P-ISSN 2586-9000
E-ISSN 2586-9027

Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia

Vol.25 No.4 October - December 2020
Page: [93-105]

Original research article

PM-10 Forecasting Models using MLPNN with a Hotspot Predictor over the Upper Northern Thailand

Rati Wongsathan*

Department of Electrical and Computer Engineering, Faculty of Engineering and Technology Research Institute, North-Chiang Mai University, Chiang Mai 50230, Thailand

Received 15 November 2019; Received in revised form 10 April 2020
Accepted 26 June 2020; Available online 24 December 2020

ABSTRACT

The local open burning and trans-boundary haze are the main sources causing serious air pollution-related to particulate matter 10 micrometers or less in diameter (PM-10), which severely affects the health of the people living in the upper Northern Thailand region. In the absence of complex emission and formation mechanisms of PM-10 information, the objective of this paper is to model and to forecast the daily PM-10 concentration using the multilayer perceptron neural networks (MLPNN). In this regard, the highly correlated hotspot fires to PM-10 were utilized as the predictor for improving model performance. For achieving a more realistic model, the related meteorological parameters and some time lags of historical PM-10 are selected into the MLPNN models by the forward selection method. Furthermore, the MLPNNs are optimized through the experiment by minimizing the number of hidden nodes to prevent the over-fitting. Data on pollution, climate, and hotspot collected during 2012-2019 and obtained from several governments and private sectors are used to train and validate the forecasting models. The optimal MLPNN with an integrated hotspot can capture the characteristics of complex PM-10 as well as the ANFIS based fuzzy rules describing the dynamic changes of PM. As a result of prediction, the proposed model provides up to 86% forecasting accuracy for one day ahead, and 11%–22% more than that for the MLPNN with a double hidden layer and the MLPNN without using the *Hotspot* predictor.

Keywords: ANFIS; Hotspot; Multilayer perceptron neural network; PM-10

1. Introduction

Nowadays, air pollution related to particulate matter (PM) is an emerging issue of disaster across the globe, including Southeast Asia. Recently, severe PM pollution has received much attention in Thailand, especially in the upper Northern Thailand (UNT) region. For a decade, the

***Corresponding author:** rati@northcm.ac.th

PMs with a diameter of less than 2.5 and 10 μm, called PM-2.5 and PM-10, respectively, have caused severe adverse effects in this area. Normally, the PM-10 can be regarded as a health indicator during the haze episode of the UNT since the existence of PM-2.5 monitoring stations is limited. Various biomass open-space burnings from both local and nearby outside areas are indicated as the major primary sources of PM [1]. PM-10 concentrations vary with the locations, emission, and transportation, and strongly depend on various exogenous factors, such as climate change and meteorological terms. They are also well correlated with the concentrations of the other toxic gases [2]. Typically, the measured PM-10 is announced daily in the morning. But this may not be thoroughly accessible and is too late to communicate awareness of the health care setting to people. As of now, there are still no environmental agencies implementing a PM-10 forecast model. Accordingly, to overcome the stated problems, this work is to propose an efficient and accurate PM-10 forecast model as one of the air quality management tools.

In the literature, the various versions of weather research and forecasting (WRF) with different coupling models have been simulated the PM-10 concentrations, such as the WRF with the California Mesoscale Puff or WRF-CALPUFF model [3], the WRF coupling with a numerical module of chemical interactions, called WRF-Chem [4], and WRF with the physical laws of motion and conservation of energy [5]. Although they provide a high resolution, slow computation and requiring extensive computer resources are disadvantages.

Alternatively, most of the researchers tend to use the statistical approach-based PM-10 models without knowing the mechanism background of PM-10 changes. Instead, they are based on the statistical relation of PM-10 and other meteorological parameters and pollutants. They often provide a higher accuracy as compared to deterministic models. Two types of this approach are parametric and non-parametric statistical methods. Most of the previously existing PM-10 forecast models have been carried out through the different regression techniques among the parametric statistical models, such as a linear regression [6, 7], a logistic regression [8], and an auto-regressive integrated moving average (ARIMA) [9]. By comparison, ARIMA models outperform the regression models in terms of accuracy. Besides, an ARIMA with exogenous variables or AIMAX model [10] can improve forecasting performance. However, these linear models may fail in making a complex nonlinear forecast since they can only capture the linear relationship.

To overcome the limitations of those classical methods, a neural network (NN)—a brain-like processor with learning and adaptive capabilities among the nonparametric stochastic model has been applied to forecast PM-10. For example, the multilayer perceptron NN (MLPNN) [11] and radial basis function NNs (RBFNN) [12] could provide reasonably accurate results. However, the massive training data required and over-fitting due to the large structure of NNs are the main drawback. Several studies have proposed the hybrid models to improve the accuracy obtained by using either linear or nonlinear models separately, such as the hybrid ARIMA-NN [9], the hybrid ARIMAX-NN [10], and the hybrid ARIMA-support vector regression [11]. More recently, another hybrid model based on the adaptive neuro-fuzzy inference system (ANFIS) [10] was implemented. The overall performance is improved, while increasing the computational complexity.

Nowadays, the advance of satellite remote sensing technology plays an important role in air quality prediction at ground level. For example, the MODerate-resolution Imaging Spectro-radiometer (MODIS) sensor provides an aerosol optical depth data [6], and the active fire hotspot

[13], which are used to assess the PM-10 emissions emitted from forest fires and to predict hourly average PM-10, respectively. In [14], the ANFIS-based PM-10 forecast model is improved by including the hotspot parameter detected by the MODIS satellite.

PM-10 is influenced by the meteorology parameters [15]. Furthermore, the open burning has been demonstrated as having significant impact on the PM-10 variations in the UNT area [16]. So, the number of hotspots representing open fires and the other weather parameters are also included as the predictors in the forecasting model. Up to now, there are few NN-based PM-10 forecasting models for the UNT. In this work, the MLPNN with a single hidden layer integrated the *Hotspot* predictor is first proposed to formulate the PM-10 forecasts. As compared with the previous work [14], the proposed forecast model gives accuracy slightly better than that of the ANFIS by 0.6%. However, the complexity of the ANFIS, in terms of the multiplication counts in computing and the number of system parameters, is exponential growth, which is the disadvantage, whereas that of the MLPNN is linear growth. Moreover, long processing times, depending on the fuzzy rules in ANFIS, are another drawback. Besides, with the use of a large amount of training and validating data, the proposed MLPNN can be improved over the ANFIS.

However, a more recent study has reported that the model of increasing complexity results in superior predictions [17]. Therefore, an attempt in improving the performance using the double hidden layer of MLPNN is further investigated. In order to achieve the optimal model, the forward selection (FS) method [18] is used to select the most significant variables to be the model inputs subject to the criterion of minimizing the cost function. Besides, the structure of MLPNN in terms of the number of hidden nodes is subsequently minimized through repeated experiments.

Three types of the simple MLPNN models, (1) with and (2) without the *Hotspot* predictors and (3) the optimal MLPNN, are modeled. In addition, the double hidden layer MLPNN, a more complex structure, is formulated accounting for the complicated PM-10 behavior. Once the different MLPNNs are constructed using the training dataset (2012-2016), they are optimized through the validating dataset (2017-2018). Their performances are evaluated using the testing dataset (2019) under the tradeoff between mean absolute error (MAE) and root mean squared error (RMSE) criteria measuring the accuracy which are further compared to those of the existing ANFIS model [14].

2. Methodology

2.1 Study area

The UNT region with a population of about 6 million (2016) consists of 9 provinces (Fig. 1): Chiang Mai (CM), Chiang Rai (CR), Mae Hong Son (MHS), Lampang (LPG), Lamphun (LP), Phrae (PR), Nan, Phayao (PYO), and Tak. It is situated in between 19° 00′ North latitude and 99° 00′ East longitude covering over a 93,000 km^2 area or 18% of Thailand, and is bound by Myanmar to the north and west, Lao to the north and east, and the lower northern provinces to the south. This area is the sub-domain of mainland Southeast Asia (MSA) where air pollution events occur due to annually recurrent fire activity. It is characterized by multiple mountain ranges and basin-like geography and influenced from the high pressure of the northeast monsoon resulting in restricted pollution dispersion. It also has a tropical savanna climate with low rainfall during the dry season that could lead to the severe forest fires. From a topography map of the UNT with PM-10 monitoring sites and the occurrence of the hotspots (dotted-blue) (Fig. 1), it is seen that people in this area suffer from the PM-10 pollution around one month a year.

The descriptions of PM-10 situations of all 9 provinces are briefly detailed as follows:

1) In CM, the largest city in UNT, the primary sources of the smoke haze mainly go beyond the open burning of forest and agricultural from the local area and neighboring Myanmar and Lao [19].

2) In CR, the northernmost province of Thailand which borders on Myanmar and Laos, the PM is mainly caused by short-range movement from open burning in this area itself. Incorporating the impact of meteorological and topographical factors, the vertical dispersion of smoke is inhibited resulting in the PM-10 accumulation [20].

3) MHS, the north-west most province covered mostly with forest which borders Myanmar, is most severely affected by PM-10 pollutions. Forest fires are seen as the main cause according to a report by the Geo-Informatics and Space Technology Development Agency (GISTDA).

4) In LPG, the province is affected by the air pollution from another source, especially a coal-fired electrical power plant in Mae Moh district and the emission of rice straw burning of 5000 tons/year.

5) LP, a small city, has the lowest PM level than the rest corresponding to a small number of hotspots occurrence. However, there are several very unhealthy days.

6) In PYO, from a hybrid single-particle Lagrangian integrated trajectory model (HYSPLIT) [21] the significant PM source is open burning from the local area, other neighbor provinces, and Myanmar.

7) In PR, the south-east most province, from the HYSPLIT model in conjunction with the potential source contribution function techniques [22], the open burning is indicated as the major PM-10 source.

8) Nan, second-largest corn-producing province has had massive deforestation and consequently high corn-waste burning. Despite continuously decreasing the detected hotspots in this area, so far the PM-10 level still exceeds the safety standard.

9) Tak, the southernmost province which borders Myanmar, has an increasing trend of the average PM-10 level as well as the number of hotspots.

2.2 Data collection and analysis

In this work, the emphasis focused on PM-10 forecasts during the high open burning season where the PM-10 variation is high and depends on various factors. It is found that meteorology has a strong impact on PM-10 accumulation in the UNT region [3]. For example (Fig. 2), in CR province, the atmospheric pressure, wind velocity, and humidity are found to be significant factors compared to the rest [21]. In MHS province, the PM-10 concentrations are significantly positively correlated to relative humidity (p-value < 0.001) [2] (Fig. 2). Moreover, the trajectory model illustrated that north or northwesterly winds from neighboring countries bring more PM, resulting in the critical haze episode period of each year. In LPG, air mass analysis by the backward-trajectory using the HYSPLIT model has pointed out that wind is a significant factor (Fig.2) that can carry the PM-10 from other areas to this province. Therefore, the previous day of the meteorological data is introduced as a predictor of the MLPNN forecast model.

However, it is evident that the open burning is the major source of PM-10 in the UNT. The correlation of the number of hotspots and the PM-10 (Fig. 3) supports the statement above. In this regard, the basic assumption of this work is that using the number of hotspots as the predictor can improve the forecast performance. Therefore, the number of hotspots of previous day is integrated into another MLPNN forecasting model.

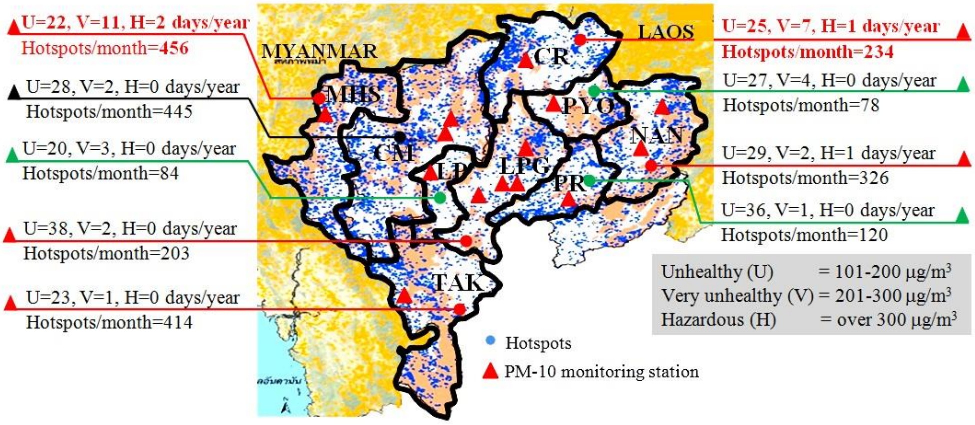


Fig. 1. Study area and a number of hotspots (blue dotted) detected daily from MODIS aboard the Terra and Aqua satellites during 2012-2019.

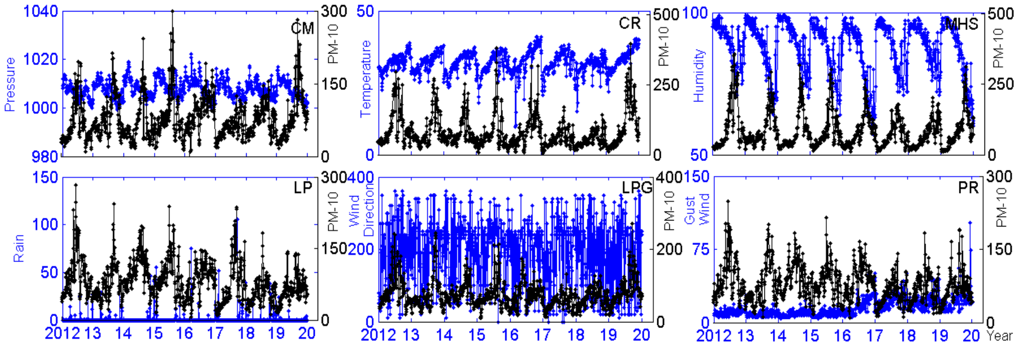


Fig. 2. PM-10 against weather parameters.

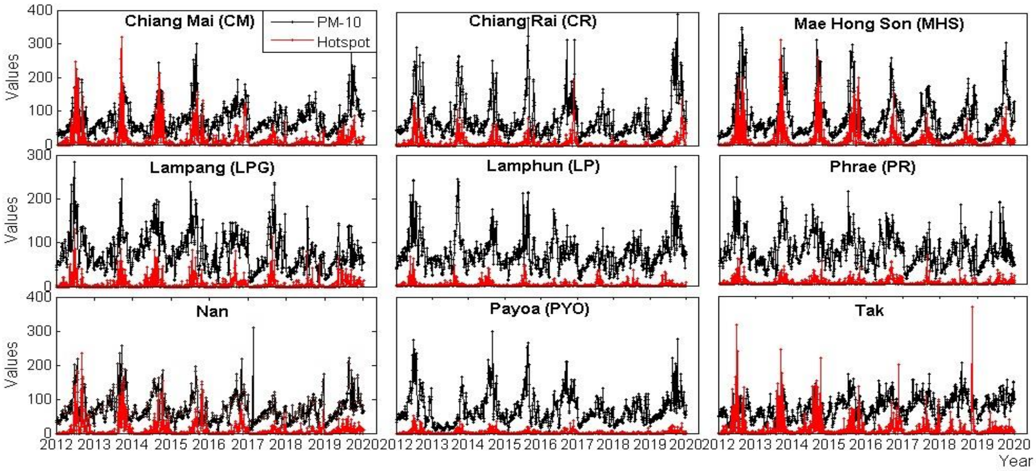


Fig. 3. The plots of the number of hotspots and PMs-10.

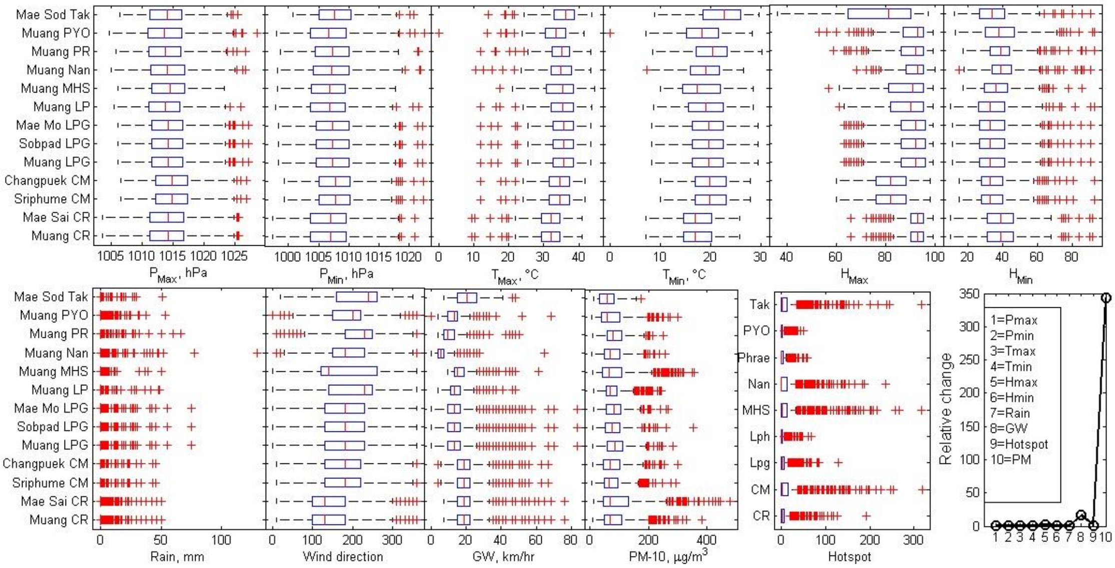


Fig. 4. Descriptive statistics through the corresponding box plots.

The meteorology variables composed of maximum and minimum pressure (P_{max} and P_{min}) in hector Pascal, maximum and minimum temperature (T_{max} and T_{min}) in degree celsius, maximum and minimum relative humidity (H_{max} and H_{min}), rain (R) in millimeters, wind direction (WD), and gust wind (GW) in km/hr are used as the predictor sets in formulating the forecast model. The weather changes data are collected from Chiang Mai meteorological department, whereas the hotspot and PM-10 data are obtained from the MODIS team during 2012-2019. The descriptive statistics of these variables are depicted by box plots in Fig. 4. After pre-processing the data for missing and outlier values, it is seen that the variables have a wide range of the relative change (ratio of variance compared to the range) from 0.43 of the T to 350 of PM-10. The MLPNN with a single hidden layer may not identify the dynamic change of the input variables. Therefore, the MLPNN with a double hidden layer is another choice to overcome the problem. However, its performances may deteriorate, as not expected, due to the overfitting problem from such a large structure.

2.3 Description of MLPNNs based PM-10 forecasting model

In this work, the MLPNN based PM-10 forecast model typically is composed of three layers: input layer, hidden layer, and output layer, as shown in Fig. 5 (a) and (b). The output of MLPNN(N_1, N_2, 1), $PM_{MLPNN}(t+1)$, referred to the forecast PM-10 one day ahead, is a weighted summation of each hidden node’s output which can be expressed in matrix-vector form as

$$\begin{aligned}&\mathrm{PM}_{MLPNN}(t+1)=\\&purelin\left(\mathbf{W}^{(2)}\times tanh\left(\mathbf{W}^{(1)}\times\mathbf{X}+\mathbf{b}^{(1)}\right)+b^{(2)}\right),\end{aligned}\qquad(1)$$

where $\mathbf{W}^{(1)}$ and $\mathbf{b}^{(1)}$ is ($N_2\times N_1$)-weight matrix and N_2-bias column vector between input and hidden layer, respectively, $\mathbf{W}^{(2)}$ and $b^{(2)}$ is N_2-weight column vector and the bias value between hidden and output layer, respectively, where N_1 is the number of input variable nodes of vector input $\mathbf{X}$ including periodic terms of $\sin(2\pi d/120)$ and $\cos(2\pi d/120)$, where $d\in[1,120]$, N_2 is the number of hidden nodes and *tanh* and *purelin* are hyperbolic tangent and linear functions, respectively.

During the training, the parameters of NN are adjusted by minimizing the following cost function,

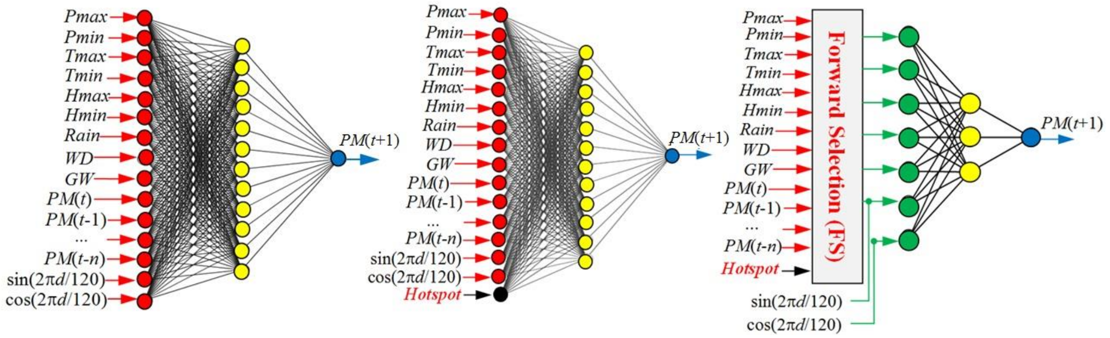
$$J(\mathbf{W}^{(1)},\mathbf{W}^{(2)},\mathbf{b}^{(1)},b^{(2)}) = \sum_{k=1}^{n}\delta^{n-k}\left(\frac{1}{N}\sum_{j=1}^{N}\left(PM_j - PM_{MLPNN}(t+1)\right)^2\right), \tag{2}$$

where PM_j is the measured PM-10, N is the number of training samples, n is the maximum number of iterations and $\delta \in (0, 1]$ is a forgetting factor that helps to prevent the local solution trapping from back propagation algorithm (BPA). In BPA, the output errors are passed back through hidden layers to train or update the connected weights and biases Eq. (4)- Eq. (10) of the network by the gradient search in minimizing the error cost function Eq. (2).

In Fig. 5 (c), the MLPNN forecasting models are optimized using the forward selection (FS) method [5] for selecting the significant input parameters and the cross-validation in choosing the appropriate number of hidden nodes (N_2). Employing the FS method, the variable with most significance observed from the correlated coefficient is selected in the beginning along with the *Hotspot* variable through the pilot models, and we continue adding one another tentative variable to the model as long as its P-value is below the pre-set level (0.05).

To further develop, the double hidden layer MLPNN (Fig. 6), shorted by DL-MLPNN(N_1, N_2, N_3, 1), where N_3 is the number of nodes in the second hidden layer, is formulated as the PM-10 forecast model. The output of the model is expressed as,

$$\mathrm{PM}_{DL\text{-}MLPNN}(t+1) = purelin\left(\mathbf{W}^{(3)} \times tanh\left(\mathbf{W}^{(2)} \times tanh\left(\mathbf{W}^{(1)} \times \mathbf{X} + \mathbf{b}^{(1)}\right) + \mathbf{b}^{(2)}\right) + b^{(3)}\right), \tag{3}$$



(a) Without Hotspot (b) With Hotspot (c) Optimized

Fig. 5. The single hidden layer MLPNN based PM-10 forecast models.

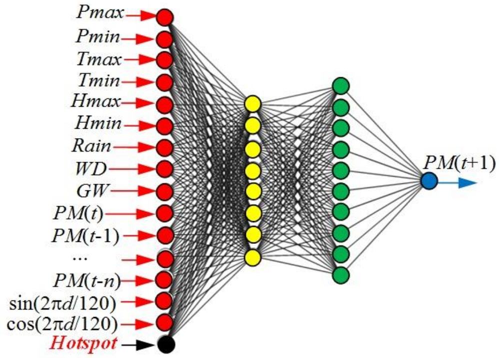


Fig. 6. The double hidden layer MLPNN-based PM-10 forecast model.

where $\mathbf{W}^{(2)}$ and $\mathbf{b}^{(2)}$ are an $(N_3 \times N_2)$-weight matrix and N_3-bias column vector between the first and second hidden layer, respectively, $\mathbf{W}^{(3)}$ and $b^{(3)}$ are an N_3-weight vector and the bias between the second hidden and output layer, respectively.

Selecting the number of hidden nodes and layers is very important in deciding the overall NNs structure. Using too few hidden nodes results in under-fitting. On the other hand, using too many hidden nodes results in over-fitting. For simplicity, the single hidden layer, normally sufficient and a large number of hidden nodes, are applied for the pilot model. After that, they are continually reduced by one node as long as the error is significantly decreased. To optimize the MLPNN(N_1, N_2, 1), N_1 is obtained from the number of inputs, in FS method, and N_2 is selected through the k-fold cross-validation. The training MLPNN with BPA and using FS method is expressed as the following algorithm:

Algorithm: Training of MLPNN(N_1, M, 1) with FS

```
Input:     Training data set {$\mathbf{X}^{(15\times1)}$, $PM(t+1)\}_i$, $\forall i \in \{1,\ldots,N\}$; $Threshold\_FS$;
           Significant variable={$PM(t)$, $Hotspot$}; Learning ratio ($\alpha$), Forgetting factor ($\delta$);
           P-value=0.05;Initial hidden node ($M$), Maximum iteration ($Max\_iteration$)

Output:  The weights ($\mathbf{W}$) and biases ($\mathbf{b}$); The set of significant predictors and $N_1$

Initialization: FS(1)={$PM(t)$, $Hotspot$}, $N_1$ = 2; $error$ =1; $epoch$ =1;RMSE(0)=100

1: For $i$ = 1:13                          %The number of tentative variables=13;
2:    if $i > 1$
3:          if  RMSE($i$)–RMSE($i$–1) < -$Threshold\_FS$
4:              $\mathbf{FS}(i) \leftarrow \mathbf{FS}(i-1) \cup \mathbf{X}(i)$;
5:              $N_1$=$N_1$+1;
6:           end if (3)
7:     end if (2)
8:     % Set the random weight matrix and vector, $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$, respectively,
9:     $\mathbf{W}^{(1)}$ =2×rand($M$, $N_1$) – 1;  $\mathbf{W}^{(2)}$ =2×rand($M$, 1) – 1 ;
10:    % Set the random bias vector and value, $\mathbf{b}^{(1)}$ and $b^{(2)}$, respectively,
11:    $\mathbf{b}^{(1)}$ =2×rand(1, $M$) – 1; $b^{(2)}$ =2×rand(1, 1) – 1;
12:   While $epoch$ < $Max\_iteration$ or $error$ > 0.0001 do
13:      For $i$ = 1:$N$
14:          $input\_hidden\_layer$=$\mathbf{W}^{(1)}\times\mathbf{X}+\mathbf{b}^{(1)}$;
15:          $output\_hidden\_layer$=$tanh$($input\_hidden\_layer$)
18:        $input\_output\_node$= $\mathbf{W}^{(2)}\times output\_hidden\_layer+b^{(2)}$;
19:        $final\_output(i)$= $purelin$($input\_output\_node$);
20:          $error\_output(i)$=$PM_i(t+1)$-$final\_output(i)$;
21:          % Back-propagation
22:          $delta$(1)=$error\_output(i)$                                          (4)
23:          $error\_hidden\_layer$=$\mathbf{W}^{(2)\prime}*delta(1)$;                     (5)
24:       $delta(2)$=$tanh'$($input\_hidden\_layer$) ×$error\_hidden\_layer$ ;            (6)
25:          $\mathbf{W}^{(1)} \leftarrow \mathbf{W}^{(2)} + \alpha\times delta(1)\times output\_hidden\_layer'$;     (7)
26:          $\mathbf{b}^{(1)} \leftarrow \mathbf{b}^{(2)} + \alpha\times delta(1)\times output\_hidden\_layer'$;     (8)
27:          $\mathbf{W}^{(2)} \leftarrow \mathbf{W}^{(2)} + \alpha\times delta(2)\times\mathbf{FS}'(i)$;          (9)
28:        $\mathbf{b}^{(2)} \leftarrow \mathbf{b}^{(2)} + \alpha\times delta(2)\times\mathbf{FS}'(i)$;            (10)
29:    end for (13)
30:   end while (12)
31:   RMSE($i$)=rmse($error\_output$);
32:   Check P-value of $\mathbf{FS}(i)$
33:end for (1)
34:Generate the MLPNN-based PM-10 forecasting model with significant predictors
```

3. Results and Discussion

In formulating the MLPNN-based forecasting PM-10 model through the experimental design, the collected data during January-April of 960 samples are divided into 3 parts: training (2012-2016), validating (2017-2018), and testing (2019). The input set comprises the historical PM-10 of 5-time lags, $PM(t)$, $PM(t\text{-}1)$,…, and $PM(t\text{-}4)$, the meteorological variables (P_{max}, P_{min}, T_{max}, T_{min}, H_{max}, H_{min}, $Rain$, WD, GW, and RH), and the hotspot.

The magnitude of the correlation coefficients between these predictors and the next day forecasting PM-10, $PM(t+1)$, is preliminarily evaluated. The results are shown in Fig. 7. It is seen that $PM(t+1)$ is well correlated with $PM(t)$ and uncorrelated with T_{min}, *Rain*, *WD*, and *GW*, which are not a consideration in the MLPNN forecasting models. The rest (11 variables) excluding *Hotspot* are used as the predictors of the first type model, named MLPNN(10, N_2, 1), whereas all variables including *Hotspot* are used as the predictors of the second type model, named MLPNN(11, N_2, 1) with *Hotspot*, where the parameter N_2 is minimized through 5-fold cross-validation. Besides, all variables are selected through the FS method to form the third type model, namely optimal MLPNN(N_1, N_2, 1). The *Hotspot*, among the others, is the exogenous variable that also correlates well with $PM(t+1)$. Therefore, $PM(t)$ and *Hotspot* are selected in the initial set in the FS through the pilot models of MLPNN(N_1, 10, 1).

The experimental results of the input selection using the FS method are shown in Fig. 7. An optimal predictor set containing the fewest significant input variables describing the PM-10 behavior is shown in Table 1. Whereas, the selected number of hidden nodes of all MLPNN models is shown in Table 2. It is seen that the MLPNN structure in terms of N_1, N_2, and N_3 for MHS province is larger than the rest, corresponding to the high variance of PM-10 levels in this area.

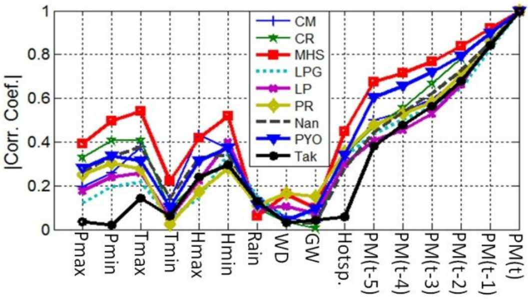


Fig. 7. The magnitude of correlation coefficients between $PM(t+1)$ with 5-time lags of historical PM-10 and previous day of exogenous variables.

Table 1. The optimal set of significant input variables from the FS method used in the optimal MLPNN-based forecast model.

Area	Selected input variables
CM	$PM(t)$, *Hotspot*, $PM(t–1)$, P_{min}
CR	$PM(t)$, *Hotspot*, $PM(t–1)$, H_{max}, P_{max}
MHS	$PM(t)$, *Hotspot*, $PM(t–1)$, H_{max}, P_{max}, T_{max}
LPG	$PM(t)$, *Hotspot*, P_{min}
LP	$PM(t)$, *Hotspot*, T_{max}
PR	$PM(t)$, *Hotspot*, T_{max}
Nan	$PM(t)$, *Hotspot*, $PM(t–1)$
PYO	$PM(t)$, *Hotspot*, $PM(t–1)$, T_{max}, H_{max}
Tak	$PM(t)$, *Hotspot*

Table 2. The number of hidden nodes selected from the experiments for MLPNNs.

Area	The number of hidden nodes			
	MLPNN without *Hotspot* (N_2)	MLPNN with *Hotspot* (N_2)	The optimized MLPNN (N_2)	DL-MLPNN (N_2, N_3)
CM	10	8	5	7, 5
CR	11	10	5	7, 7
MHS	12	12	7	10, 7
LPG	9	7	3	7, 7
LP	8	6	4	5, 4
PR	7	5	4	5, 4
Nan	9	8	5	6, 6
PYO	8	8	6	7, 5
Tak	5	4	3	5, 3

In the test, the error performances in terms of MAE and RMSE between MLPNN without *Hotspot*, MLPNN with *Hotspot*, the optimal MLPNN with *Hotspot*, DL-MLPNN and ANFIS [14] are compared in Table 3. It is seen that the optimal MLPNN with *Hotspot* performs the best prediction, whereas the MLPNN without *Hotspot* performs the worst prediction. The forecasting performance of the optimal MLPNN is slightly better than that of the ANFIS model [14] and the MLPNN with *Hotspot* variable about 0.6% and 2.4%, respectively, but significantly better than that of the DL-MLPNN and the MLPNN without *Hotspot* about 11% and 22%, respectively. The results verify the assumption that the *Hotspot* variable is the strong predictor of the model that helps to improve the forecasting performance significantly. Moreover, the results support the performance of the ANFIS model based on fuzzy rules describing the dynamic changes of the input variables, including PM-10 [14].

However, the performance of the DL-MLPNN seems to be less satisfactory than expected. It is due to the over-fitting from using a large number of hidden nodes that leads a large error for some observations in the testing stage (Fig. 8) corresponding to the highest RMSE value, whereas it is not observed in the training and validating stages. In this regard, it is possible to overcome the problem by using the dropout learning technique [23] to eliminate the co-dependent hidden nodes.

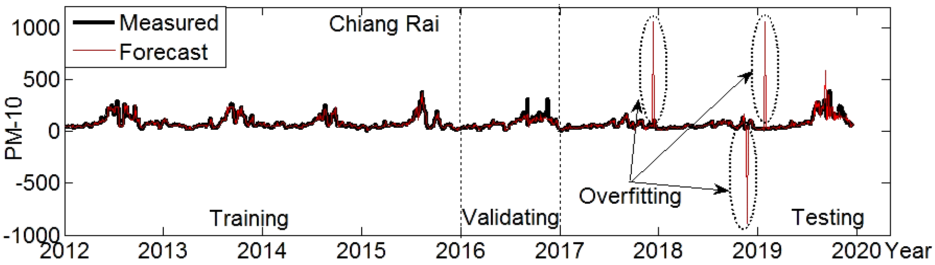


Fig. 8. The over-fitting due to the complex DL-MLPNN model.

Table 3. Comparison of the performances between PM-10 forecasting models based on the proposed MLPNNs, MLPNN with a double hidden layer, and the ANFIS [14].

Area	(1) MLPNN model without hotspot variable						(2) MLPNN model with hotspot variable					
	Train		Validation		Test		Train		Validation		Test	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
CR	19.01	12.45	39.74	23.40	31.43	17.05	20.57	12.84	37.50	20.54	23.80	15.20
CM	16.43	11.29	29.01	19.09	23.31	15.01	17.19	11.47	21.83	14.43	21.48	13.51
LPG	17.94	12.38	19.07	15.14	23.30	16.72	19.92	13.98	16.31	12.87	22.08	15.64
LP	17.24	12.35	15.12	11.11	22.46	14.83	18.01	12.72	14.25	10.43	21.89	14.52
MHS	17.98	11.99	26.25	16.90	37.80	19.10	19.26	13.01	24.55	16.10	27.97	18.38
Nan	17.30	11.76	21.65	16.27	20.92	14.93	18.28	12.38	21.47	14.64	20.53	14.95
Phr	15.86	11.84	17.9	13.77	23.94	17.59	18.08	12.89	12.23	9.79	19.29	14.19
PYO	14.84	10.33	39.63	23.23	28.12	16.55	18.18	12.16	20.82	15.16	21.22	13.95
Tak	12.44	9.30	18.59	14.202	24.90	18.65	13.93	10.35	16.63	12.44	22.38	16.90
Average	**16.5**	**11.5**	**25.2**	**17.0**	**26.2**	**16.7**	**18.1**	**12.0**	**18.3**	**13.2**	**21.1**	**15.0**
Area	(3) Optimized MLPNN				(4) Double hidden layer MLPNN				(5) ANFIS [14]			
	Validation		Test		Validation		Test		Validation		Test	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
CR	21.83	13.52	27.43	16.15	35.05	21.78	32.61	18.15	22.82	14.11	26.55	16.31
CM	17.27	11.49	20.79	13.88	25.74	15.57	22.14	14.27	17.57	11.52	21.49	13.82
LPG	18.48	13.20	22.11	15.13	15.48	12.31	21.99	15.36	18.87	13.22	22.38	15.23
LP	18.63	12.10	20.19	13.84	16.42	11.84	24.61	16.59	18.55	12.29	20.53	13.78
MHS	22.34	14.08	24.77	15.62	24.06	15.29	25.62	17.32	22.70	13.81	24.59	15.43
Nan	16.92	11.77	20.56	13.56	20.01	14.14	21.50	15.30	17.09	12.05	20.50	13.47
PHR	16.85	12.04	18.85	13.14	15.00	11.02	20.30	14.60	17.10	11.99	19.00	13.38
PYO	19.23	12.26	21.82	13.89	24.83	15.72	21.95	14.32	19.48	12.45	21.88	13.95
Tak	16.48	12.63	15.40	11.11	16.29	12.36	17.55	13.35	16.59	12.66	15.40	11.14
Average	**18.6**	**12.5**	**21.3**	**14.0**	**21.4**	**14.4**	**23.1**	**15.4**	**18.9**	**12.6**	**21.3**	**14.0**

The PM-10 forecasting results from the optimal MLPNN model of the validating (2017-2018) and testing (2019) data are shown in Fig. 9 and Fig. 10, respectively. It is seen that the proposed MLPNN models are capable of capturing the complex nonlinear characteristics of the PM-10 well even at the peak, which is difficult to achieve. Also,it can forecast the unseen data more correctly.

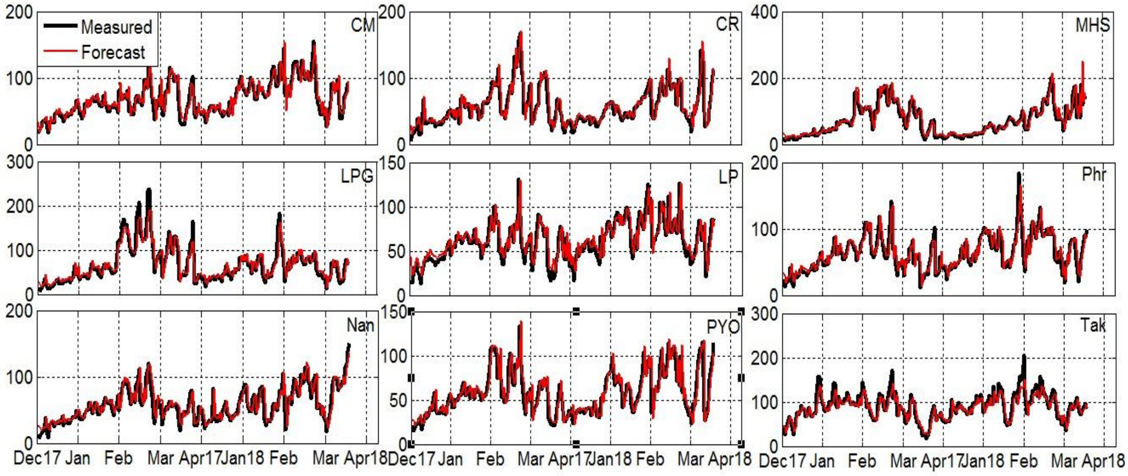


Fig. 9. The forecast results of the validating data (2017-2018) using the optimal MLPNN.

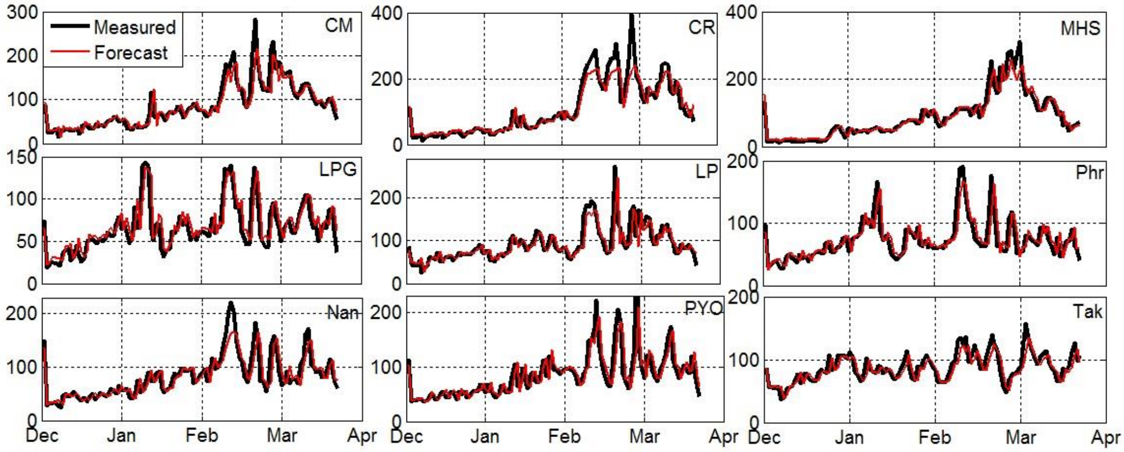


Fig. 10. The forecast results of the testing data (2019) using the optimal MLPNN.

4. Conclusions

The number of hotspots corresponding to the open burning as the main PM source in the UNT region is the key parameter in formulating the PM-10 forecast models. It is incorporated as the predictor among other parameters selected using the FS method. After optimization of the MLPNN by minimizing the number of hidden nodes, the optimal MLPNN can capture the nonlinear characteristics of the PM-10 very well when compared to the rest. The performances of the proposed model in terms of accuracy are slightly better than that of the ANFIS model but are much better than that of the MLPNN with a double hidden layer and MLPNN without using the *Hotspot* parameter.

PM-2.5, a subset of PM-10, is more dangerous in adversely affecting humans than PM-10, and it is now adopted in estimating the air quality index (AQI) together with PM-10. Therefore, in the future, the MLPNN might be used to forecast PM-2.5 from the related PM-10 for the area where the PM-2.5 monitoring station is lacked. Besides, due to the great advantage of deep learning technology in the forecast field, the deep neural network (DNN) is our choice for applying to PM forecasts.

References

[1] Pasukphun N. Environmental health burden of open burning in northern Thailand: A review. PSRU J of Science and Technology 2018;3(3):11-28.

[2] Mao M, Zhang X, Yin Y. Particulate matter and gaseous pollutions in three metropolises along the Chinese Yangtze river: Situation and implications. Int J Environ Res Public Heath 2018;15:1-29.

[3] Amnuaylojaroen T, Kreasuwun J. Investigation of fine and coarse particulate matter from burning areas in Chiang Mai, Thailand using the WRF/CALPUFF investigation of fine and coarse particulate matter. Chiang Mai J of Science 2011;39(2):311-26.

[4] Sooktawee S, Patpai A, Boonyapitak S, Kongsong R, Piemyai N, Humphries U. Influence of PM10 from the outside area affecting on the northern part of Thailand. The 3rd Environment and Natural Resources Int Conf (ENRIC 2018), Chonburi, Thailand, 2018:30-41.

[5] Macatangay R, Gagtasa G, Sonkaew T. Non-chemistry coupled PM10 modeling in Chiang Mai city Northern Thailand: A fast operational approach for aerosol forecasts. J of Physics: Conference 2017;901:1-6.

[6] Kanabkaew T. Prediction of hourly particulate matter concentrations in Chiangmai, Thailand using MODIS aerosol optical depth and ground-based meteorological data. Environmental Asia 2013;6(2):65-70.

[7] Trang, NH, Tripathi NK. Spatial correlation analysis between particulate matter 10 (PM10) hazard and respiratory disease in Chiang Mai province, Thailand. The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2014;XL-8:185-91.

[8] Pimpunchat B, Sirimangkhala K, Junyapoon S. Modeling haze problems in the North of Thailand using logistic regression. J of Mathematical and Fundamental Sciences 2014;46(2):183-93.

[9] Wongsathan R, Seedadan I, Wanasri S. Hybrid forecast model for PM-10 prediction: A case study of Chiang Mai city of Thailand during high season. KKU Eng J 2016; 43(S2):203-6.

[10] Wongsathan R. The Hybrid Neural Networks-ARIMA/X Models and ANFIS Model for PM-10 Forecasting: A Case Study of Chiang Mai, Thailand's High Season. Engng J CMU 2018;25(1):203-13.

[11] Chuentawat R, Kerdprasop N, Kerdprasop K. The forecast of PM10 Pollutant by using a hybrid model. Int J of Future Computer and Commun. 2017;6(3):128-32.

[12] Wongsathan R, Seedadan I. Prediction modeling of PM-10 in Chiangmai city moat by using artificial networks. J of Applied Mech and Mat 2015;781:628-31.

[13] Junpen A, Garivait S, Bonnet S, Pongpullponsak A. Spatial and temporal distribution of forest fire PM10 emission estimation by using remote sensing information. Int J of Environmental Science and Development 2011:1-7.

[14] Wongsathan R. Improvement of PM-10 forecast using ANFIS model with an integrated hotspots, Science & Technology Asia 2018;23(3):62-71.

[15] Kliengchuay W, Meeyai AC, Worakhunpiset S, Tantrakarnapa K. Relationships between meteorological parameters and particulate matter in Mae Hong Son Province, Thailand Int J of Environmental Research and Public Health 2018;15:1-13

[16] Sirimongkonlertkun N. Effect form open burning at greater Mekong sub-region nations to the PM10 concentration in northern Thailand: A case study of

backward trajectories in March 2012 at Chiang Rai province. In Proc. 1st Mae Fah Luang University Int Conf 2012:1-13.

[17] Park K, Kim J, Lee J. Visual field prediction using recurrent neural network. Scientific Reports 2019;9(8335).

[18] Blanchet FG, Legendre P, Borcard D. Forward selection of explanatory variables. Ecology 2008;89(9):2623-32.

[19] Kiatwattacharoen S, Prapamontol T, Singharat S, Chantara S, Thavornyutikarn P. Exploring the source of PM10 burning-season haze in Northern Thailand using nuclear analytical techniques. CMU J Nat Sci 2017;16(4): 307-25.

[20] Sirimongkonlertkul N, Upayokhin P, Phonekeo V. Multi-temporal analysis of haze problem in Northern Thailand: A case study in Chiang Rai Province. Kasetsart J (Nat Sci) 2013;47:768-80.

[21] Pimonsree S, Muangjai P. Situation and changes air pollution in Phayao province. Naresuan Phayao J 2011;4(3):1-9.

[22] Bonyapitak S, Kongsong R, Patpai A, Piemyai N, Sooktawee S. Estimation of potential areas impacting PM10 concentration during haze episode: Case study on Phrae Province. The 1st Naresuan Conf on Natural Resources Geoinformation and Environment 2018:9-15.

[23] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Sukakhutdinov R. Dropout: A simple way to prevent neural network from overfitting. J of Machine Learning Research 2014;15:1929-58.