ENVIRONMENTAL ENGINEERING SCIENCE
Volume 30, Number 12, 2013

DOI: 10.1089/ees.2013.0164

Improving Neural Network Prediction Accuracy for PM_{10} Individual Air Quality Index Pollution Levels

Qi Feng,[1,*] Shengjun Wu,[2] Yun Du,[1] Huaiping Xue,[1] Fei Xiao,[1] Xuan Ban,[1] and Xiaodong Li[1]

[1]*Key Laboratory for Environment and Disaster Monitoring and Evaluation, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China.*
[2]*Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China.*

Received: May 22, 2013 *Accepted in revised form: September 10, 2013*

Abstract

Fugitive dust deriving from construction sites is a serious local source of particulate matter (PM) that leads to air pollution in cities undergoing rapid urbanization in China. In spite of this fact, no study has yet been published relating to prediction of high levels of PM with diameters $<10\,\mu$m (PM_{10}) as adjudicated by the Individual Air Quality Index (IAQI) on fugitive dust from nearby construction sites. To combat this problem, the Construction Influence Index (*Ci*) is introduced in this article to improve forecasting models based on three neural network models (multilayer perceptron, Elman, and support vector machine) in predicting daily PM_{10} IAQI one day in advance. To obtain acceptable forecasting accuracy, measured time series data were decomposed into wavelet representations and wavelet coefficients were predicted. Effectiveness of these forecasters were tested using a time series recorded between January 1, 2005, and December 31, 2011, at six monitoring stations situated within the urban area of the city of Wuhan, China. Experimental trials showed that the improved models provided low root mean square error values and mean absolute error values in comparison to the original models. In addition, these improved models resulted in higher values of coefficients of determination and *AHPC* (the accuracy rate of high PM_{10} IAQI caused by nearby construction activity) compared to the original models when predicting high PM_{10} IAQI levels attributable to fugitive dust from nearby construction sites.

Key words: construction site; fugitive dust; neural network; PM_{10}; pollution

Introduction

Construction of buildings and infrastructure can produce significant emissions as a result of activities common to construction sites. Throughout the construction period, uncontrolled fugitive dust emissions can present serious environmental, health, and operational problems that impact both site personnel and nearby communities (Ashbaugh *et al.*, 2003; Ho *et al.*, 2003; Dorevitch *et al.*, 2006; Kumar *et al.*, 2012).

An air quality index (AQI) is a quantitative measure used to uniformly report on the air quality of different constituents with respect to human health (Ministry of Environmental Protection, 2012). PM_{10} (particulate matter with a diameter $<10\,\mu$m) The Individual Air Quality Index (IAQI) is a conversion of PM_{10} (Ministry of Environmental Protection, 2012), one of the primary pollutants afflicting China today (Chan and Yao, 2008; Ministry of Environmental Protection, 2009–2011). It is measured at sampling stations on a 0:500 scale. A PM_{10} IAQI result of 100 corresponds to the short-term "PM_{10} air quality objective" established by the Air Pollution Control Ordinance. The Ministry of Environmental Protection of the People's Republic of China classifies PM_{10} air quality standards into six major categories with respect to PM_{10} IAQI values (Table 1): I (clean), II (good), III (low-level pollution), IV (mid-level pollution), V (high-level pollution), and VI (serious pollution).

Forecasting models can be used to identify in advance what regulations should be enforced when an AQI exceeds acceptable values. This would prevent unnecessary annoyances and potential health risks to urban inhabitants.

Recently, model forecasting using various artificial neural networks (ANNs) has been shown to be an effective tool when planning health warning systems related to air quality and PM_{10} pollution (Brunelli *et al.*, 2007). For example, Morabito and Versaci (2003) have proposed the use of hybrid fuzzy neural systems for modeling and predicting time series of pollutant concentration levels in Italy. Similarly, Kukkonen *et al.* (2003) compared the performance of five different NN models for the prediction of PM_{10} concentrations in Helsinki. Results obtained showed that NN models performed better

*Corresponding author: Key Laboratory for Environment and Disaster Monitoring and Evaluation, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, No. 340 Xudong Road, Wuhan 430077, China. *Phone:* +862768881075; *Fax:* +862768881362; *E-mail:* fengqi@asch.whigg.ac.cn

Table 1. Daily PM_{10} Individual Air Quality Index and Air Quality Management in China

PM_{10} IAQI	*Daily PM_{10} concentration ($\mu g/m^3$)*	*Air quality classification*	*Health influence*	*Air quality description and management*
≤50	≤50	I (clean)	No	No action is required.
51–100	50–150	II (good)	No	No action is required.
101–150	150–250	III (low-level pollution)	Minor but aggravating symptoms in healthy people.	People with respiratory disease should be cautioned when participating in outdoor activities.
151–200	250–350	IV (mid-level pollution)	Symptoms start to become evident in healthy people.	Healthy people are advised to take appropriate action to reduce outdoor activities.
201–300	350–420	V (high-level pollution)	Patients with heart disease and pulmonary symptoms are notably affected. A reduction in endurance commonly appears in healthy people when active outdoors.	Air pollution is severe. Consequently, the general public is advised to reduce physical exertion and outdoor activities.
>300	>420	VI (serious pollution)	Healthy people exhibit obvious and intense symptoms, while participating in outdoor activities. Certain diseases develop prematurely.	The general public is advised to avoid outdoor activities altogether.

PM_{10}, particulate matter with a diameter $<10\,\mu$m; IAQI, Individual Air Quality Index.

than linear models. In addition, Jiang *et al.* (2004) used an enhanced multilayer perceptron (MLP) network to formulate API predictions in Shanghai, while Hooyberghs *et al.* (2005) described the development of an MLP NN to forecast daily average PM_{10} concentrations in urban areas in Belgium one day in advance.

One main benefit in PM_{10} prediction is its ability to predict pollution events or high pollution concentrations so that local residents or commuters can adjust their activities in response. Accordingly, a few studies have been published on models that can forecast high levels of PM_{10} pollution. For example, Grivas *et al.* (2006) used a genetic algorithm optimization procedure to select input variables to improve MLP network performance. It was reported to perform well in predicting high PM_{10} concentrations in Greece. In addition, Perez and Reyes (2006) developed an integrated ANN to forecast maximum values of daily PM_{10} concentrations in Santiago, Chile. Cai *et al.* (2009) presented methods in forecasting hourly air pollutant concentrations in Guangzhou, China, using a backpropagation NN. Paschalidou *et al.* (2011) used MLP and radial basis function NN, as well as a principal component regression analysis to make reliable forecasting of hourly PM_{10} concentrations in Cyprus. Wu *et al.* (2011) considered dust storms when improving the Elman network in predicting PM_{10} API in Wuhan, China. Nejadkoorki and Baroutian (2012) used the Levenberg–Marquardt method to optimize MLP, while also incorporating gaseous pollutants to predict maximum PM_{10} in Tehran, Iran. Chan and Jian (2013) used NN to identify key factors (meteorological, traffic, etc.) that affected air pollution levels in Hangzhou, China. Siwek and Osowski (2012) applied wavelet transform and NN ensemble averaging to improve accuracy of daily PM_{10} concentration predictions.

While exiting PM_{10}, prediction models have utilized these and other variables (meteorological, vehicle exhaust, etc.) as inputs; no one has incorporated fugitive dust from construction sites. Even though construction related activities are considered to be important sources of pollution, particulate sources and how they influence surrounding areas have been less quantified to date (Kumar *et al.*, 2012).

Wuhan (Fig. 1a) is the capital of Hubei Province located in central China. The Yangtze River (the third longest river in the world) meets its largest tributary, Hanshui, at Wuhan, dividing the city into three sections: Hankou, Wuchang, and Hanyang—commonly referred to as the Three Towns of Wuhan. The population of Wuhan is ~8.6 million, and its total area is ~8500 km^2. Wuhan is situated within a humid subtropical monsoon climate and is consequently subject to hot and humid summers. As well as being the political, economic, and cultural center of Hubei Province, Wuhan is one of the largest junctions of land, water, and air transportation in China. Accordingly, the city has embarked on a path of rapid urbanization. Data have been published on the air quality problem the city has been experiencing in recent years. With the growing number of construction sites, the contribution from fugitive dust (having an approximate ratio of 30%) to overall PM_{10} concentration is increasing (Zhu *et al.*, 2009; Feng *et al.*, 2011a; Yang *et al.*, 2011). Fugitive dust from construction sites has become one of the most significant sources of PM_{10} pollution in megacities in China (Chan and Yao, 2008). Figure 2 and Table 2 list some information regarding construction sites surrounding St-2 (one of six PM_{10} monitoring stations in Wuhan). Additionally, Table 3 provides the total number of days PM_{10} IAQI exceeded 100 between 2003 and 2011 in the area surrounding St-2. The figures and tables provided clearly show that the intense processes governing building construction activity cannot be ignored in pollution modeling.

The initial aim of this study was to predict PM_{10} IAQI one day in advance using meteorological and construction pollutant-related parameters taken from the previous day.

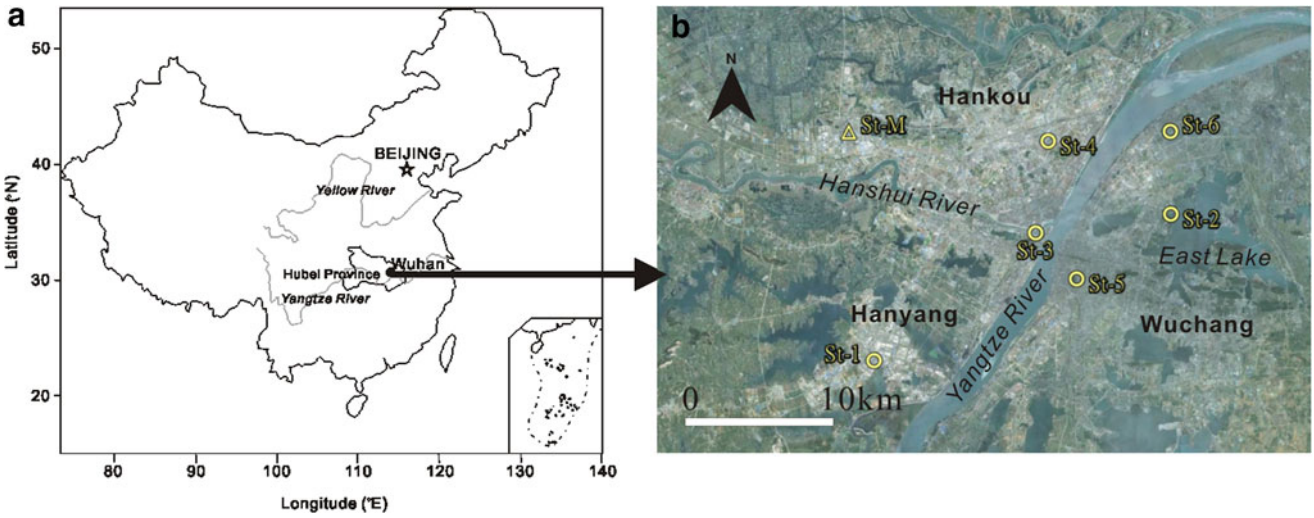

FIG. 1. Wuhan: **(a)** map and **(b)** location of stations.

Three NN-based forecasters (MLP, Elman, and support vector machine [SVM]) were used. Experimental trials were aimed to improve existing neural models (Wu *et al.*, 2011) to enhance prediction accuracy of high PM_{10} IAQI levels caused by fugitive dust derived from construction sites.

Networks were assembled using a time series recorded between January 1, 2005, and December 31, 2011, at six monitoring stations (St-1 to St-6; Fig. 1b) situated around the city of Wuhan. Model validation was carried out by comparing model prediction values to a different set of recorded data not used in model training. A cross-validation strategy was used for validation. Both existing and modified models were tested and compared for performance in achieving a one day advanced forecast of a high level IAQI event attributable to nearby construction site fugitive dust.

Data and Methodology

Data preparation

Network training was based on data taken during a 7-year period between January 1, 2005, and December 31, 2011. Daily PM_{10} IAQI data acquired at the six monitoring stations were made available by the Wuhan Environmental Protection Bureau. Meteorological variables of average daily temperature (T [°C]), relative humidity (RH, %), wind speed (Ws, m/s), barometric pressure (P [bar]), rainfall amount (RF, mm), and sunshine duration (SD, hours) were monitored at a meteorological station located within the Wuhan Meteorological Bureau. A description of monitoring data from 2005 to 2011 is provided in Table 4.

FIG. 2. Construction site distribution surrounding St-2.

TABLE 2. BRIEF DESCRIPTION OF THE CONSTRUCTION SITES SURROUNDING THE ST-2 MONITORING STATION

Construction site	*Area (m^2)*	*Distance[a] (m)*	*Construction duration*
a	38,160	1570	2009–current
b	54,760	1915	2009–current
c	48,508	1145	2009–current
d	143,250	1314	2009–current
e	37,200	615	2009–current
f	83,450	870	2009–current
g	56,287	445	2005–2007
h	61,864	1153	2007–2009
I	132,400	1010	2006–2009
j	543,800	1495	2005–current
k	118,030	1948	2007–current
l	96,500	1200	2009–2011
m	64,310	1170	2006–current
n	164,500	1990	2005–2008
o	25,510	868	2005–2006
p	221,774	1627	2007–2009

[a]Numbers in the distance column correspond to the shortest distances between construction site and St-2.

TABLE 3. NUMBER OF DAYS WHEN PM_{10} INDIVIDUAL AIR QUALITY INDEX WAS OVER 100 BETWEEN 2005 AND 2011 AS DETERMINED BY ST-2

2005	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*
27	49	59	62	43	67	49

Table 4. Description of PM_{10} Individual Air Quality Index Monitoring and Meteorological Sites Around the Study Area

Site	*Location*	*Arithmetic mean*	*Median*	*Range*
St-1	Residential suburban area	79.2	76	11–498
St-2	East Lake Park	73.7	71	10–448
St-3	Residential area in Hanyang	81.2	78	10–469
St-4	Residential area in Hankou	80.7	76	11–457
St-5	Commercial area in Wuchang	83.7	79	9–449
St-6	Heavy industrial area	86.3	82	10–478
St-M	Residential area in Hankou			

St-1 to St-6 are six PM_{10} monitoring sites; St-M is the meteorological station.

In this study, meteorological parameter input values used in model development corresponded to the actual time for which the prediction applies in the absence of available data from numerical weather forecasts.

Information related to construction area and duration was provided by the Wuhan Urban Construction Archives. Distances between construction site boundary and PM_{10} monitoring stations were measured using the ArcGIS system.

Methodology

To quantify nearby construction activity influence, this study consulted certain published literature (Watson and Chow, 2000; Muleski *et al.*, 2005; Tian *et al.*, 2008a, 2008b; Zhao *et al.*, 2009; Mensink *et al.*, 2011). The following variables, such as construction site area (A, m^2), distance between construction site boundaries and PM_{10} monitoring stations (D, m), and wind speed (Ws, m/s) were therefore, introduced.

A sigmoid was adopted to qualify the influence of one construction site when the corresponding PM_{10} monitoring station was located downwind from it:

$$Ci = \begin{cases} 0 & \text{If it is has rained in the previous or following day} \\ 0 & \text{If } D > 2000\,\text{m or } Ws \le 0.185\,\text{m/s} \\ \dfrac{A \times (114.6Ws^3 - 393.4Ws^2 + 538.8Ws - 87.03)}{D^2} & \text{otherwise,} \end{cases} \quad (1)$$

where A (m^2) is the area of the construction site; D (m) is the distance between construction site boundary and a specific monitoring station situated downwind from it; and Ws is wind speed (m/s). Ci is the Construction Influence Index of construction site i relating to a specific monitoring station.

When more than one construction site was situated upwind from a monitoring station, a sigmoid function was applied to the Construction Influence Index as follows:

$$Ci = \begin{cases} 0 & \text{If } Ws \text{ is} \le 0.185\,\text{m/s or it is has rained in the previous or following day} \\ 0 & \text{If a construction site is not upwind from a monitoring station} \\ \sum_{i=1}^{n} Ci & \text{otherwise,} \end{cases} \quad (2)$$

where Ci is the Construction Influence Index of n construction sites to a specific monitoring station.

Neural type networks for prediction

The aim of this study was to improve ANN prediction accuracy by introducing Ci. Three classical types of NNs were chosen since they individually represent independent approaches to the paradigm. MLP, one of the best known of these networks, applies the sigmoidal activation function (Hornik *et al.*, 1989). SVM is a universal solution that applies kernel principle analysis with a sophisticated, robust statistical learning algorithm. Both MLP and SVM use the feedforward structure of signal processing. The Elman network has a feedback structure (Elman, 1990) and has proven to perform well when modeling complex processes related to pollution prediction (Brunelli *et al.*, 2007). All three networks have demonstrated good performance when modeling complex processes related to air pollution formation (Brunelli *et al.*, 2007; Osowski and Garanty, 2007; Paschalidou *et al.*, 2011).

Accurate predictions are difficult due to high variability. A solution is to decompose the predicted time series into terms of lower variability. Since the wavelet application in time series analysis and prediction has been applied successfully in the past (Osowski and Garanty, 2007; Siwek *et al.*, 2009; Feng *et al.*, 2011b), wavelet decomposition of the original PM_{10} IAQI time series was used for this study. Detailed methodology regarding wavelet decomposition of the original signals has been previously described by Osowski and Garanty (2007). Figure 3 illustrates results of exemplary five level wavelet decomposition of real data related to PM_{10} IAQI from St-2 in 2005 (the upper curve) obtained by applying Daubechies (db4) wavelets implemented on the MATLAB platform. All signals (the first five levels of wavelet coefficients from D1 to D5 and the coarse approximation A5 on the fifth level) are illustrated in their original resolutions.

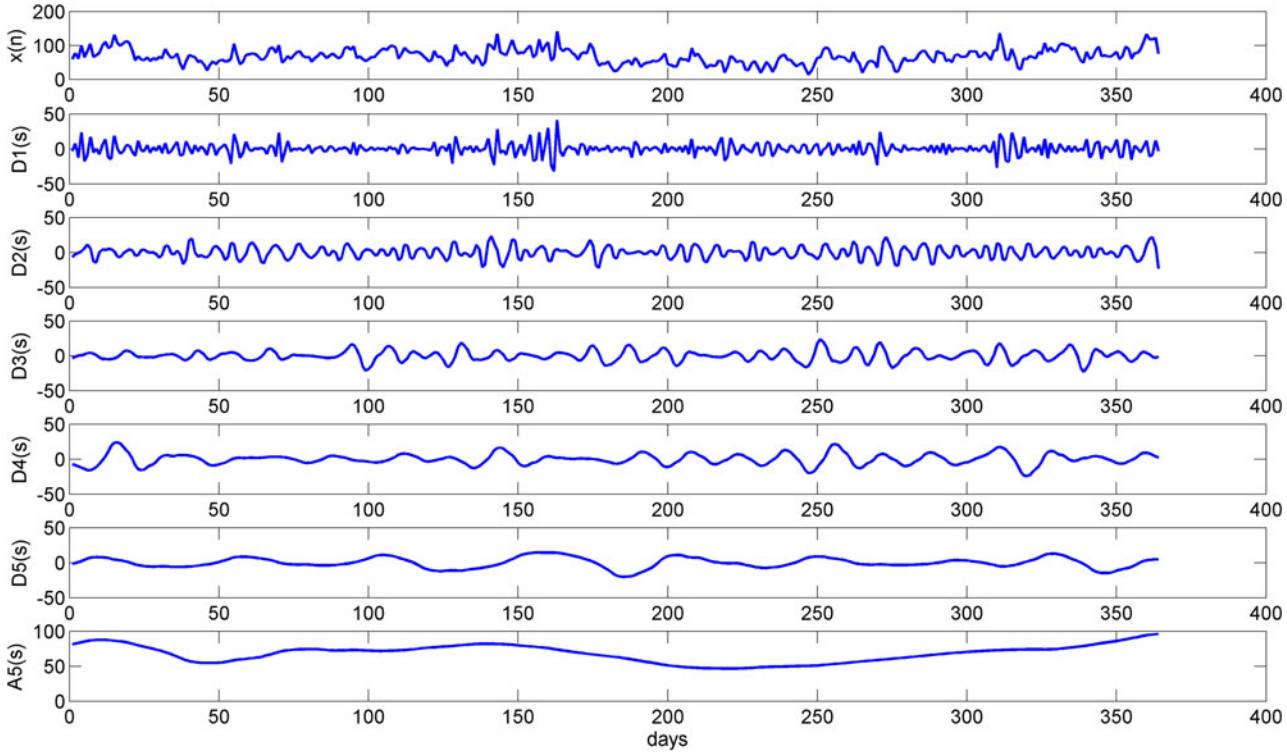

FIG. 3. Wavelet decomposition of the measured time series $x(n)$ of PM_{10} IAQI from St-2 in 2005; D1–D5 represent the detailed coefficients and A5 the coarse approximation of $x(n)$ on the fifth level. PM_{10}, particulate matter with a diameter $<10\,\mu$m; IAQI, Individual Air Quality Index.

Experiment

To evaluate the effectiveness of *Ci*, three types of NNs (MLP, Elman, and SVM) were applied separately. Wavelet coefficient prediction on each level required the use of one specific network. An additional network was needed to predict a coarse approximation of the data. Since five levels of wavelet coefficients were chosen along with A5 for coarse approximation on the fifth level, six networks were used altogether.

Each mode input pattern for each station contained a set of daily values for prediction of *Di* ($i=1,2,\ldots5$) or A5 from a specific station. One value was applied to both the current and subsequent day and where the final value was the specific PM_{10} IAQI monitoring station under consideration. Therefore, each neural mode input pattern had a total of 15 values: average temperature (*T*), relative humidity (*RH*), wind speed (*WS*), barometric pressure (*P*), rainfall amount (*RF*), sunshine duration (*SD*) from St-M, the Construction Influence Index of *n* construction sites (*Ci*), and the PM_{10} IAQI *Di* ($i=1,2,\ldots5$) or A5 from specific stations (St-1 to St-6) as illustrated in Fig. 4.

On the basis of these predicted coefficients, the real prediction of PM_{10} IAQI from specific stations for the following day is made by simply adding them together as reported in published literature (Osowski and Garanty, 2007). Equation (3) shows the recovery process of the original PM_{10} IAQI signal:

$$PM_{10}IAQI = D_1 + D_2 + D_3 + D_4 + D_5 + A_5 \tag{3}$$

The data set used to build the NN database constituted daily values related to a period between January 1, 2005, and

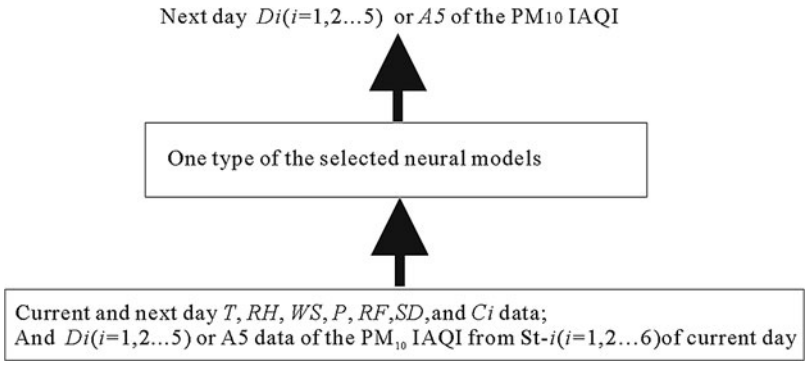

FIG. 4. Neural network architecture for St-*i* ($i=1, 2,\ldots6$). *T* (average daily temperature); *RH* (relative humidity); *WS* (wind speed); *P* (barometric pressure); *RF* (rainfall amount); *SD* (sunshine duration); *Di* (wavelet coefficients from level *i*); A5 (the coarse approximation on the fifth level); *Ci* (Influence Index of construction site); St (station).

Table 5. Training and Test Sets Used for Cross Validation

	2005			2006			…	2010			2011		
	Jan.–Apr.	May–Aug.	Sept.–Dec.	Jan.–Apr.	May–Aug.	Sept.–Dec.	…	Jan.–Apr.	May–Aug.	Sept.–Dec.	Jan.–Apr.	May–Aug.	Sept.–Dec.
Set 1	*	*	*	*	*	*	…	*	*	*		*	*
Set 2	*	*	*	*	*	*	…	*	*	*	*		*
Set 3	*	*	*	*	*	*	…	*	*	*	*	*	

Data from four successive months were cyclically used as test sets.
*Dates marked by asterisks were used for network training.

December 31, 2011. Neural model performance was evaluated by applying a cross-validation strategy by which to test the effectiveness of the tested model for prediction accuracy. The entire data set between January 1, 2005, and December 31, 2010, was used as a training set, while the 2011 data set was shared between the three subsets, using two out of the three subsets to complete the training set. The remaining subset was applied as a test set. Accordingly, three different training and test sets were used to guarantee robust performance, and test set selection independency attributed for all models that were developed and tuned. The different training and test sets used are provided in Table 5.

Data were preprocessed to eliminate instrumental errors. This was accomplished by replacing holes in the established time series with values before or after a hole occurred. In addition, each value in the NN was normalized within the specified range [0, 1], using the following linear transformation:

$$X' = (X - V_{\min})/(V_{\max} - V_{\min}), \quad (4)$$

where X' is the new normalized value; X is the old value; $V_{\max}$ is the maximum of the data set under consideration; and $V_{\min}$ is the minimum of the data set under consideration. The normalized value set was used as the NN input.

For experiments pertaining to nonlinear models of prediction, the same structures were used for predicting pollution and wavelet coefficients. Developed nonlinear network structures were as follows: 15-15-1 for MLP and 15-24-24-1 for Elman. They were established after a series of additional introductory trials. Gaussian kernel numbers of the SVM network were automatically adjusted by the learning procedures applied (Osowski and Garanty, 2007), which was different for each experiment.

Results and discussion

Trials were carried out with and without the *Ci* input to promote training and optimization, as well as to evaluate the forecasting task for daily PM_{10} IAQI. Accordingly, the training set, given the previous description, comprised of a value of 80 months, while the test set comprised of a value of 4 months.

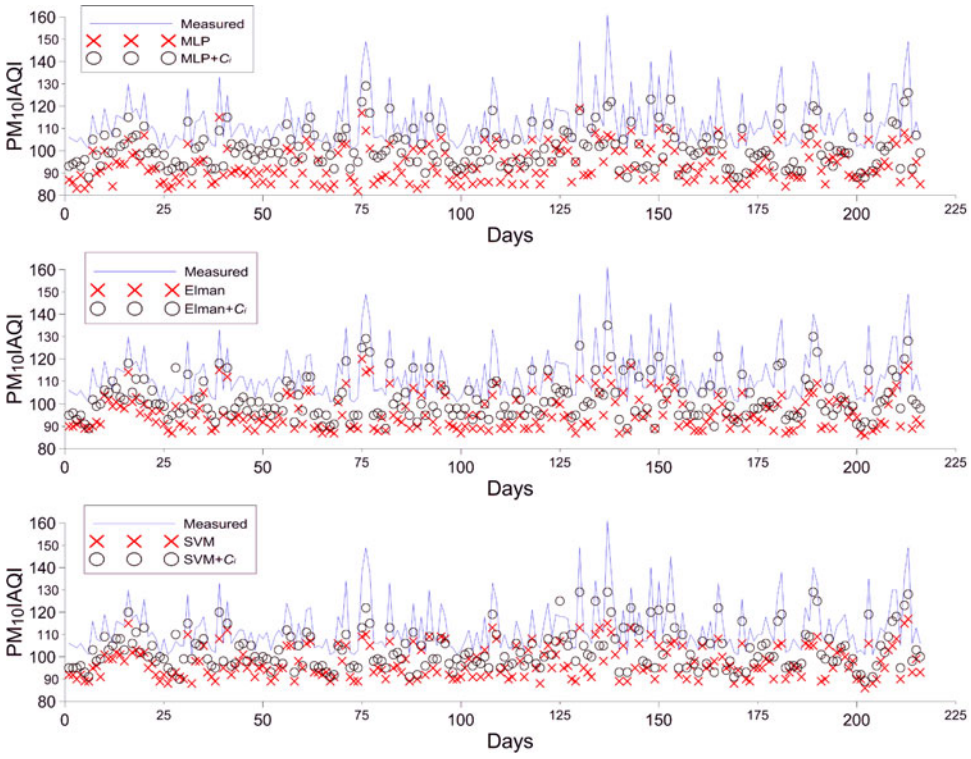

FIG. 5. Prediction of high level PM_{10} IAQI attributable to nearby construction site activity by applying three neural networks with and without *Ci* input.

Table 6. Static Index Applied to Models With and Without Construction Influence Index Input for High PM_{10} Individual Air Quality Index Event Predictions Attributable to Construction Fugitive Dust from Construction Sites

	MLP	*MLP + Ci*	*Elman*	*Elman + Ci*	*SVM*	*SVM + Ci*
r	0.88	0.91	0.89	0.93	0.89	0.92
MAE	19.39	11.99	16.95	10.95	15.53	10.20
RMSE	20.23	12.93	17.85	11.63	16.65	11.03
Mean value	93.70	101.19	96.14	102.26	97.56	103.08
MAPE (%)	16.96	10.43	14.79	9.56	13.48	8.86
n_{100}	10	99	13	132	18	161
AHPC (%)	4.63	43.98	6.01	60.82	8.33	74.53

The mean value of measured PM_{10} IAQI is 113.09.

MLP, multilayer perceptron; Ci, Construction Influence Index; SVM, support vector machine; *r*, correlation coefficient; *MAE*, mean absolute error; *RMSE*, root mean square error; *MAPE*, mean absolute percentage error; *AHPC*, accuracy rate of high PM_{10} IAQI from nearby constrcution.

The aim of experimental trials was to establish optimized architecture for each model. Model performance was evaluated using the following parameters: the correlation coefficient (*r*), mean absolute error (*MAE*), root mean square error (*RMSE*), and mean absolute percentage error (*MAPE*).

Model performance evaluations were extended to include the prediction of high PM_{10} IAQI attributable to nearby construction activity. This task is of particular importance to administrators since successfully predicting high PM_{10} values in a timely manner provides the time to restrict and confine activities that put the health and welfare of local residents at risk.

The prediction accuracy rate of high PM_{10} IAQI attributable to nearby construction activity (*AHPC*) was introduced in this study to evaluate the six models investigated.

$$AHPC = n_{>100}/N_{>100} \times 100\%, \quad (5)$$

where *AHPC* is the accuracy rate of high PM_{10} IAQI attributable to nearby construction activity. $n_{>100}$ is the total predicted number of records from the six stations wherein PM_{10} IAQI values exceeded 100 (attributable to nearby construction activity), while $N_{>100}$ (with a value of 216 for this study) is the total number of records from the six stations wherein the PM_{10} IAQI value exceeded 100 (attributable to nearby construction activity) in 2011. Prediction accuracy was identified when the relative error of the prediction value and the record was $<10\%$. High PM_{10} pollution caused by remote sources, such as dust storms was excluded from this study.

Figure 5 provides a comparison between forecasting performance of the models with and without the *Ci* input as they relate to high PM_{10} IAQI attributable to construction activity.

When taking into account Fig. 5 and Table 6, which compare predicted and observed values, correlation coefficient (*r*) values and *AHPC* were higher for the models using the *Ci* input than the original models. Improved models also outperformed original models in other parameters.

Conclusion

The aim of this study was to improve the accuracy of neural models in forecasting high PM_{10} air pollutant values in rapidly urbanizing cities. The approach used was essentially to define a warning system as it relates to information regarding PM_{10} pollution to provide local residents the capacity to choose whether to reduce unnecessary risks during outbreaks of severe pollution. The authors of this study built a predictive model using three classical neural models (MLP, Elman, and SVM) and wavelet application. While most studies have exclusively focused on the use of meteorological variables, this study also considered construction pollutants in predicting high PM_{10} one day in advance of an outbreak. Prediction tasks were related to daily PM_{10} IAQI forecasting. Five statistical indicators (*r*, *MAE*, *RMSE*, *MAPE*, and *AHPC*) were utilized to estimate output results. Improved models outperformed the original models when carrying out forecasting tasks related to high PM_{10} IAQI attributable to nearby construction activity. The benefit of the improved models is their potential in predicting PM_{10} IAQI parameters within rapidly urbanizing cities, making this forecaster an effective tool supporting other systems designed for high PM_{10} pollution management. Air pollution is complex in the Wuhan urban area as it is elsewhere. Therefore, it is necessary to add other key particle source indexes (such as vehicle exhaust) in the ongoing development of prediction models to achieve more accuracy in forecasting tasks as it pertains to local urban areas.

Acknowledgments

This study was jointly supported by one program from Hubei province and three programs from National Natural Science Foundation of China (grant No. 41301098, grant No. 41271125, grant No. 51109195) PM_{10} IAQI data used in this study were provided by the Wuhan Environmental Protection Bureau. Meteorological data were provided by the Wuhan Meteorological Station. Information related to construction area and construction project duration was provided by the Wuhan Urban Construction Archives. The authors would like to express their gratitude to the anonymous reviewers for their constructive and insightful comments.

Author Disclosure Statement

No competing financial interests exist.

References

Ashbaugh, L.L., Carvacho, O.F., Brown, M.S., Chow, J.C., Watson, J.G., and Magliano, K.C. (2003). Soil sample collection and analysis for the Fugitive Dust Characterization Study. *Atmos. Environ.* 37, 1163.

Brunelli, U., Piazza, V., Pignato, L., Sorbello, F., and Vitabile, S. (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.

Cai, M., Yin, Y.F., and Xie, M. (2009). Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. *Transport. Res. D* 14, 32.

Chan, C.K., and Yao, X.H. (2008). Air pollution in mega cities in China. *Atmos. Environ.* 42, 1.

Chan, K.Y., and Jian, L. (2013). Identification of significant factors for air pollution levels using a neural network based knowledge discovery system. *Neurocomputing* 99, 564.

Dorevitch, S., Demirtas, H., Perksy, V.W., Erdal, S., Conroy, L., Schoonover, T., and Scheff, P.A. (2006). Demolition of high-rise public housing increases particulate matter air pollution in communities of high-risk asthmatics. *J. Air Waste Manage. Assoc.* 56, 1022.

Elman, J.L. (1990). Finding structure in time. *Cogn. Sci.* 14, 179.

Feng, Q., Wu, S.J., Du, Y., Li, X.D., Ling, F., Xue, H.P., and Cai, S.M., (2011a). Variations of PM_{10} concentrations in Wuhan, China. *Environ. Monit. Assess.* 176, 259.

Feng, Q., Wu, S.J., Du, Y., Li, X.D., Xue, H.P., and Cai, S.M. (2011b). The application of wavelet analysis in the API of PM_{10} time series of Wuhan urban. *J. Huazhong Normal Univ. (Nat. Sci.)* 44, 678 (In Chinese).

Grivas, G., and Chaloulakou, A. (2006). Artificial neural network models for prediction of PM_{10} hourly concentrations, in the Greater Area of Athens, Greece. *Atmos. Environ.* 40, 1216.

Ho, K.F., Lee, S.C., Chow, J.C., and Watson, J.G. (2003). Characterization of PM_{10} and $PM_{2.5}$ source profiles for fugitive dust in Hong Kong. *Atmos. Environ.* 37, 1023.

Hooyberghs, J., Mensink, C., Dumont, G., Fierens, F., and Brasserur, O. (2005). A neural network forecast for daily average PM_{10} concentrations in Belgium. *Atmos. Environ.* 39, 3279.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Netw.* 2, 359.

Jiang, D.H., Zhang, Y., Hu, X., Zeng, Y., Tan, J.G., and Shao, D.M. (2004). Progress in developing an ANN model for air pollution index forecast. *Atmos. Environ.* 38, 7055.

Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R., and Cawley, G. (2003). Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmos. Environ.* 37, 4549.

Kumar, P., Mulheron, M., Fisher, B., and Harrison, R.M. (2012). New directions: Airborne ultrafine particle dust from building activities—a source in need of quantification. *Atmos. Environ.* 56, 262.

Mensink, C., Cosemans, G., Bleux, N., Berghmans, P., Deutsch, F., Janssen, L., Liekens, I., Torfs, R., and Van Rompaey, H. (2011). Quantification of diffuse and fugitive PM_{10} sources by integrated "hot spot" method. *Atmos. Environ.* 45, 2233.

Ministry of Environmental Protection of the People's Republic of China. (2009–2011). State of the Environment in China. http://jcs.mep.gov.cn/hjzl/zkgb

Ministry of Environmental Protection of the People's Republic of China. (2012). Technical Regulation on Ambient Air Quality Index (on trial). (HJ 633–2012).

Morabito, F.C., and Versaci, M. (2003). Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data. *Neural Netw.* 16, 493.

Muleski, G.E., Cowherd, C. Jr., and Kinsey, J.S. (2005). Particulate emissions from construction activities. *J. Air Waste Manage. Assoc.* 55, 772.

Nejadkoorki, F., and Baroutian, S. (2012). Forecasting extreme PM_{10} concentrations using artificial neural networks. *Int. J. Environ. Res.* 6, 1735.

Osowski, S., and Garanty, K. (2007). Forecasting of the daily meteorological pollution using wavelets and support vector machine. *Eng. Appl. Artif. Intel.* 20, 745.

Paschalidou, A.K., Karaloysios, S., Kleanthous, S., and Kassomenos, P.A. (2011). Forecasting hourly PM_{10} concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. *Environ. Sci. Pollut. R.* 18, 316.

Perez, P., and Reyes, J. (2006). An integrated neural network model for PM_{10} forecasting. *Atmos. Environ.* 40, 2845.

Siwek, K., Osowski, S., and Sowinski, M. (2009). Neural predictor ensemble for accurate forecasting of PM_{10} pollution. *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Barcelona. p. 1.

Siwek, K., and Osowski, S. (2012). Improving the accuracy of prediction of PM_{10} pollution by the wavelet transformation and an ensemble of neural predictors. *Eng. Appl. Artif. Intel.* 25, 1246.

Tian, G., Fan, S.B., Huang, Y.H., Nie, L., and Li, G. (2008a). Relationship between wind velocity and PM_{10} concentration & emission flux of fugitive dust source. *Environ. Sci.* 10, 2983. (In Chinese)

Tian, G., Li, G., Yan, B.L., Huang, Y.H., and Qin, J.P. (2008b). Spatial dispersion laws of fugitive dust from construction sites. *Environ. Sci.* 1, 259. (In Chinese)

Watson, J.G., and Chow, J.C. (2000). Reconciling urban fugitive dust emissions inventory and ambient source contribution estimates: summary of current knowledge and needed research. Desert Research Institute, DRI Document No. 6110.4F.

Wu, S.J., Feng, Q., Du, Y., and Li, X.D. (2011). Artificial neural network models for daily PM_{10} air pollution index prediction in the urban area of Wuhan, China. *Environ. Eng. Sci.* 28, 357.

Yang, T., Zeng, Q.L., Liu, Z.F., and Liu, Q.S. (2011). Magnetic properties of the road dusts from two parks in Wuhan city, China: implications for mapping urban environment. *Environ. Monit. Assess.* 177, 637.

Zhao, P.S., Feng, Y.G., Zhang, Y.F., Zhu, T., Jin, J., and Zhang, X.L. (2009). Modeling and impact study of fugitive dust emissions from building construction sites. *China Environmental Science* 6, 567. (In Chinese)

Zhu, Z.C., Kong, L.L., and Xia, K. (2009). Analysis of PM_{10} source in Wuhan and its countermeasures [J]. *Environ. Sci. Technol.* 9, 64. (In Chinese)