.-J. Moreno-Vazquez, F. Trujillo-Romero and A. E. Violante-Gavira, "Enhanced Fuzzy Inference System for PM10 Concentration Prediction Using Genetic Algorithms," in *Proc. 21st Int. Conf. Electr. Eng., Comput. Sci. Automat. Control (CCE)*, Mexico City, Mexico, pp. 1–6, 2024.

[22] Z. S. Mousavi Fard, H. Asilian Mahabadi, F. Khajehnasiri and M. A. Rashidi, "Modeling the concentration of suspended particles by fuzzy inference system (FIS) and adaptive neuro-fuzzy inference system (ANFIS) techniques: A case study in the metro stations," *Environ. Health Eng. Manag. J.*, vol. 10, no. 3, pp. 311–319, Sep. 2023.

[23] J. Saini, M. Dutta and G. Marques, "Modeling indoor PM. using Adaptive Dynamic Fuzzy Inference System Tree (ADFIST) on Internet of Things-based sensor network data," *Internet Things*, vol. 20, pp. 1–13, 2022.

[24] T. Supasri *et al.*, "Association between ambient air particulate matter and human health impacts in northern Thailand," *Scientific Reports*, vol. 13, p. 12753, pp. 1–15, 2023.

[25] R. Wongsathan, I. Seedadan and M. Kavilkrue, "Prediction modeling of PM-10 in Chiangmai City Moat by using artificial neural networks," *Appl. Mech. Mater.*, vol. 781, pp. 628–631, 2015.

[26] R. Wongsathan and I. Seedadan, "A hybrid ARIMA and neural networks model for PM-10 pollution estimation: The case of Chiang Mai City Moat area," *Procedia Comput. Sci.*, vol. 86, pp. 273–276, 2016.

[27] R. Wongsathan, I. Seedadan and S. Wanasri, "Hybrid forecast models for PM-10 prediction: A case study of Chiang Mai city of Thailand

during high season," *Eng. Appl. Sci. Res.*, vol. 43, pp. 203–206, 2016.

[28] R. Wongsathan and S. Chankham, "Improvement on PM-10 forecast by using hybrid ARIMAX and neural networks model for the summer season in Chiang Mai," *Procedia Comput. Sci.*, vol. 86, pp. 277–280, 2016.

[29] R. Wongsathan, "Improvement of PM-10 forecast using ANFIS model with an integrated hotspot," *Sci. Technol. Asia*, vol. 23, no. 3, pp. 61–70, 2018.

[30] R. Wongsathan, "PM-10 forecasting models using MLPNN with a hotspot predictor over the Upper Northern Thailand," *Sci. Technol. Asia*, vol. 25, no. 4, pp. 93–105, 2020.

[31] C. Chairungrueang and R. Wongsathan, "Forecasting the PM-10 using a deep neural network," Songklanakarin J. Sci. Technol., vol. 43, no. 3, pp. 687–695, 2021.

[32] R. Rakholia, Q. Le, B. Ho, K. Vu and R. Simon, "The HelthyAir Dataset: Outdoor Air Quality in Ho Chi Minh City, Vietnam," *Mendeley Data*, vol. 1, 2022.

[33] J. Chen, D. Li, G. Liu, Y. Li, A. Zhang, S. Lu and M. Zhou, "Development of a coal dust concentration sensor based on the electrostatic induction method," *ACS Omega*, vol. 8, no. 13, pp. 13060–13067, 2023.

[34] V. Rachnarong, "Performance evaluation of low-cost particulate matter sensors for PM2.5 and PM10 measurement," Engineering and Technology Horizons, vol. 40, no. 3, pp. 1–18, Sep. 2023.

[35] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Stat.*, vol. 13, no. 3, pp. 253–263, 1995.

APPENDIX A

Table 10: *Chiang Mai $PM_{2.5}$ Data (1/1/2012-5/31/2024) (Top-Down, Left-Right Reading Order).*

19	72	49	35	38	23	159	42	100	49	18	54	15	60	37	13	94	6	41	56	40	32.3
15	90	45	41	36	21	64	56	114	54	7	48	20	65	43	17	71	26	46	55	50	42.9
16	58	43	42	41	23	46	58	144	48	10	61	19	73	44	26	50	21	39	49	40	32.7
15	55	33	49	56	31	38	67	104	43	12	74	20	88	47	28	87	17	13	38	40	35.1
18	58	30	53	80	21	30	48	94	39	9.6	83	23	93	63	19	116	19	7	50	49	43.4
17	68	21	54	72	33	22	19	103	39	8	87	19	86	54	25	108	16	15	66	48	54.2
17	31	25	41	67	32	18	6.9	75	44	13	95	21	64	45	20	103	16	28	69	28	50.2
21	31	23	30	60	20	23	6	53	50	24	73	13	54	41	10	105	16	39	75	29	42.7
23	16	28	25	52	19	26	6	59	37	30	59	22	57	38	12	114	21	43	76	25	47.5
19	23	30	26	48	22	42	13	67	39	38	83	29	53	33	19	102	18	59	58	25	51.1
18	24	29	25	50	36	41	27	81	26	27	74	31	37	50	16	104	19	64	30	43	42.1
20	27	34	21	58	29	52	33	111	34	22	69	25	49	81	16	94	17	42	27	51	37.9
26	29	46	27	45	42	91	32	109	38	28	64	21	53	86	20	74	24	44	34	52	52.7
24	36	57	27	65	53	84	31	53	47	15	72	26	58	78	28	66	30	48	51	60	94.4
16	46	65	29	61	50	84	25	45	41	15	71	31	44	56	25	67	28	62	42	14	122.2
17	43	49	28	62	33	115	31	30	39	17	60	23	20	56	11	76	30	72	44	19	100.4
19	44	52	29	78	34	90	34	29	38	18	74	25	29	71	11	73	17	52	49	16	56.1
20	39	65	31	79	51	95	42	42	48	21	108	27	45	56	16	69	10	65	59	20	62.9
17	36	67	24	77	23	69	45	50	67	16	107	44	43	55	27	60	14	60	78	14	81
20	38	69	21	73	11	48	47	49	48	20	72	44	45	56	38	55	19	42	116	13	62.7
14	38	68	20	96	10	41	32	52	55	18	45	46	42	68	39	36	15	16	49	24	42.5
13	36	81	25	104	10	56	60	53	54	16	52	47	56	69	38	42	17	21	16	21	94.1
13	48	49	19	93	15	33	66	66	54	21	34	38	55	54	27	59	12	18	15	25	152.7
14	48	34	25	107	27	20	86	60	61	22	43	42	56	59	21	52	16	30	19	32	101.9
14	40	51	18	137	34	30	93	58	42	39	37	47	59	55	11	64	19	35	47	31	95.9
14	42	51	33	109	33	42	85	65	36	40	28	29	63	42	10	64	15	38	64	30	120.6
16	49	65	34	106	48	55	85	61	34	35	20	22	47	66	6	73	18	39	55	32	115.5
16	56	42	26	188	58	45	66	49	33	24	14	18	52	57	9	93	20	42	54	32	47
21	42	45	20	159	62	58	80	40	36	23	27	24	54	48	24	97	18	39	39	26	29.8
21	36	57	17	47	60	60	69	34	45	30	43	24	55	40	26	91	14	51	58	22	40.5
27	22	76	23	66	44	52	72	30	45	22.2	76	21	52	41	34	111	18	27	57	26	40.3
28	19	76	25	62	32	56	69	22	38	15	69	24	44	41	40	99	23	30	54	25	57.7
25	18	66	23	74	27	54	59	18	47	25	76	26	40	54	45	76	24	21	68	18	59.1
27	22	58	23	111	34	30	56	19	34.9	34	86	38	39	50	30	76	19	19	85	23	82.2
22	28	57	19	116	37	26	70	13.9	40	36	76	74	32	53	28	45	15	17	83	17	77.9
21	23	50	22	97	41	19	61	9	84	33	61	72	25	57	32	12	22	24	93	15	71
22	14	48	52	84	45	11	62	12	73	29	33	42	18	49	29	14	19	29	90	12	82.6
25	8	36	39	89	42	8	59	17	64	27	25	38	17	113	36	24	16	26	81	26.8	87
28	10	58	45	115	46	10	40	19	98	32	34	46	29	174	45	20	12	10	72	19.2	98.2
34	16	32	42	119	50	10	26	19.3	96	34	50	41	36	160	34	12	16	9	76	18.3	93.1
29	23	30	47	127	60	11	16	20	57	23	57	45	40	66	24	16	17	8	81	20.2	106.7
29	22	17	40	92	73	13	17	22	51	26	80	52	33	68	38	18	18	10	75	21.9	131.6
27	31	28	28	42	63	20	20	21	32	18	73	51	34	77	55	22	22	7	90	19.4	116.5
23	29	49	29	52	54	17	31	14	29	19	82	56	27	116	45	24	25	9	91	22.9	113.7
29	28	56	34	69	47	16	52	17	41	19	71	63	10	129	47	31	30	18	100	28.7	163.9
45	28	73	33	57	45	13	76	13	34	26	45	64	15	69	52	17	32	15	53	24.2	143.3
41	28	100	40	39	52	16	57	10.4	49	24	37	48	20	70	45	5	32	16	43	21.3	116.3
88	27	108	35	51	54	46	55	11.3	49	25	34	46	22	93	53	14	20	21	41	21.3	136
97	25	102	33	54	59	56	74	13	17	21	22	58	23	78	57	22	20	10	38	22.8	51.7
52	19	94	36	55	70	43	87	13.5	27	30	17	62	21	88	56	32	18	9	35	23.7	28.1
48	21	74	35	56	65	50	57	13.3	38	26	16	66	18	87	58	41	12	9	47	27	47.5
51	21	62	35	43	62	60	63	13	45	32	9	47	12	94	58	51	13	9	47	29.9	33.5
53	16	112	38	31	61	38	74	15	61	50	10	42	8	123	40	53	21	5	35	21.3	62.6
75	20	188	33	38	52	41	85	23	69	51	15	41	12	117	43	35	27	13	52	19.9	49.9
82	25	141	31	35	50	34	80	25.2	74	54	15	36	28	107	47	25	26	18	68	20.7	45.9
122	21	128	41	42	74	24	70	20	72	65	22	25	17	86	43	17	27	10	73	15.2	53.7
117	30	132	50	42	76	19	80	12	52	57	33	23	16	87	50	11	28	15	72	15.2	88.2
81	24	123	46	47	56	14	85	12	50	51	38	31	30	84	38	8	35	20	94	20.1	91.1
79	18	110	53	50	56	10	55	14	66	47	41	43	55	92	29	7	41	20	109	28.4	84.4
83	16	125	46	39	52	8	59	14	48	39	34	35	36	123	37	10	48	25	101	29.7	70
89	12	120	45	42	47	9	63	18	15	42	27	52	27	143	28	14	49	23	94	35.7	66.1
100	11	101	41	48	52	11	59	19	13	54	25	80	51	125	20	14	49	21	149	36	66.2
86	15	101	61	46	73	9	86	26	14	65	21	122	58	83	32	19	51	17	140	31.8	61.6
74	38	88	63	34	67	10	94	25	15	64	22	119	47	90	47	19	56	32	173	26.6	63.9
85	47	82	43	33	56	14	58	33	11	49	24	126	35	81	59	20	48	24	213	38.4	85.3
105	40	89	46	37	64	21	86	27	22	34	23	146	40	67	59	13	38	22	137	27.2	65.4
120	44	102	43	36	56	19	95	29	24	37	19	144	80	79	40	19	41	20	86	27.2	72
120	41	112	34	39	77	15	138	21	39	48.4	16	122	45	87	9	12	21	21	98	15	75.4
147	49	77	28	38	96	22	98	24	55	57	19	88	34	65	12	17	26	16	143	14.7	79.5
110	46	90	27	34	156	20	81	21	58	59	22	56	30	35	20	11	32	14	210	17.5	71.2
34	46	95	32	31	168	15	63	19	60	61	26	47	38	42	25	12	32	28	236	22.7	84.4
34	49	82	32	33	131	12	58	15	57	43	16	51	51	43	31	12	34	30	160	44.7	56.7
31	53	88	34	39	80	25	75	20	57	39	18	103	43	49	34	17	37	27	143	36.4	88.5
29	55	93	31	31	56	25	97	24	64	39	12	210	51	56	40	23	39	28	109	32.3	65.9
38	43	84	29	27	92	27	62	28	53	37	12	139	57	57	42	24	30	29	135	27.6	89.8
40	38	72	31	30	98	26	63	28	13	49	14	101	43	58	41	26	38	25	154	41.7	61.7
66	50	48	38	24	103	27	85	40	17	43	7	75	31	64	50	24	42	24	145	38.3	31.6
120	41	50	46	32	202	38	81	39	32	75	12	60	34	58	38	22	48	25	92	36.5	31.1
127	52	41	46	27	256	45	70	33	33	82	10	63	34	46	30	19	39	36	130	29.7	24.5
104	53	28	38	28	174	45	78	26	31	74	11	110	46	39	37	22	32	37	130	37	23.3
85	55	61	41	20	129	37	99	27	33	58	21.9	205	45	30	32	18	19	45	83	34.2	18.2
74	46	56	34	25	108	30	91	25	25	51.8	27	92	50	15	42	9	12	44	98	35.2	18.9
73	46	61	34	27	92	40	67	36	18	66	17	115	42	9	53	5	9	36	92	31.5	16.2
67	44	70	31	27	173	44	70	40	22	86	16	119	56	8	38	12	16	36	82	36.8	19
60	51	50	23	25	266	48	73	39	24	68	16	108	70	9	46	13	26	38	78	34.2	28.7
50	44	32	32	25	257	44	79	41	31	100	16	116	54	14	63	17	34	35	75	28.2	17.1
46	32	24	32	15	160	30	70	40	32	72	18	114	38	16	76	16	45	42	39	24.7	13.6
51	33	30	56	23	135	27	69	47	27	38	22	99	37	11	72	14	51	47	41	23.2	16.7
64	38	27	38	22	155	35	69	51	22	59	17	68	40	9	66	6	42	51	40	25.1	16

Rati Wongsathan received his B.Eng. (Hons.) in Electrical Engineering, M.Eng. in Electrical Engineering, M.Sc. in Applied Mathematics from Chiang Mai University (CMU), Chiang Mai, Thailand, in 1997, 2000, and 2005, respectively, and PhD in Electrical Engineering from King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in 2020. His research interests include renewable energy, mathematical modeling, telecommunications, data storage, optimization, AI, and engineering education. He has published over 40 journal articles indexed by SCOPUS and ISI with impact factors, 30 conference papers, and 2 books. He currently served as an assistant professor at the Faculty of Engineering and Technology of North-Chiang Mai University (NCU), Chiang Mai, Thailand.

Win Mya Thway received her B.C.Sc. (Hons.) in Computer Science from the University of Computer Studies, Yangon, Myanmar, in 2009, and her M.Sc. in Information Technology from Shinawatra International University, Bangkok, Thailand, in 2014. She also obtained a Diploma in Development Studies from the Yangon Institute of Economics, Myanmar, in 2018. Her research interests include artificial intelligence (AI), software engineering, information systems, and development studies. She is currently serving as a lecturer in the Department of Software Engineering, Faculty of Engineering and Technology, North-Chiang Mai University (NCU), Chiang Mai, Thailand.

Wutthichai Puangmanee was born in Lampang, Thailand. He received his B.Ed. in Computer Education from Lampang Rajabhat University, Thailand, in 2004, and his M.Sc. in Internet and Information Technology from Naresuan University, Phitsanulok, Thailand, in 2007. He is currently pursuing a Ph.D. in Computer Engineering at Kasetsart University, Bangkok, Thailand. His research interests include embedded systems, automation systems, and machine learning.