©2016 Published in 4th International Symposium on Innovative Technologies in Engineering and Science 3-5 November 2016 (ISITES2016 Alanya/Antalya - Turkey)

Prediction of Air Pollutant Levels by Using Artificial Neural Networks and Statistical Methods

[1]Fatih Taşpınar, [2]Memiş Işık and [3]Kamil Kaygusuz
[*1] Faculty of Engineering, Environmental Engineering Dept., Düzce University, Konuralp Campus, Düzce, Turkey
[2]Faculty of Engineering, Mechanical Engineering Department, Karabük University, Karabük, Turkey
[3]Physical Chemistry, Department of Chemistry, Karadeniz Technical University, Trabzon, Turkey

Abstract:

One of the environmental problems adversely affecting human health and welfare in many part of world is air pollution. Pollution monitoring data can be utilized to predict concentrations of air pollutants for short-term using artificial intelligence approaches and multivariate regression analysis. In this paper, artificial neural networks (ANNs) and multivariate regression modeling (MRM) techniques have been comparatively employed to forecast one-hour ahead concentration of particulate air pollution (PM_{10}). An hourly based data was composed by including meteorological factors and particulate concentrations for the years 2015-2016. ANN(12-7-1) model with R^2 of 0.887 and RMSE of 19.89 yielded fairly rational predictions over hourly dataset while the best MRM models produced lower scores with R^2 of 0.848 and RMSE of 21.33 in general. ANN models simulating the time series data better among the models identified have been chosen in further tuning in the prediction of next hour's concentration of PM_{10}.

Key words: Air pollution, Modeling, Artificial Neural Networks, Regression Analysis.

1. Introduction

Particulate matter (PM) air pollution is caused by a mixture of organic and inorganic particles which are solid and liquid phase spreading out from variable sources [1, 2]. The particles with an aerodynamic diameter less than or equal to 10 µm, namely PM_{10}, are emitted to the atmosphere mainly from the fuel combustion [3-5]. The highest particulate pollution levels are associated to stable meteorological conditions with thermal inversion in urban and industrial areas. Epidemiological studies indicated a close relationship between atmospheric particulate concentrations and increased mortality and morbidity [6, 7]. High levels of these pollutants can be harmful for goods, and also decrease visibility. The air quality standards are thus set for PM_{10}, declaring hourly, daily and annual limits. According to EU standards for PM_{10}, the annual average limit value of 40 $\mu g.m^{-3}$ and 24-h limit value is declared as 50 $\mu g.m^{-3}$, and also the limit values should not be exceeded by the specified number of times in a year [8].

Elevated levels of air pollutants in the air may cause acute or chronic health effects, and even cause premature deaths in the elderly people. The air quality forecasting studies is an important research topic in air pollution science for public health. Many functional alert systems were employed by utilizing statistical and hybrid models, to take precautions before and during air pollution episodes. In this context, long-term or short-term air pollution forecasting models have

*Corresponding author
Address: Faculty of Engineering, Environmental Engineering Department., Düzce University, Konuralp Campus, Düzce, Turkey 81620, Düzce/TURKEY. E-mail address: fatihtaspinar@duzce.edu.tr, Phone: +90380 542 10 36

been utilized as an aid for air quality management. Artificial neural networks (ANNs), multiple regression models (MRM), time series models and mixed models are mostly preferred approaches in air quality forecasting researches [9-11]. With nonlinear simulation and learning abilities, ANNs, are powerful tools for regression and pattern recognition problems. A real-life problem such as short-term air pollution prediction, covering complex nonlinear relations with meteorological factors, can be handled by ANN models very well. ANNs consist of neurons that are interrelated connections artificial processing units and they can process information by error minimization within a finite computation loop. ANNs can thus be trained to learn a complex relationship between two or more variables recorded in training datasets. Among the available ANNs, the feedforward error backpropagation neural networks are the most employed ANN types, of which inputs has a nonlinear transfer function. By this means, they have been used in many successful studies in local air pollution modelling for forecasting pollutants NO_2, O_3, SO_2, CO and PM_{10} [12, 13].

Time series modeling approaches for short-term air pollution prediction phenomena are also employed, of which results are comparable to other artificial intelligence methods. They mostly applied on continues time series datasets. These datasets include some degree of randomness, for example, random changes in meteorological parameters due to atmospheric events during diurnal changes and seasonal variations. Some studies have revealed that the air quality data are stochastic time series by making short-term estimations possible by exploring historical data patterns [14, 15]. The most widely employed time series models are autoregressive integrated moving average and a type of them with external parameter models (e.g. ARMA, ARIMA, ARIMAX) in time series analysis [16, 17]. In the case of conventional air pollutants non-seasonal and seasonal time series models have been successfully applied to monitored datasets that are based on mostly daily or monthly averaged values [18, 19]. Generally, the accuracy and quality of models can vary on individual experience of issue, knowledge of statistical analysis methods in the model identification stage. The visualization of time series forecasting plots leads to establish several models for the same dataset and most stable one can used in tests further.

In the present study, one-hour ahead concentration prediction of PM_{10} using statistical multiple regression and ANN based models were employed on the hourly dataset for the period of 2015-2016. Well tuned models were then applied in short-term predictions of PM_{10} air pollution levels to identify a model best explains the variance in data with reduced inputs.

2. Materials and Methods

2.1 Dataset used and explanatory statistics

An hourly dataset was composed for Düzce province in Turkey for the period of 2015-2016. The dataset contained information about local meteorological parameters such as air temperature (AT, °C), wind speed (WS, m/s), relative humidity (RH, %) and mass concentration of particulate matter (PM_{10}, $\mu g/m^3$). The meteorological data was taken from the General Directorate of Meteorological Affairs of Turkey and PM_{10} data was taken from the Ministry of Environment and Urban Planning, using the online web service of the National Air Quality Monitoring

Network of Turkey. Table 1 gives the descriptive statistics of these variables and Fig. 1 visualizes the hourly time series plot for PM_{10} over air temperature.

Table 1. Descriptive statistics of hourly dataset (2015-2016) used for investigation.

	Cases (N)	Mean	Median	Mode	Min.	Max.	25% Perc.	75% Perc.	Range	Std.Dev.
PM_{10}	8881	98.41	60.00	37	0.00	891	39	104	891	112.81
AT	8926	16.02	17.00	22	-13.00	42	8	23	55	9.82
WS	8926	0.62	1.00	1	0.00	1	0	1	1	0.48
RH	8926	79.95	88.00	103	12.00	103	63	100	91	22.78

In the hourly dataset, one step forward-lagged set of these variables were constructed for including the prior data from one-hour before. The peak levels of PM_{10} can be seen during winter due to residential heating by fossil fuels such as coal, lignite and wood, particularly at least five months from October to March in contrast to the levels observed during the summer periods. PM_{10} and temperature values were ranged in [0-891] µg/m^3 and [–13-42] ^{0}C, respectively. The mean and 75% percentile of PM_{10} level were 98.41±112.81 and 104 µg/m^3, respectively, however, which is higher than the acceptable limit of 90 µg/m^3 declared in National Air Quality Standard of Turkey. The statistics showed that the atmosphere over Düzce is highly polluted by particulate matter and the pollution episodes particularly during winter periods can affect human health adversely. Therefore, air pollution forecasting models can serve a tool in identifying emergency periods and short-term pollutant levels.

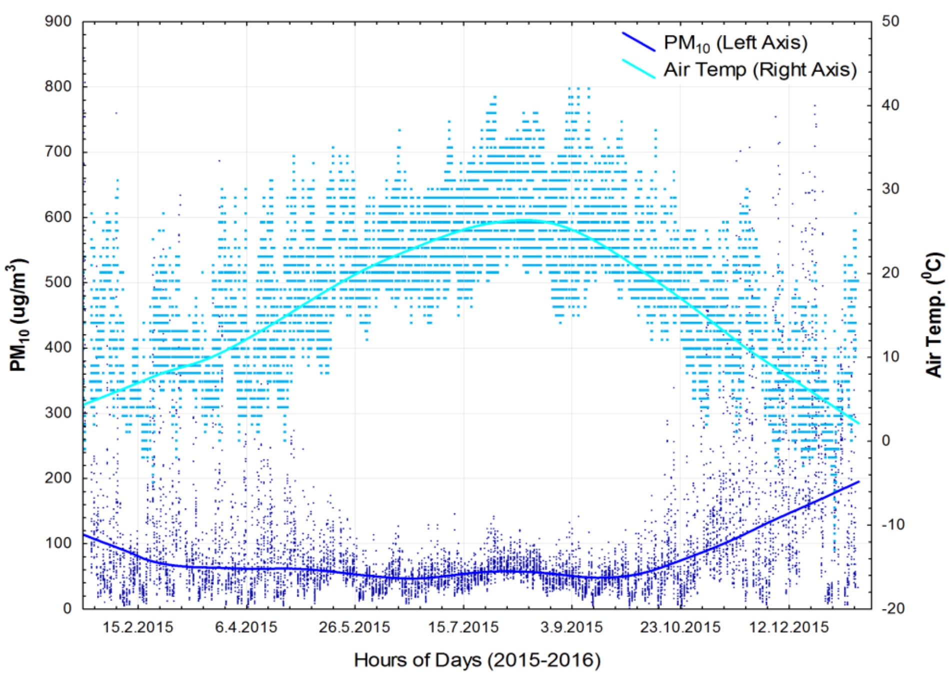

Figure 1. Hourly time series plots of PM_{10} and air temperature for 2015-2016 period.

2.2. Multivariate regression models

Regression is a statistical tool and one of the most frequently applied technique for the investigation of relationships between variables. It is valuable for quantifying the impact of various simultaneous influences upon a single dependent variable [20]. With *n* independent observations, one can write a model for each dependent variable by organizing independent variables into vectors and matrices as follows.

$$Y = Z.\beta + \varepsilon \tag{1}$$

where Y is dependent variable, Z is independent matrices, β is coefficient matrices and ε is residual matrices. The equation may be an empirical model (simply descriptive) or a mechanistic model. A response variable or dependent variable (Y) has been measured at several settings of one or more independent variables (X), also called input variables, regressors or predictor variables. Sometimes, regression is called curve fitting or parameter estimation. A first-order multiple-linear regression model can be written as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{2}$$

where β_0 is the intercept, β_i's are the model coefficients or variable weights, X_i's are dependent variables. In order to investigate the predominant factors affecting PM_{10} levels using meteorological factors, empirical models were also developed incorporating the SWR method. In this method, the best-fitted combination of independent variables for the prediction of dependent variable can be selected with forward-adding and backward deleting variables [21].

As a part of the study, we used MRM and stepwise regression (SWR) techniques to develop linear models in which PM_{10} was dependent variable whereas all lagged-PM_{10} and meteorological variables were explanatory variables. These multivariate linear regression models, in this case, are considered as base models for comparison with the other tuned models.

2.3. ANN modeling

The artificial neural networks are adaptive nonlinear systems capable to approximate any function. ANNs are used in regression and classification studies in general, in which the inspired model that does not have a clear relationship between its inputs and outputs [22]. ANNs are built on a network of simple processing elements, namely neurons that exhibit complex global behavior determined by the connections between the processing elements and element parameters. ANNs are made up of a number of layers with neurons. The ANN neurons are located in input, hidden and output layers, which is thus called as multi-layer perceptron (MLP) ANN in general. Fig. 2 shows a typical three-layered ANN, which is also used in this study.

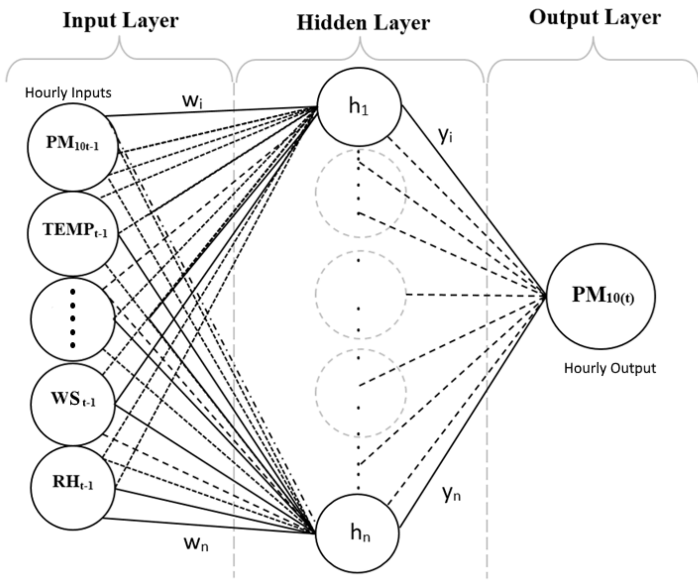

Figure 2. General structure and inputs of ANN model used in hourly PM_{10} modeling

The first layer corresponds to the input variables to the problem with one node for each input variable. The second layer used to capture nonlinear relationships among variables by interconnections. The third layer provides the predicted values. All weights are usually initialized with random values drawn from a standard normal distribution. During an iterative training process, ANN calculates an output o(x) for given inputs and current weights. If the training process is not yet completed, the predicted output (o) will differ from the input (y). An error function, like the root mean squared error (RMSE) which measures the difference between predicted and observed output. Finally, the process stops if a pre-specified criterion is fulfilled such as checking early stopping conditions by calculating global error. A single neuron processes multiple inputs applying an activation function on a linear combination of the inputs as follows:

$$y_i = f\left(\sum_{q=1}^{l} w_{iq} . f\left(\sum_{j=1}^{m} (v_{qj} x_i + b_j)\right) + b_q\right) \tag{3}$$

where x_j is the set of inputs, w_{iq} and v_{qj} ate the synaptic weights connecting the q_{th} input to the j_{th} neuron, b is bias term, f is the activation or transfer function, and y_i is the output of the i_{th} neuron. Weights are the knowledge base of the ANN system, which represents the non-linear properties of the neuron by its activation function. The activation function is usually non-linear, with a sigmoid shape such as logistic function as follows:

$$sig(x) = \frac{1}{1+e^{-X}} \tag{4}$$

Generally, feedforward MLP networks are trained using error back propagation (BP) algorithm [23], which covers heuristic and numerical optimization algorithms. Heuristic techniques include gradient descent and the resilient algorithm [24]. Some ANN parameters such as learning rate, learning momentum and hidden layer neuron count etc. have been determined before training stage and ANN model should then be executed. The inputs to the ANN models also have to be selected appropriately to better simulate the problem under consideration. Later, these parameters were determined by testing several ANN models on the same dataset.

2.4. Data pre-processing and performance measures

Time series dataset including the variables PM_{10}, AT, WS and RH were pre-processed prior to use in the models. Firstly, it’s applied to a list-wise local linear regression to fill the missing values up to six cells by columns, but, the bigger missing areas were remained blank. Thus, the average valid data was about 90% of the entire dataset. The blank inputs can be skipped when training which is a powerful attribute of ANN models.

Additionally, in order to make input variables intercomparable before executing on the modelling framework, the variables were normalized in the range of 0.05-0.95 using min-max normalization given in Eq. (7) as follows:

$$y^{'} = 0.05 + \frac{(y - y_{\min})}{(y_{\max} - y_{\min})} * 0.95 \tag{5}$$

where $y^{'}$ is the normalized value, y_{min} is minimum value, y_{max} is maximum value and y is the actual value.

All hourly variables at hour *t* can be represented as $PM10_t$, AT_t, WS_t and RH_t for simplicity. Here, in order to make input vectors utilizing from the prior hours to predict actual $PM10_t$ concentration a time *t*, all these variables were 3 hours lagged by back-shifting. $PM10_t$ shows the actual PM_{10} concentration to being predicted. So, we obtained a dataset including the variables $PM10_t$ *to* $PM10_{t-3}$, AT_t *to* AT_{t-3}, WS_t *to* WS_{t-3} and RH_t *to* RH_{t-3}, which was 16 variables in total.

In order to compare model accuracy and to identify the best models among tested, coefficient of determination (R^2, 1.0 is the perfect model) indicating how well the model results fit to observed data points, RMSE (the minimum is the best) showing the overall accuracy of the actual model and Index-of-Agreement (IA) showing the overall fitting accuracy of the model (IA values close 1.0 indicates the better agreement with the selected model) is used. Equations of these indices are given as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \tag{6}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n} (O_i - \overline{O})^2 . (\mathrm{P}_i - \overline{\mathrm{P}})}{n.\sigma_O . \sigma_P} \right)^2 \qquad (7)$$

$$IA = 1 - \left(\frac{\sum_{i=1}^{n} \left(P_i - O_i \right)^2}{\sum_{i=1}^{n} \left(\left| P_i - \overline{O} \right| + \left| O_i - \overline{O} \right| \right)^2} \right) \qquad (8)$$

where *n* total number of annual measurements, P_i is predicted values, O_i is observed values, $\overline{P}$ is mean of the predicted values, $\bar{O}$ is mean of the observed values, σ_P is standard deviation of the predicted values, σ_O is standard deviation of the observed values.

3. Results and Discussions

3.1. Multivariate regression models with performance evaluations

Constructing statistical models based on multivariate linear regression analysis can be resulted in many models with different input variables. In order to select the best model with reduced inputs as initially it was included 16 variables, we thus employed step-wise regression method. SWR technique yielded three models with much lower inputs and better error measures comparing to MRMs, which was given in Table 2. Here, these models were coded as SWR-1 to SWR-3 and R^2 and RMSE values were also given. Alternatively, a simple model inputted with first-lag of all the variables coded as MRM, which is considered as a base model for comparison. For the better accuracy, all regression models included a constant term.

Table 2. The best models identified by multivariate regression analysis

Model	Predictors	Equation*	R^2	RMSE
SWR-1	$PM10_{t-1}$, $PM10_{t-2}$	$PM10_t = 7.402+1.060*PM10_{t-1}-0.136*PM10_{t-2}$	0.826	32.67
SWR-2	$PM10_{t-1}$, $PM10_{t-2}$, RH_{t-3}, AT_{t-1}	$PM10_t = 48.773+1.033*PM10_{t-1}-0.128*PM10_{t-2} - 0.351*RH_{t-3} - 0.704*AT_{t-1}$	0.831	25.51
SWR-3	$PM10_{t-1}$, $PM10_{t-2}$, RH_{t-3}, AT_{t-1}, AT_{t-3}	$PM10_t = 42.591+1.030*PM10_{t-1}-0.123*PM10_{t-2} - 0.298*RH_{t-3}-1.307*AT_{t-1}+0.713*AT_{t-3}$	*0.848*	21.33
MRM	$PM10_{t-1}$, AT_{t-1}, WS_{t-1}**, RH_{t-3},	$PM10_t = 35.618+0.921*PM10_{t-1}-0.533*AT_{t-1} - 0.234*RH_{t-3}$	0.781	38.64

*The coefficients were significant at $p<0.05$,** The coefficient of WS_{t-1} was not significant at $p<0.05$

R^2 values for SWR models were ranged in [0.826-0.848] and RMSE values were ranged in [21.33-32.67]. SWR analysis selected maximum 5 variables, that is $PM10_{t-1}$, $PM10_{t-2}$, RH_{t-3}, $AT_{t-1,}$ AT_{t-3}, for the SWR-3 model with an R^2 of 848 and RMSE of 21.33, which was the best scores obtained within this modeling method. The base model MRM was also yielded a reasonable R^2 (0.781) and RMSE (38.65). Generally, predictor variables $PM10_{t-1}$, $PM10_{t-2}$, AT_{t-1}, AT_{t-3} and RH_{t-3} were included in the models whereas no lags of WS was selected. This interesting situation was associated with the recorded values of WS that was 0 or 1 mostly. It might be a systematic error due to wind speed measurement equipment. The inputs from the WS were then discarded by SWR, because WS values did not significantly change the R^2 in regression analysis. A scatter plot for the predictions of SWR-3 model and PM_{10} data was visualized in Fig. 3. Also it shows a linear regression line that fits fairly well to the observed PM_{10} data. In order to see how this model represented actual data, time series data for the predictions of SWR-3 model and observed PM_{10} data were plotted in Fig. 4. The predicted data points by the model followed historical pattern of actual data reasonably well with reaching peak and valley points. However, it mostly tended to underfit according to time series plot. Because, hourly fluctuations in data was much bigger than that model could explain, which was the reason of increasing model residual errors. Ultimately, obtained SWR models produced acceptable accuracy scores considering their simplicity with reduced inputs.

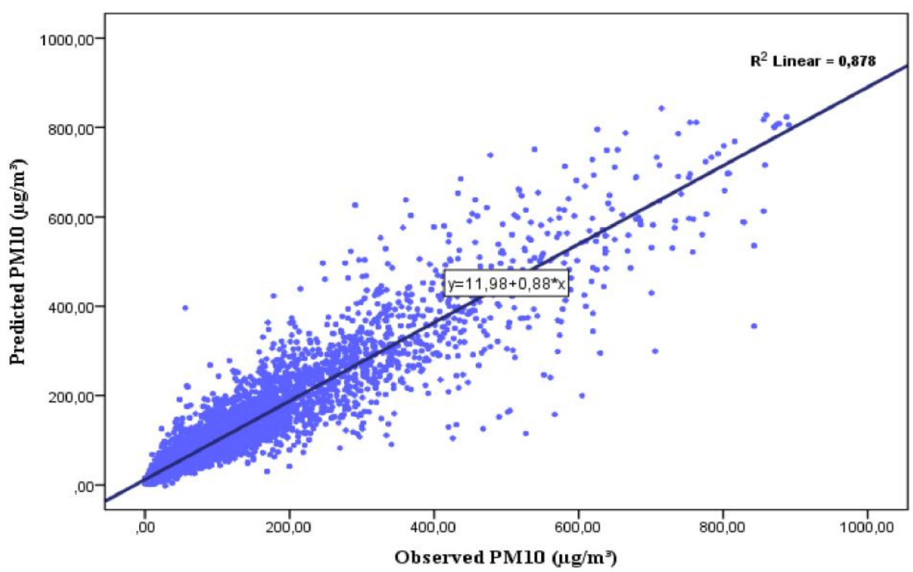

Figure 3. Scatter plot of predictions of SWR-3 model and observed PM_{10}

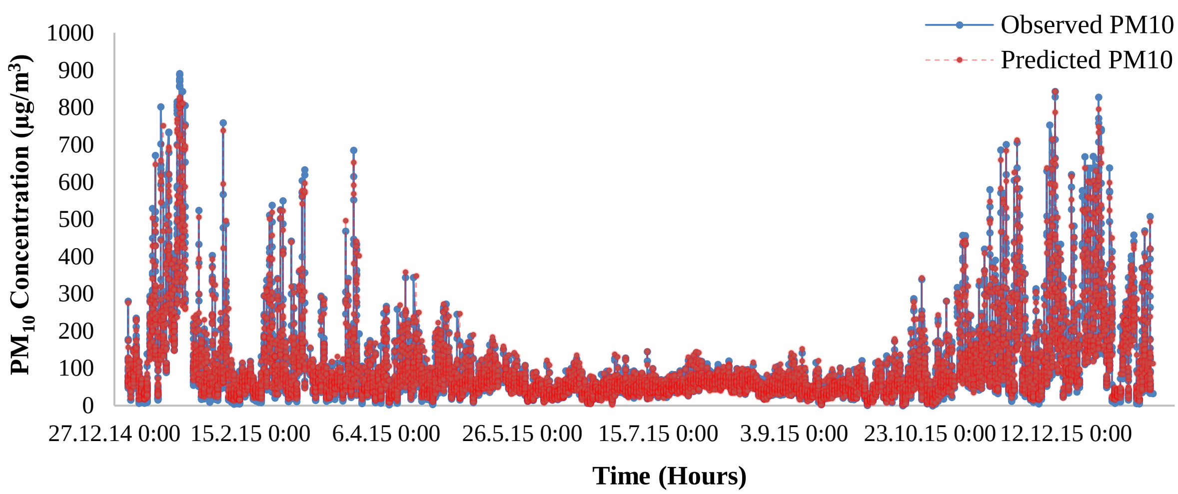

Figure 4. Time series plot of SWR-3 model predictions over actual PM_{10}

3.2. ANN models with performance evaluations

Construction of ANN models are merely complex comparing to MRMs because some internal parameters such as hidden layer neuron count and learning rate should be determined. Firstly, before the training, the entire dataset is divided into training (75%), test (15%) and validation sets (10%). Later, ANN models were designated and model parameters were set. In the present study, an open source library, Fast Artificial Neural Network (*FANN*) implemented by Nissen (2003), was utilized as ANN modeling engine in the prediction of one-hour ahead PM_{10} concentration with lagged input vectors. The input vector includes the first to the third lags of all variables as shown in Fig. 2, which can be written as $PM10_t = f_{net}(PM10_{t-1} \ldots PM10_{t-3},\ AT_{t-1} \ldots AT_{t-3},\ WS_{t-1} \ldots WS_{t-3},,\ RH_{t-1} \ldots RH_{t-3})$ model. These models were coded as ANN-1 to ANN-3. Also, we tried to construct ANN models with some selected inputs according to inputs of SWR models, which were coded as ANN-R1 and ANN-R2. Table 3 shows constructed ANN models with training parameters. The *FANN* library offers an automated training method, so-called cascading-training procedure, which provides a way to determine the final neural network structure consists of a number of hidden layers with one shortcut connected neuron in each. Ultimately, ANN models were set by utilizing cascading-training method provided by this library.

Table 3. The ANN models constructed in the prediction of $PM10_t$

Model	Input Vector	Target Learning Rate	Hidden Layer Neuron Count	R^2	IA	RMSE
ANN-1 (4-5-1)	$PM10_{t-1}, AT_{t-1}, WS_{t-1}, RH_{t-1}$	0.1	5	0.843	0.67	25.55
ANN-2 (8-6-1)	$PM10_{t-1} \ldots PM10_{t-2}, AT_{t-1} \ldots AT_{t-2}, WS_{t-1} \ldots WS_{t-2}, RH_{t-1} \ldots RH_{t-2}$	0.1	6	0.861	0.73	22.39
ANN-3 (12-7-1)	$PM10_{t-1} \ldots PM10_{t-3}, AT_{t-1} \ldots AT_{t-3}, WS_{t-1} \ldots WS_{t-3}, RH_{t-1} \ldots RH_{t-3}$	0.1	*7*	*0.887*	0.79	19.89
ANN-R1 (4-6-1)	$PM10_{t-1}$, $PM10_{t-2}$, RH_{t-3}, AT_{t-1}	0.1	*6*	*0.865*	*0.77*	*20.57*
ANN-R2 (5-6-1)	$PM10_{t-1}$, $PM10_{t-2}$, RH_{t-3}, AT_{t-1}, AT_{t-3}	0.1	6	0.866	0.77	20.31

Feed-forward backpropagation type ANN with sigmoid function for transfer functions of input layer and tanh for hidden layer were then determined in training of networks. Maximum number of epochs was set to 1000, applying an early stopping criterion to avoid over fitting or underfitting by setting the validation process at every 10 training epochs. A starting learning rate of 0.52 was gradually decreased by 1.21% to reach a learning rate of 0.1 at every epoch during the cascading-training procedure. Training results showed that R^2 values were ranged in 0.843-0.887 and RMSE were in 19.89 -25.55. The ANN-3 model with 12 inputs and 7 hidden neurons, namely ANN(12-7-1) model, yielded the best R^2 value (0.887) and the minimum RMSE (19.89). The second ANN model produced the best scores was ANN-R2 model with 5 inputs and 6 hidden neurons as in the form of ANN(5-6-1) (R^2 of 0.866 and RMSE of 20.31). Several experiments with different structures were also tried as shown in Table 3, however, the models ANN(12-7-1) and ANN(5-6-1) produced the best error measures and model accuracy according to IA values among the others. IA measures prediction errors to test fitting accuracy and quality of fit. Here, IA values were generally close to 1.0, which were 0.79 for ANN-3 and 0.77 for ANN-R2, suggesting a well agreement with the selected model. Thus, the validation of ANN model did not tend to underfitting or overfitting on average. A performance plot obtained from predicted values from ANN models was visualized in Fig. 5. The red line indicates an exact fit of R^2=1.0, dashed gray lines indicates 95% confidence band limits were also shown in Fig. 5, clearly indicating the most of the data points fall in these limits.

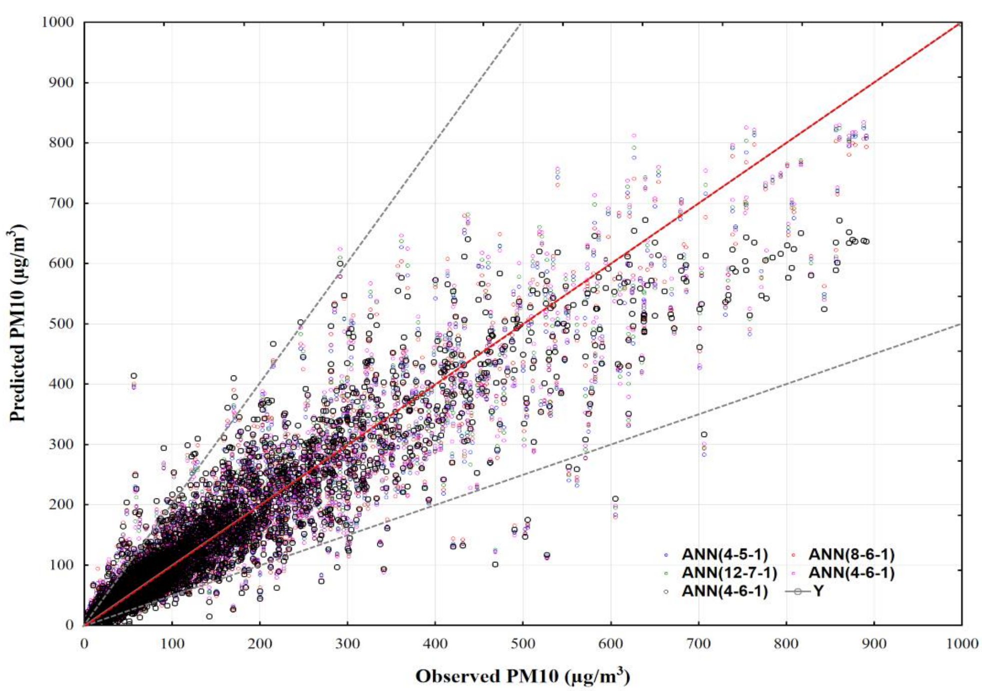

Figure 5. Scatter plot of ANN model's predictions over actual PM_{10}

Hourly time series dataset was plotted against to predictions of ANN-3 and ANN-R2 models in Fig. 6. All the data is well followed by ANN model, simulating historical pattern of hourly PM_{10}

concentration. As it can be seen in Fig. 6 that at some extreme conditions with elevated PM_{10} levels, particularly the levels higher than 550 $\mu g/m^3$ observed during strictly calm days. Due to air circulation issues occurred in Düzce province in calm nights, extreme conditions can be observed. Elevated PM_{10} levels thus mostly occurred during cold months with a rate of 2% for levels higher than 520 $\mu g/m^3$. However, the most frequent PM_{10} data within a range of $\mu \pm 3\sigma$ were predicted reasonably successfully. Extreme value problem for ANN model is a well-known issue, because ANNs cannot successfully evaluate less trained input values or less frequent data observed at extreme conditions [26, 27]. However, all the ANN models executed in the tests were very successful in the predictions of $PM10_t$ comparing to tested multivariate regression models.

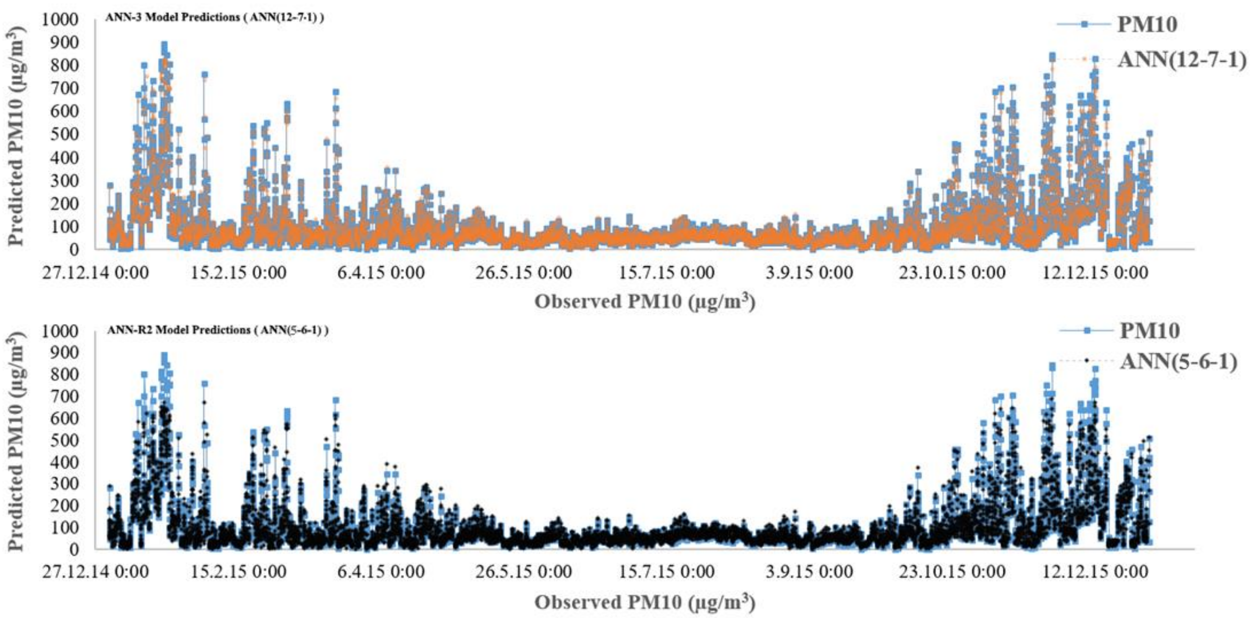

Figure 6. Hourly time series plots for the predictions of ANN-3 and ANN-R2 models over actual PM_{10} data

Conclusions

The present study investigated short-term air pollution modeling through PM_{10} based on hourly monitoring data of meteorological factors and particle concentrations. In order to obtained a reasonable model for one-hour ahead PM_{10} forecasting using a time series dataset, we have employed multivariate regression analysis technique and ANN modeling approaches. R^2 values for MRM-SWR models were varied in 0.826-0.848 and RMSE values were varied in 21.03-32.67. A regression model with 5 predictors constructed to predict $PM10_t$, including variables $PM10_{t-1}$, $PM10_{t-2}$, RH_{t-3}, AT_{t-1}, AT_{t-3}, produced acceptable results according to R^2 of 848 and RMSE of 21.33. These models were considered as base models in comparisons. Also, analysis of recorded WS data showed a systematic error, so it was discarded in regression models. In addition to regression models, ANN models for predicting PM_{10} air pollution level one hour-ahead were designed. ANN models with 4, 8, 12 and 5 inputs were constructed with different hidden neuron counts varied in 5-7. Identified ANN models produced better R^2 values varied in 0.843-0.887 and RMSE values varied in 19.89 -25.55. The best ANN model, ANN-3, was in the form of 12-7-1 with 12 inputs and 7 hidden neurons, producing R^2 of 0.887 and RMSE of 19.89. The inputs to ANN-3 were $PM10_{t-1}...PM10_{t-3}$, $AT_{t-1}...AT_{t-3}$, $WS_{t-1}...WS_{t-3}$, $RH_{t-1}...RH_{t-3}$ and the output was $PM10_t$. Generally, all the tested ANN models were reasonably successful comparing to the MRM and SWR models. Performance plot of hourly time series dataset showed that the predictions of ANN models followed the extreme data points well at peaks and valleys in the dataset. Cascading-training method of *FANN* library was utilized to determine ANN model parameters such as neuron count and learning rate. ANN model is slightly tended to underpredict mostly at extreme conditions but yielded better prediction results than regression models in general. On the other hand, ANN model did not tend to overpredict as time series plots indicated. Comparing to regression based models, ANN models were very flexible in handling of the model and execution, because ANN models obtained after a training step can be managed on any data with proper input vector. We thus concluded that a real life application for emergency alert system for hourly air pollution predictions can be realized utilizing ANN modeling. However, to properly tackle with the extreme PM_{10} pollution levels, particularly observed during pollution episodes during winter periods, some other methods such as time based hybrid models or discrete ANN models for handling elevated air pollution levels should be considered.

References

[1] WHO. Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment. WHO, 2006, Geneva.

[2] Sfetsos A, Vlachogiannis D. A new methodology development for the regulatory forecasting of PM10. Application in the Greater Athens Area, Greece. Atmospheric Environment, 2010, 44, 3159-3172.

[3] Aneja VP, Agarwal A, Roelle PA, Phillips SB, Tong Q, Watkins N et al. Measurements and analysis of criteria pollutants in New Delhi, India. Environment International, 2001, 27,35-42.

[4] Kampa M, Castanas E. Human health effects of air pollution. Envir. Pollution, 2008, 151, 362–367.

[5] Vahlsing C, Smith KR. Global review of national ambient air quality standards for PM10 and SO2 (24 h). Air Qual. Atmos. Health, 2012, 5, 393–399.

[6] Shang Y, Sun Z, Cao J, Wang X, Zhong L, Bi X, Li H, Liu W, Zhu T. and Huang W. Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. Environment International, 2013, 54, 100–111.

[7] Pope CA, Dockery DW. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc, 2006, 56, 709–742.

[8] EC, Directive 2008/50/EC of the European Parliament and of the Council on Ambient Air Quality and Cleaner Air for Europe, 2008, Brussels, Belgium.

[9] Niska H, Hiltunen T, Karppinen A, Ruuskanen, J. and Kolehmainen, M. Evolving the neural network model for forecasting air pollution time series. Eng. App. of Artificial Intelligence, 2004, 17, 159-167.

[10] Schlink U, Dorling S, Pelikan E, Nunnari G, Cawley G, Junninen H et al. A rigorous intercomparison of ground-level ozone predictions. Atmospheric Environment, 2003, 37, 3237–3253.

[11] Perez P, Trier A, Reyes J. Prediction of $PM_{2.5}$ concentrations several hours in advance

using neural networks in Santiago, Chile. Atmospheric Environment, 2000, 34, 1189–96.
[12] Kukkonen J, Partanena L, Karppinen A, Ruuskanen J, Junninen H, Kolehmainen M, et al. Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki, Atmospheric Environment, 2003, 37, 4539–50.
[13] Kurt A, Gulbagci B, Karaca F, Alagha O. An online air pollution forecasting system using neural networks, Environment International, 2008, 34, 592–98.
[14] Kao J, Huang S. Forecasts Using Neural Network versus Box-Jenkins Methodology for Ambient Air Quality Monitoring Data, J. Air Waste Mang. Assoc., 2000, 50 (2), 219-226.
[15] Horowitz J, Barakat S. Statistical analysis of the maximum concentration of an air pollutant: effects of autocorrelation and non-stationarity, Atmos. Env., 1979, 13, 811-18.
[16] Kumar A, Goyal P. Forecasting of Daily Air Quality Index in Delhi, Sci. Total Environ, 2011, 409, 5517–5523.
[17] Goyal P, Chan AT, Jaiswal N. Statistical models for the prediction of respirable suspended particulate matter in urban cities, Atmos. Environ, 2006, 40, 2068–2077.
[18] Jian L, Zhao Y, Zhu YP, Zhang MB, Bertolatti D. An application of ARIMA model to predict submicron particle concentrations from meteorological factors at a busy roadside in Hangzhou, China, Sci. Total Environ, 2012, 426, 336–345.
[19] Modarres R, Dehkordi AK. Daily air pollution time series analysis of Isfahan City, Int. J. Environ. Sci. Technol, 2005, 2(3), 259-267.
[20] Chen Y, Shi R, Shu S, Gao W. Ensemble and enhanced PM10 concentration forecast model based on stepwise regression and wavelet analysis. Atmospheric Environment, 2013, 74, 346-359.
[21] Ul-Saufie AZ, Yahaya AS, Ramli NA, Rosaida N, Hamid HA. Future daily PM_{10} concentrations prediction by combining regression models and feed forward back propagation models with principle component analysis (PCA). Atmospheric Environment, 2013, 77, 621-630.
[22] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors, Nature, 1986, 323(6088): 533-536.
[23] Lahmiri S. A Comparative Study of Backpropagation algorithms in financial prediction, International Journal of Computer Science, Engineering and Applications, 2011, 1(4), 1-21.
[24] Dong X, Zhou D X. Learning gradients by a gradient descent algorithm. Journal of Mathematical Analysis and Applications, 2008, 341(2): 1018-1027.
[25] Nissen S. Implementation of a fast artificial neural network library (FANN). Department of Computer Science University of Copenhagen, 2003, Online: [Avail.: 02.09.2016] http://leenissen.dk/fann/wp/
[26] Taspinar F, Bozkurt Z. Application of Artificial Neural Networks and Regression Models in The Prediction of Daily Maximum PM10 Concentration in Düzce, Turkey. Fresenius Environmental Bulletin, 2014, 23(10): 2450-2459.
[27] Taşpınar F. Improving Artificial Neural Network Model Predictions of Daily Average PM10 Concentrations by Applying PCA and Implementing Seasonal Models, Journal of The Air& Waste Mang. Assoc., 2015, doi: 10.1080/10962247.2015.1019652.