Journal of Environmental Science and Technology

ISSN 1994-7887

Journal of Environmental Science and Technology 8 (2): 59-73, 2015
ISSN 1994-7887 / DOI: 10.3923/jest.2015.59.73

PM_{10} Concentrations Short Term Prediction Using Feedforward Backpropagation and General Regression Neural Network in a Sub-urban Area

[1]Ahmad Zia Ul-Saufie, [2]Ahmad Shukri Yahaya, [2]Nor Azam Ramli and [2]Hazrul Abdul Hamid

[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, 13500 Permatang Pauh, Pulau Pinang, Malaysia

[2]Clean Air Research Group, School of Civil Engineering, Universiti Sains Malaysia, Engineering Campus, 14300 Nibong Tebal, Seberang Perai Selatan, Pulau Pinang Malaysia

Corresponding Author: Ahmad Zia Ul-Saufie, Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, 13500 Permatang Pauh, Pulau Pinang, Malaysia Tel: + (60)4-3823428 Fax: + (60)4-3822766

ABSTRACT

Particulate matter has a significant impact on human health when the concentration levels of this substance exceed Malaysia Ambient Air Quality Guidelines (MAAQG). This research focused only on particulate matter with an aerodynamic diameter less than 10 μm, namely PM_{10}. Statistical modeling is required to predict future PM_{10} concentration. The aim of this study is to develop and predict next day, next two-day and next three-day PM_{10} concentration in a sub-urban area (Seberang Jaya) of Malaysia. This study used daily average monitoring records from 2001 to 2010. Two main models for predicting PM_{10} concentration were used: feedforward backpropagation and general regression neural network models. The models for the artificial neural network show that feedforward backpropagation is better than the general regression neural network with fewer errors; as much as 5.6% for next day, 3.5% for next two-day and 2.5% for next three-day predictions. These models will help local authorities to take an appropriate course of action to reduce PM_{10} concentration and could also be used as an early warning system.

Key words: Feedforward backpropagation, general regression neural network, PM_{10}, sub-urban area, future prediction

INTRODUCTION

The Department of Environment (DOE., 2006) is the government body responsible for monitoring air quality in Malaysia. Department of Environment Malaysia monitors air quality continuously via 52 stations located in urban, sub-urban and industrial areas. These monitoring stations are located in strategic locations to detect any significant change of air quality. The Malaysia Ambient Air Quality Guidelines (MAAQG, DOE., 2011) set out the allowable levels of various air pollutants. The MAAQG set the safe threshold value for PM_{10} at 150 μg m^{-3} for a 24 h averaging period and 50 μg m^{-3} per year.

Both acute and chronic human health impacts may occur when the concentration levels of air pollutants exceed the air quality guidelines (QUARG., 1996; Lee, 2010). Nasir *et al.* (1998) reported in 1997 (haze episode in Malaysia) the estimated negative effect to health for asthma attacks was

285,277. There were 118,804 cases of bronchitis in children and 3889 cases in adults, in addition to 2003 respiratory-related hospital admissions, 26,864 emergency room visits and 5,000,760 limited activity days.

Thus, particulate matter (PM_{10}) has become a challenge in managing Malaysia's air quality. Developing PM_{10} forecasting models is of key importance in PM_{10} monitoring. Artificial neural network (ANN) models can be used for predicting future PM_{10} concentrations in urban and sub-urban areas in Malaysia.

Neural network modeling has become popular as a tool for prediction modeling. This study used two types of neural networks, multi-layer perception (Feedforward Backpropagation (FFBP) and radial basis functions (General Regression Neural Network (GRNN). The FFBP model for predicting future PM_{10} concentrations has not yet been developed in Malaysia. GRNN studies for prediction of PM_{10} concentration are very few and have also not yet been developed for Malaysia. This research would bring great impact to the air quality studies in Malaysia.

MATERIALS AND METHODS

Site description: Seberang Jaya is located in North-West coast of Peninsular Malaysia. This monitoring station is located in the residential area of Sek Keb Seberang Jaya 2. Vehicular traffic is the main source of emission in Seberang Jaya. Power plants and industrial activities in a 10 km radius of the site can also be considered as possible sources of PM_{10}.

Parameter selection: Seven daily-averaged parameters from the ten years of the monitoring record base (2001-2010) were used in the study in order to gain a better understanding of PM_{10} variability. All the monitoring records provided by Department of Environment Malaysia followed the quality assurance procedure as recommended by Department of Environment Malaysia. The parameters that were used in this research are particulate matter with an aerodynamic diameter less than 10 μm (PM_{10}, μg m^{-3}), nitrogen dioxide (NO_2, ppm), wind speed (WS, km h^{-1}), sulphur dioxide (SO_2, ppm), Relative Humidity (RH %), Carbon Monoxide (CO, ppm) and ambient temperature (T, °C). Parameters influencing future PM_{10} concentrations used by previous researchers are shown in Table 1.

Feedforward backpropagation: Feedforward backpropagation (FFBP) neural networks have been selected because the model minimizes the error between target and obtained values and has in common the learning algorithms of neural networks (Sun *et al*., 2008). Figure 1 shows the procedure for the development of FFBP models.

Table 1: Summary of parameter selection in previous research

Authors	Year	WS	RH	T	SO_2	NO_2	CO	$PM_{10,t-1}$	Others
Papanastasiou *et al*. (2007)	2001-2003	x		X				x	
Perez and Reyes (2002)	1998-2000	x	x	X				x	Max T, ΔT, average T and WS for next day
Paschalidou *et al*. (2011)	2006-2008		x	X	x	x	x	x	O_3, WD, solar radiation, rainfall,
Sfetsos and Vlachogiannis (2010)	2000-2004		x	X				x	barometric pressure
Hooyberghs *et al*. (2004)	1997-2001	x		X				x	DoW, cloud cover, WD
Brunelli *et al*. (2007)	2003-2004			X					Barometric pressure, WD
Slini *et al*. (2006)	1994-2000	x		X					Mean and max of WS and T
Nejadkoorki and Baroutian (2012)	2001-2009	x		X			x	x	NO, WD, solar radiation
Grivas and Choloulakou (2006)	2001-2002	x	x	X				x	WD, Rainfall, DoW

DoW: Day of the week and ΔT: T_{max}-T_{min}

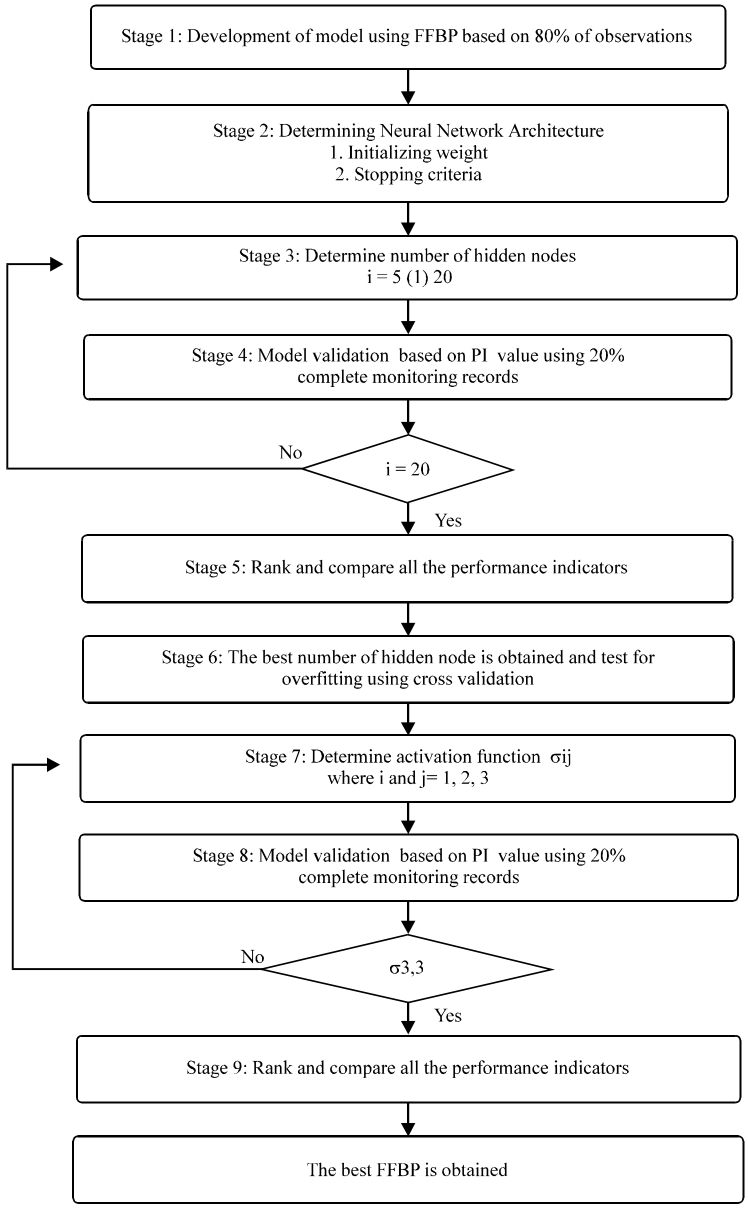

Fig. 1: Procedure for development of FFBP models

In the first stage, development of the model using FFBP is based on 80% of observations, where future PM_{10} is considered as a dependent variable and WS, RH, T, SO_2, NO_2, PM_{10} and CO as independent variables. Defining neural network architecture is a very important step to enable FFBP models to achieve the optimum result. In this research three layers of neural network were used, i.e. an input layer, a hidden layer and output layer (Fig. 2) (Singh and Deo, 2007).

There are two main phases in FFBP, forward pass phase and backward pass phase. First, in forward pass phase, the training monitoring record set is propagated through the hidden layer and comes out of the neural network through the output layer. The output values are then compared to actual target output values. The residual (error) between output layer and actual values is calculated and propagated back towards the hidden layer. In this phase, the modifications to the

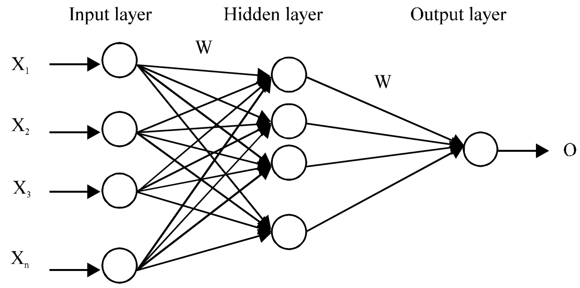

Fig. 2: Architecture of a feedforward backpropagation neural network (FFBP) (Singh and Deo, 2007)

connection strengths are made based on the differences between computed and observed information monitoring records at the output units (Viotti *et al*., 2002; Cigizoglu and Kisi, 2006). Adjustment to the weighting was made, so that error was reduced with each alteration and the FFBP models approaching the desired outputs (Sun *et al*., 2008). The advantage of this method lies in the ability of the model to adjust itself during the training phase, by comparing the actual and obtained values with low error rate.

In stage two, initializing weight and stopping criteria is also considered. Initializing weight and bias are normally selected randomly; this research used small random values. Maximum error used as stopping criteria is set at 0.05. In this study the training process was set at maximum 10,000 iterations, or until it reached the maximum error.

Levenberg-Marquardt optimization was used as a network training function to update weight and bias values. This method was chosen because this is the fastest algorithm in the Matlab toolbox and is recommended for supervised learning (MATLAB., 2012).

Stage three is an important in FFBP because the number of neurons in the hidden layer has strong influence on the output. The optimum number of neurons is very important, because too few neurons will contribute to under-fitting while too many neurons will lead to over-fitting. The number of neurons is also related to the complexity of mapping between input and output, compared to the amount of noise and the amount of training monitoring record.

A variety of techniques and methods have been suggested by researchers to determine the number of hidden nodes. A combination of techniques suggested by Heaton (2008) and Charytoniuk and Chen (2000) were used in this study to determine the number of nodes in the hidden layer. After running a trial and error method with a maximum twice input layer, ranking and comparison of all performance indicators in stage five, the best FFBP based on the number of hidden nodes was obtained in stage six. In stage six, tests for overfitting are required using cross validation, as recommended by Sterlin (2007), for the purpose of preventing overfitting in FFBP models. Then, the monitoring records were randomly divided into five folders. Figure 3 illustrates the cross validation technique. This study used 20% for each folder. All five combinations were run using 80% for training and 20% for testing, then the error was calculated for all of the sets. If the values error for training were less than the values error for the test set, this indicated that no overfitting occurred.

Stage seven focused on determining the best activation function for the FFBP model. Activation functions are important because they have a significant impact on the applicability of training algorithms. Kamruzzaman and Aziz (2002) suggested the use of a sigmoid activation function

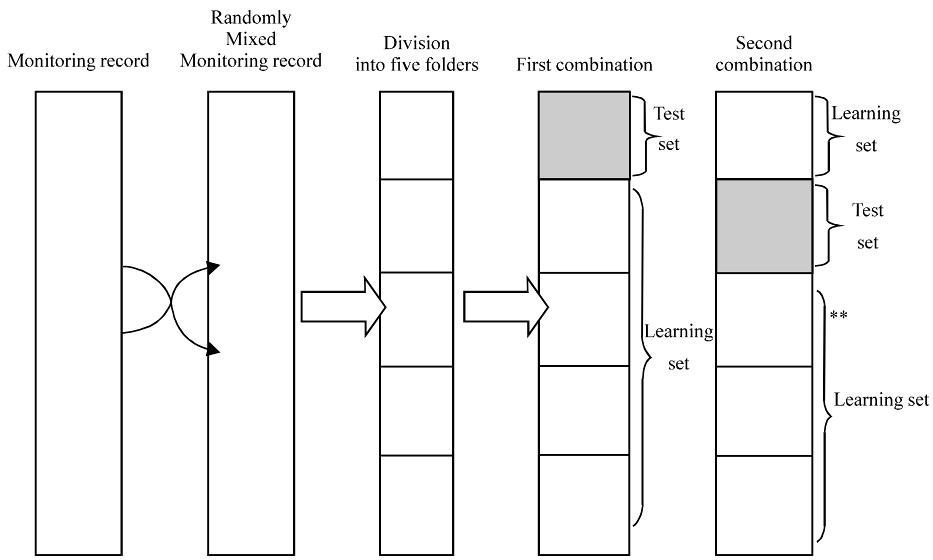

Fig. 3: Illustration of cross validation technique (Sterlin, 2007)

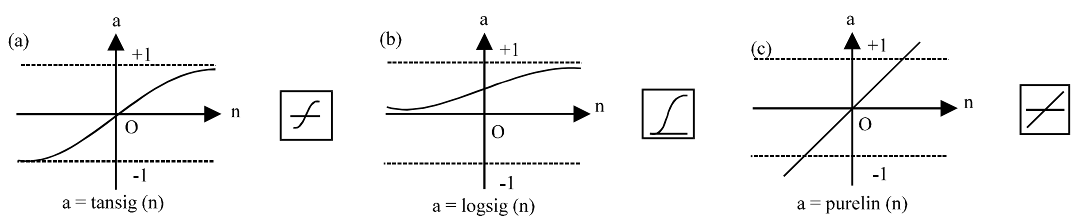

Fig. 4(a-c): Sigmoid transfer function (MATLAB., 2012)

because sigmoid units are easier to train than other activation functions. Besides that, a small change in the weights will usually produce a change in outputs, making it possible to tell whether that change in the weights is accurate or not. All the sigmoid transfer functions that were used are shown in Fig. 4. They are the tangent sigmoid (Fig. 4a), log sigmoid (Fig. 4b) and linear transfer functions (Fig. 4c).

The FFBP model was obtained and the model was validated based on the performance indicator value using 20% complete observed monitoring records. All the procedures were repeated until all combinations of sigmoid transfer function were completed. Then, all performance indicators were ranked to obtain the best FFBP.

General regression neural network: A general regression neural network is one of the methods for curve fits. This method tries to adjust the values of the function coefficient such that the function closely approximates the available monitoring record. FFBP is based on equations that are connected using weighting factors but GRNN uses statistical distribution to select the equation within the structure. Figure 5 shows the procedure for development of GRNN models.

Development of GRNN models using 80% of the monitoring records is shown in stage one. Stage two was to obtain the GRNN model, using smoothness function (σ) values of 0.1(0.1)0.9. In GRNN the critical step for good prediction is selection of the smoothness parameter (σ), so this study used a trial and error method to get the best smoothness function for predicting future PM_{10} concentrations.

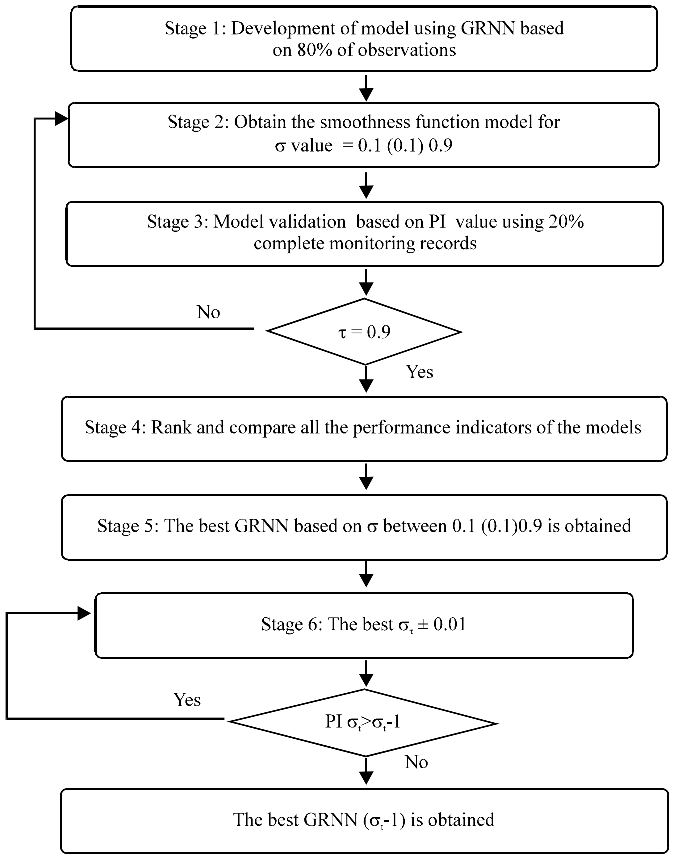

Fig. 5: Procedure for development of GRNN models

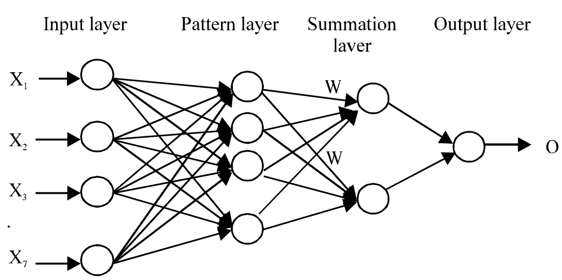

Fig. 6: Architecture of general regression neural network (Chen and Yu, 2009)

Figure 6 shows the architecture of the general regression neural network. The input layer used all the independent variables, i.e. SO_2, NO_2, CO, RH, T, WS and PM_{10}. The relation between input and output is shown as (Chen and Yu, 2009):

$$y = \frac{\sum_{j=1}^{m} w_j \phi_j(X)}{\sum_{j=1}^{m} \phi_j(X)} = \frac{\alpha}{\beta}$$

Table 2: Performance indicators for measuring accuracy and errors of model

Performance indicators	Equation	Description
Root Mean Square Error (RMSE)	$\text{RMSE} = \frac{1}{n-1}\sum_{i=1}^{n}(P_i - O_i)^2$	The best model should have values closer to zero
Normalized Absolute Error (NAE)	$\text{NAE} = \frac{\sum_{i=1}^{n}\text{Abs}(P_i - O_i)}{\sum_{i=1}^{n} O_i}$	
Index of agreement	$\text{IA} = 1 - \left[\frac{\sum_{i=1}^{n}(P - O_i)^2}{\sum_{i=1}^{n}\left(\lvert P_i - \bar{O}\rvert + \lvert O_i - \bar{O}\rvert\right)^2}\right]$	The best model should have values closer to one
Prediction accuracy	$\text{PA} = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(P_i - \bar{P})(O_i - \bar{O})}{\sigma_p \sigma_O}$	
Coefficient of Determination (R^2)	$R^2 = \left(\frac{\sum_{i=1}^{n}(P_i - \bar{P})(O_i - \bar{O})}{n.S_{pred}.S_{obs}}\right)^2$	

n: Number of observations, O_i: Observed values, P_i: Predicted values, $\bar{O}$: Mean of the observed values, $\bar{P}$: Mean of the predicted values, S_{obs}: Standard deviation of the observed values and S_{pred}: Standard deviation of the predicted values

where, w_j is weight between the jth pattern and summation layer node $X = (x_1, x_2,..., x_7)$ is dimensional of input vector and φ is Gaussian function. After the GRNN model was obtained in stage two, the model needed to be validated using the remaining 20% of the monitoring record. In stage three, performance indicators were used to measure the accuracy and error of the model.

Ranking and comparison of all the performance indicators at all smoothness functions was conducted in stage four. The best smoothness function was σ_t found for the range 0.1(0.1)0.9 in stage five. To increase accuracy, further analysis was carried out by addition or subtraction of the best smoothness function in the range of $\sigma_t(0.01)\sigma_{t+1}$ and $\sigma_{t-1}(0.01)\sigma_t$, where, σ_t is the best smoothness function and $\sigma_{t\pm1}$ is the smoothness function after or before the best smoothness function shown in stage six. Performance indicators were compared and the best GRNN models were obtained.

Performance indicators: The model's performance in both development and validation processes was evaluated by calculating performance indicators. Performance indicators that have been used are, (a) coefficient of determination (R^2), Root Mean Square Error (RMSE), Index of Agreement (IA), Normalized Absolute Error (NAE) and Prediction Accuracy (PA). The equations used are given in Table 2 (Yusof *et al.*, 2010; Lu, 2004).

RESULT AND DISCUSSION

The Seberang Jaya monitoring station is located near Seberang Perai monitoring station. However, the Seberang Jaya monitoring station is surrounded by major roads such as Projek Lebuhraya Utara-Selatan Berhad (PLUS) Expressway, Butterworth-Kulim Expressway (BKE) and Federal Road. On 8 August 2005 Seberang Jaya was affected by land and forest fires in the Riau Province of Central Sumatra, Indonesia, with a maximum concentration of 247.583 μg m^{-3} as shown in Table 3. The average PM_{10} concentration exceeds than 50 μg m^{-3}.

Feedforward backpropagation models: The main problem faced using the FFBP method was deciding on the FFBP network architecture, such as activation function and number of neurons in hidden layer. Table 4 gives the validation models using different numbers of neurons based on a few performance indicators. The initial value of neurons was five and the number was increased by one unit until optimal value was achieved. Table 4 shows that 12 is the best number of neurons in hidden layer for next-day predictions at Seberang Jaya.

Table 5 shows the result for testing overfitting using cross validation. Table 5 shows that there was not a problem overfitting for Seberang Jaya, because the average result for testing validation (0.1721) was larger than the training set (0.1602).

Table 3: Descriptive statistics for Seberang Jaya

	PM_{10} (μg m^{-3})	WS (km h^{-1})	T (°C)	RH (%)	SO_2 (ppm)	NO_2 (ppm)	CO (ppm)
Mean	55.073	4.944	71.411	27.801	0.003	0.015	0.757
Median	51.000	4.908	72.000	27.838	0.003	0.015	0.725
S.D.*	20.103	1.103	7.147	1.362	0.002	0.003	0.234
Skewness	1.924	-0.058	-1.044	-0.322	1.364	0.809	1.064
Kurtosis	7.499	1.166	2.894	0.047	3.723	4.556	2.386
Max	247.583	11.000	91.000	31.500	0.016	0.049	2.350

*S.D: Standard deviation

Table 4: Validation models using different number of neurons

No. of hidden node	NAE	RMSE	IA	PA	R^2
5	0.1639	12.7980	0.8550	0.7514	0.5623
6	0.1641	13.0185	0.8486	0.7416	0.5477
7	0.1641	13.0185	0.8486	0.7416	0.5477
8	0.1711	13.2974	0.8500	0.7385	0.5430
9	0.1676	12.9083	0.8527	0.7508	0.5613
10	0.1718	13.6519	0.8418	0.7279	0.5276
11	0.1675	13.0132	0.8538	0.7466	0.5550
12	0.1657	12.8132	0.8547	0.7519	0.5630
13	0.1770	13.5319	0.8476	0.7366	0.5403
14	0.1780	13.7073	0.8394	0.7297	0.5302
15	0.2171	35.0265	0.4434	0.2820	0.0792
16	0.1694	12.9507	0.8628	0.7542	0.5664

Table 5: Results for overfitting problem using cross validation

Set	Group	NAE
First set	Training	0.1597
	Test	0.1860
Second set	Training	0.1572
	Test	0.1736
Third set	Training	0.1593
	Test	0.1718
Fourth set	Training	0.1611
	Test	0.1701
Fifth set	Training	0.1639
	Test	0.1588

Table 6: Results using different activation functions using original parameters as inputs

Transfer function A	Transfer function B	NAE	RMSE	IA	PA	R^2
Tansig	Purelin	0.1657	12.8132	0.8547	0.7519	0.5630
Tansig	Logsig	1.4757	80.9425	0.2409	0.2604	0.0676
Tansig	Tansig	0.1670	12.8556	0.8523	0.7390	0.5441
Logsig	Logsig	1.4757	80.9425	0.2409	0.0000	0.0000
Logsig	Tansig	0.1721	13.1380	0.8454	0.7326	0.5346
Logsig	Purelin	0.1731	13.2769	0.8473	0.7431	0.5502
Purelin	Purelin	0.1651	12.5254	0.8499	0.7478	0.5571

Table 7: Summary of FFBP model for next-day

Day	No. of neurons	Transfer function A*	Transfer function B**	NAE	RMSE	IA	PA	R^2
Next day	12	Tansig	Purelin	0.1657	12.8132	0.8547	0.7519	0.5630
Next two-days	12	Tansig	Purelin	0.2002	14.7135	0.7796	0.6514	0.4225
Next three-days	12	Tansig	Purelin	0.2154	18.8226	0.6033	0.5406	0.2910

*Transfer function A: Transfer function from input layer to hidden layer, **Transfer function B : Transfer function from hidden layer to output layer

After determining the best number of neurons (hidden node) in the hidden layer, the best transfer function was obtained. The best transfer functions are obtained using sigmoid transfer functions. This gives the best weights for the activation function. In this study three sigmoid functions were used i.e., linear (purelin), log-sigmoid (logsig) and hyperbolic tangent sigmoid (tansig).

Table 6 shows performance indicators using the three different transfer functions. The best transfer function for Seberang Jaya (next-day) from input to hidden layer is Tangent Sigmoid (tansig) and activation function from hidden layer to output layer is linear transfer function (purelin).

Figure 7 shows the scatter plot of predicted against observed PM_{10} concentrations in Seberang Jaya. UL and LL are upper and lower 95% confident limits for FFBP models. The results shows that 95.83% of next day (Fig. 7a), 96.03% of next two-day (Fig. 7b) and 97.51% of next three-day (Fig. 7c) monitoring records fall within the confidence limits.

Repeating the procedure revealed the best number of neurons and transfer functions for next two-day and next three-day predictions for Seberang Jaya. Table 7 showed the best model using FFBP for (a) Next-day, (b) Next two-days and (c) Next three-days for all monitoring stations. None of the FFBP models had a problem with over-fitting after being tested using cross validation. For transfer functions from the hidden layer and output layer, linear transfer function (purelin) gives the best result. For the input layer to output layer, tangent sigmoid function is usually chosen as the best transfer function.

General regression neural network: In GRNN the critical step for good prediction is selection of the smoothness parameter (spread). This study used a trial and error method to get the best smoothness function for predicting future PM_{10} concentrations. Table 8 show performance indicators for the GRNN model using different smoothing factors. Performance indicators were used to select the best smoothing function for predicting PM_{10} concentration for next-day at Seberang Jaya. Ranking all five performance indicators, the 0.2 smoothing function gives the best GRNN models for PM_{10} concentration.

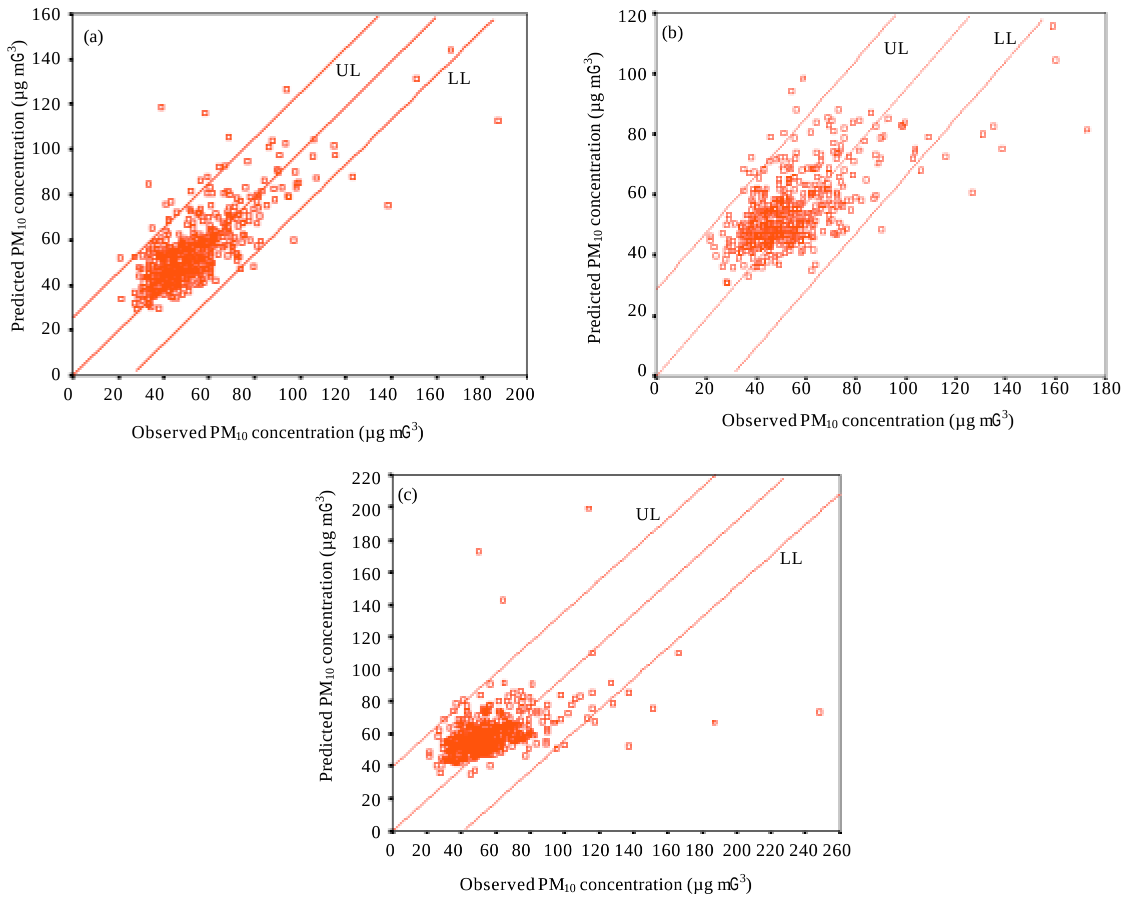

Fig. 7(a-c): Scatter plot of predicted against observed PM_{10} concentrations (μg m^{-3}) for FFBP models

Table 8: GRNN result using different smoothing functions for next-day (step one)

Smoothing function	NAE	RMSE	IA	PA	R^2
0.1	0.1903	14.1536	0.8050	0.6718	0.4496
0.2	0.1789	13.2534	0.7999	0.7069	0.4978
0.3	0.1848	13.5454	0.7527	0.7128	0.5062
0.4	0.1974	14.7772	0.6636	0.7060	0.4963
0.5	0.2077	15.7066	0.5668	0.6865	0.4693
0.6	0.2164	16.4222	0.4795	0.6693	0.4460
0.7	0.2234	16.9531	0.4065	0.6599	0.4336
0.8	0.2289	17.3542	0.3473	0.6556	0.4281
0.9	0.2332	17.6626	0.3006	0.6530	0.4246

Therefore, 0.2 was used as smoothing function to represent PM10 concentration models in the Seberang Jaya monitoring station for next-day prediction. To obtain a more accurate model, further analysis was carried out taking a smoothing function in the range of 0.2±0.01 i.e., 0.1(0.01)0.2 or 0.2(0.01)0.3. Table 9 shows performance indicators for smoothing functions values 0.2±0.01 and which shows the best smoothing function to be 0.19. This step will improve the accuracy of GRNN models.

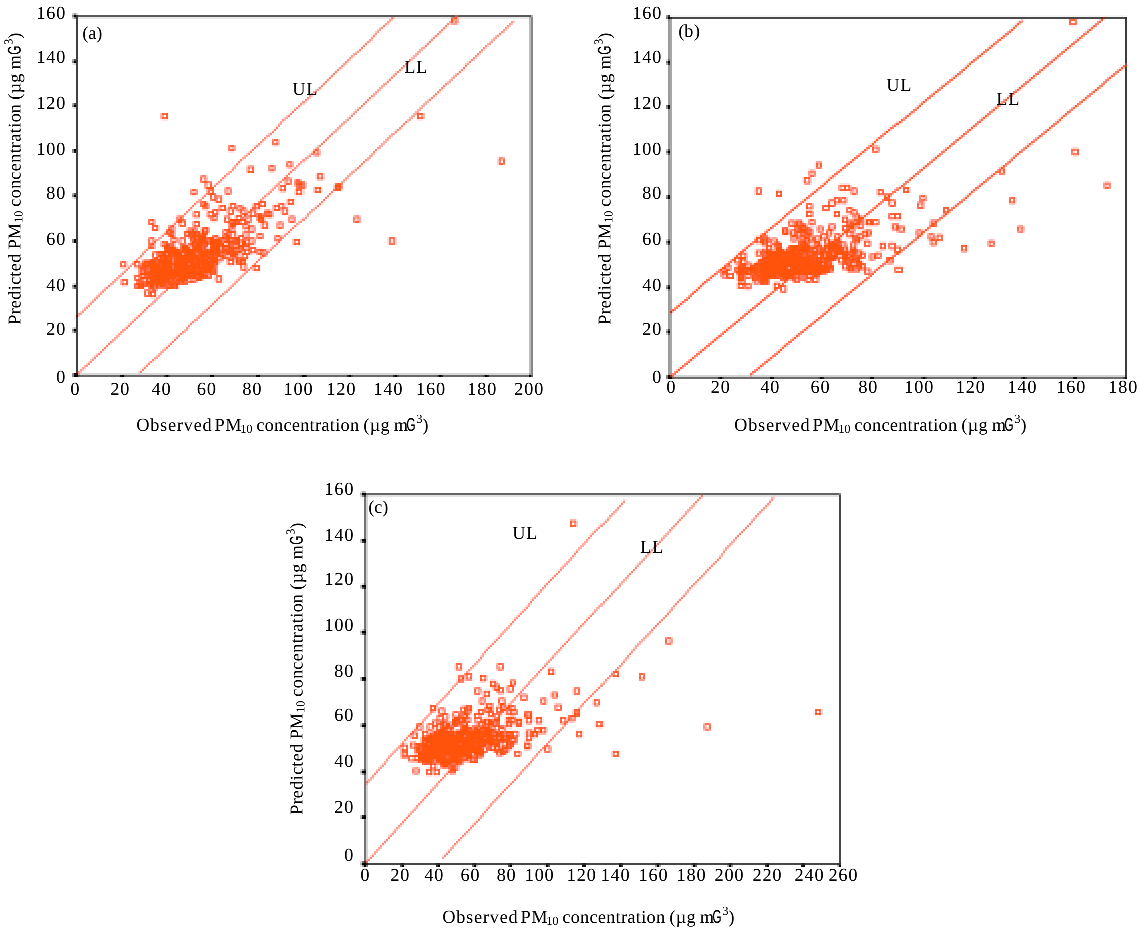

Fig. 8(a-c): Scatter plot of predicted against observed PM_{10} concentrations (µg m^{-3}) for GRNN models

Table 9: GRNN result using different smoothing functions for next-day (step two)

Smoothing function	NAE	RMSE	IA	PA	R^2
0.2	0.1789	13.2534	0.7999	0.7069	0.4978
0.21	0.1793	13.2442	0.7970	0.7081	0.4995
0.19	0.1785	13.2698	0.8024	0.7056	0.4960
0.18	0.1783	13.2936	0.8046	0.7042	0.4941

Table 10: Summary of GRNN models for all prediction days

Day	Smoothing function	NAE	RMSE	IA	PA	R^2
Next day	0.19	0.1785	13.2698	0.8024	0.7056	0.4960
Next two-days	0.21	0.2078	15.1767	0.6939	0.5996	0.3581
Next three-days	0.22	0.2247	18.9447	0.5460	0.5008	0.2499

The scatter plots were plotted for the predicted versus observed values of PM_{10} for next-day, next two-days and next three-days as shown in Fig. 8. The results show that 97.08% of next-day (Fig. 8a), 96.66% of next two-day (Fig. 8b) and 97.07% of next three-day (Fig. 8c) monitoring records fall within the confidence limits.

Table 11: Comparison performance indicator for FFBP and GRNN models

	Next day		Next two-days		Next three-days	
	FFBP	GRNN	FFBP	GRNN	FFBP	GRNN
NAE	0.1657	0.1785	0.2002	0.2078	0.2154	0.2247
RMSE	12.8132	13.2698	14.7135	15.1767	18.8226	18.9447
IA	0.8547	0.8024	0.7796	0.6939	0.6033	0.5460
PA	0.7519	0.7056	0.6514	0.5996	0.5406	0.5008
R^2	0.5630	0.4960	0.4225	0.3581	0.2910	0.2499

Table 12: Comparison of PM_{10} models using artificial neural network approaches

Author	Area	Methodology	R2	IA
Corani (2005)	Milan, Italy	MLP	n/a	0.94
Papanastasiou *et al*. (2007)	Volos, Greece	MLP (FFNN)	0.61	0.87
	Warsaw, Poland	MLR	0.55	0.86
Siwek and Osowski (2012)	Tehran, Iran	FFBP	0.53	0.69
	Chile	RBF	0.58	0.74
Nejadkoorki and Baroutian (2012)	Brazil	MLP (FFBP)	0.51-0.73	0.83
Diaz-Robles *et al*. (2008)	Athens and Helsinki, Greece	MLP (FFBP)	0.65	0.77
Lira *et al*. (2007)	Seberang Jaya,	MLP (FFBP)	0.7591	0.914
Vlachogianni *et al*. (2011)	Malaysia	MLR	0.60-0.90	n/a
This paper		MLP	0.67-0.91	0.8547
		FFBP (next day)	0.563	0.7796
		FFBP (next two-days)	0.4225	0.6033
		FFBP (next three-days)	0.291	0.8024
		GRNN (next day)	0.496	0.6939
		GRNN (next two-days)	0.3581	0.546
		GRNN (next three-days)	0.2499	

Repeating the procedure revealed the best number of neurons and transfer functions for next two-days and next three-days for Seberang Jaya. Table 10 shows the best model using GRNN next-day, next two-days and next three-days. Table 10 shows the range of the smoothing factor for all monitoring stations is from 0.07-0.24.

Comparison of performance: Table 11 showed the comparison of performance indicators between (Multiple Linear Regressions (MLR) and GRNN models. In general, both models show the same results for performance indicators, reflecting greater accuracy in next-day (0.6956) PM_{10} concentration predictions compared with average accuracy in the next two-day (0.5842) and next three-day predictions (0.4553). Table 11 also shows FFBP models perform better than GRNN models for all three days in predicting PM_{10} concentration, with fewer errors, as much as 5.6% for next day, 3.5% for next two-day and 2.5% for next three-day predictions. The accuracy of the model appears to be far greater than the GRNN model, with improvements of 8.1% for next day, 11.4% for next two-day and 10.3% for next three-day predictions. The results also showed that FFBP and GRNN could predict future PM_{10} concentration accurately for the next three-days for this monitoring station.

Table 12 describes the use of ANN in predicting PM_{10} concentrations. In Table 12, it was found the ANNs are used widely for predicting PM_{10} concentration, but from the current knowledge none of the FFBP and GRNN models for predicting PM_{10} concentration have been used in Malaysia. The results of this study show the value of Index of Agreement (IA) is from 0.69-0.94 for next day prediction and coefficients of determination (R^2) were between 0.51-0.91 for next day prediction only. These results show similar values of R^2 and IA with previous researchers, as shown in Table 12.

CONCLUSION

The FFBP and GRNN approaches have been proven to be effective as techniques to predict future next day, next two-day and next three-day PM_{10} concentrations in advance for a sub-urban area in Malaysia. The results show that FFBP performs better than GRNN because FFBP has a more complex structure than GRNN. FFBP gives more accurate results when many alterations have been performed to reduce the error between observed and predicted values. GRNN has a simple structure, is less time-consuming than FFBP and is based on curve-fitting prediction. Both FFBP and GRNN give useful information on PM_{10} status and could be used government bodies to reduce the impact of air pollution in Seberang Jaya, Malaysia.

ACKNOWLEDGMENT

This study was funded by Universiti Teknologi MARA under Grant FRGS/1/2014/STWN01/UITM/02/4. Thank you to Universiti Sains Malaysia and Universiti Teknologi MARA for providing financial support to carry out this study and also thanks to the Department of Environment Malaysia for their support.

REFERENCES

Brunelli, U., V. Piazza, L. Pignato, F. Sorbello and S. Vitabile, 2007. Two-days ahead prediction of daily maximum concentrations of SO_2, O_3, PM_{10}, NO_2, CO in the urban area of Palermo, Italy. Atmos. Environ., 41: 2967-2995.

Charytoniuk, W. and M.S. Chen, 2000. Neural network design for shirt term load forecasting. Proceedings of the Electric Utility Deregulation and Restructing and Power Technologies, April 4-7, 2000, London, pp: 554-561.

Chen, T.C. and C.H. Yu, 2009. Motion control with deadzone estimation and compensation using grnn for TWUSM drive system. Expert Syst. Applic., 36: 10931-10941.

Cigizoglu, H.K. and O. Kisi, 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol., 317: 221-238.

Corani, G., 2005. Air quality prediction in Milan: Feed-forward neural networks, pruned neural networks and lazy learning. Ecol. Mod., 185: 513-529.

DOE., 2006. Malaysia environmental quality report 2005. Department of Environment, Ministry of Natural Resources and Environment, Putrajaya, Malaysia.

DOE., 2011. Malaysia environmental quality report 2010. Department of Environment, Ministry of Natural Resources and Environment, Kuala Lumpur, Malaysia.

Diaz-Robles, L.A., J.C. Ortega, J.S. Fu, G.D. Reed, J.C. Chow, J.G. Watson and J.A. Moncada-Herrera, 2008. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. Atmos. Environ., 42: 8331-8340.

Grivas, G. and A. Chaloulakou, 2006. Artificial neural network models for prediction of PM_{10} hourly concentrations, in the greater area of Athens, Greece. Atmos. Environ., 40: 1216-1229.

Heaton, J., 2008. Introduction to Neural Networks for Java. 2nd Edn., Heaton Research, Inc., USA., ISBN: 9781604390087, Pages: 440.

Hooyberghs, J., C. Mensink, G. Dumont, F. Fierens and O. Brasseur, 2004. Short term PM_{10} forecasting: A survey of possible input variables. Proceedings of the International Conference on Modelling, Monitoring and Management of Air Pollution, June 30-July 2, 2004, Rhodes, Greece, pp: 171-178.

Kamruzzaman, J. and S.M. Aziz, 2002. A note on activation function in multilayer feedforward learning. Proceeding of the International Joint Conference on Neural Network, May 12-17, 2002, Honolulu, USA., pp: 519-523.

Lee, S.W., 2010. Fine particulate matter measurement and international standardization for air quality and emissions from stationary sources. Fuel, 89: 874-882.

Lira, T.S., M.A.S. Barrozo and A.J. Assis, 2007. Air quality prediction in Uberlandia, Brazil, using linear models and neural networks. Comput. Aided Chem. Eng., 24: 51-56.

Lu, H.C., 2004. Estimating the emission source reduction of PM_{10} in central Taiwan. Chemosphere, 54: 805-814.

MATLAB, 2012. MATLAB Version R2010a: Computer Software. The MathWorks Inc., USA.

Nasir, M.H., M. Awang, R. Abdullah, M.S. Othman and J. Othman *et al*., 1998. The damage costs of haze episodes in Malaysia-concepts, empirical studies and improvement policies. Proceedings of the Regional Workshop on Transboundary Pollution from Fires: Haze Impacts on Health, October 27, 1998, Universiti Putra Malaysia.

Nejadkoorki, F. and S. Baroutian, 2012. Forecasting extreme PM_{10} concentrations using artificial neural networks. Int. J. Environ. Res., 6: 277-284.

Papanastasiou, D.K., D. Melas and I. Kioutsioukis, 2007. Development and assessment of neural network and multiple regression models in order to predict PM_{10} levels in a medium-sized Mediterranean city. Water Air Soil Pollut., 182: 325-334.

Paschalidou, A.K., S. Karakitsios, S. Kleanthous and P.A. Kassomenos, 2011. Hourly PM_{10} concentration in cyprus through artificial neural networks and multiple regression models: Implications to local environmental management. Environ. Sci. Poll. Res., 18: 316-327.

Perez, P. and J. Reyes, 2002. Prediction of maximum of 24-h average of PM_{10} concentrations 30h in advance in Santiago, Chile. Atmos. Environ., 36: 4555-4561.

QUARG., 1996. Airborne particulate matter in the United Kingdom. Third Report of the Quality of Urban Air Review Group (QUARG), Prepared at the Request of the Department of the Environment, London, UK., May 1996.

Sfetsos, A. and D. Vlachogiannis, 2010. A new methodology development for the regulatory forecasting of PM_{10}. Application in the Greater Athens Area, Greece. Atmospheric Environ., 44: 3159-3172.

Singh, P. and M.C. Deo, 2007. Suitability of different neural networks in daily flow forecasting. Applied Soft Comput., 7: 968-978.

Siwek, K. and S. Osowski, 2012. Improving the accuracy of prediction of PM_{10} pollution by the wavelet transformation and an ensemble of neural predictors. Eng. Applic. Artif. Intell., 25: 1246-1258.

Slini, T., A. Kaprara, K. Karatzas and N. Moussiopoulos, 2006. PM_{10} forecasting for Thessaloniki, Greece. Environ. Model. Software, 21: 559-565.

Sterlin, P., 2007. Overfitting prevention with cross-validation. M.Sc. Thesis, University Pierre and Marie Curie, Paris, France.

Sun, G., S.J. Hoff, B.C. Zelle and M.A. Nelson, 2008. Development and comparison of backpropagation and generalized regression neural network models to predict diurnal and seasonal gas and PM_{10} concentrations and emissions from swine buildings. Trans. Am. Soc. Agric. Biol. Eng., 51: 685-694.

Viotti, P., G. Liuti and P. Di Genova, 2002. Atmospheric urban pollution: Applications of an Artificial Neural Network (ANN) to the city of Perugia. Ecol. Model., 148: 27-46.

Vlachogianni, A, P. Kassomenos, A. Karppinen, S. Karakitsios and J. Kukkonen, 2011. Evaluation of a multiple regression model for the forecasting of the concentrations of NO_x and PM_{10} in Athens and Helsinki. Sci. Total Environ., 409: 1559-1571.

Yusof, N.F.F.M., N.A. Ramli, A.S. Yahaya, N. Sansuddin, N.A. Ghazali and W. Al Madhoun, 2010. Monsoonal differences and probability distribution of PM_{10} concentration. Environ. Monitor. Assess., 163: 655-667.