COMPARATIVELY USE OF TIME SERIES AND ARTIFICIAL INTELLIGENCE METHODS IN THE PREDICTION OF AIR POLLUTANTS

[1]Fatih TAŞPINAR*, [2]Kamran ABDOLLAHI

[1]Duzce University, Department of Environmental Engineering, Düzce, TURKEY
[2]Southern University and A&M College, Urban Forestry Program, Baton Rouge, LA, USA

[1]fatihtaspinar@duzce.edu.tr
[2]kamrana664@cs.com

Abstract: Air pollution is a continuing environmental problem in many part of world which affects welfare adversely. Air pollution monitoring data can thus be used to forecast concentrations of air pollutants for short-term using time series and artificial intelligence approaches. In this paper, time series modelling techniques, auto regressive integrated moving average model and another type of it with exogenous variables (ARIMA and ARIMAX), and artificial neural networks (ANNs) have been comparatively used to model particulate air pollution (PM_{10}) for predicting one-hour ahead concentration of particles in the air. An hourly based data for the years 2015-2016 was composed with including meteorological factors and air particulate concentration. The models were structured with inputting external parameters to simulate air pollution better. ARIMAX(3,1,2) model with R^2 of 0.667 and ANN(5-13-1) model with R^2 of 0.857 produced reasonable predictions over hourly dataset. The best fitting model among these models have been chosen in further tests in the prediction of one-hour ahead PM_{10} concentrations.

Keywords: Air pollution, Time Series Methods, Artificial Neural Networks.

Introduction

Air pollution problem due to particulate matter (PM) is caused by a mixture of organic and inorganic particles which are solid and liquid phase spreading out from variable sources (WHO, 2006; Sfetsos and Vlachogiannis, 2010.) These particles with an aerodynamic diameter less than or equal to 10 µm, namely PM_{10}, arise in the atmosphere mainly from the fuel combustion (Aneja et al., 2001; Kampa, M. and Castanas, 2008; Vahlsing and Smith, 2012). The highest PM_{10} levels are associated to stable meteorological conditions with thermal inversion in urban and industrial areas. Epidemiological studies showed a close relationship between outdoor particulate matter concentration and increased mortality and morbidity (Shang et al., 2013; Pope and Dockery, 2006). 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 (EC, 2008).

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 scope, long-term or short-term air pollution forecasting models have been used as an aid for air quality management. Time series models, artificial neural networks (ANNs), multiple linear regression (MLR) and hybrid models are mostly preferred approaches in air quality forecasting researches (Schlink et al., 2003; Niska et al., 2004; Perez and Reyes, 2000). 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} (Kukkonen et al., 2003; Kurt et al. 2008).

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 (Kao and Huang, 2000; Horowitz and Barakat, 1979). The most widely employed time series models (TSMs) are the non-seasonal and seasonal autoregressive integrated moving average and a type of them with external parameter models (e.g. ARMA, ARIMA, ARIMAX) in time series analysis (Goyal et al., 2006; Kumar and Goyal, 2011). 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 (Modarres and Dehkordi, 2005, Jian et al., 2012). Generally, the quality of models can vary on individual experience of issue, knowledge of time series 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, an air pollutant, PM_{10}, one-hour ahead concentration prediction of PM_{10} using ARIMA, ARIMAX and ANN based models were studied for the period of 2015-2016. Well-tuned models were then applied in short-term predictions of PM_{10} to determine a model best explains the variance in data with reduced inputs.

Materials and Methods

2.1 Data with explanatory statistics

An hourly dataset for Düzce province in Turkey was composed containing information about local meteorological parameters such as air temperature (AT, °C), wind direction (WD), wind speed (WS, m/s), relative humidity (RH, %) and mass concentration of particulate matter (PM_{10}, µg/m^3) for the period of 2015-2016. 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 shows the descriptive statistics of these variables and Fig. 1 visualizes an hourly time series plot for PM_{10} over air temperature.

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

	Valid (N)	Min.	Max.	Mean	Median	Mode	Freq. of Mode	25% Perc.	75% Perc.	Range	Std.Dev.
PM_{10}	8782	98.41	60.00	37	121	0.00	891	39	104	891	112.81
AT	8926	16.02	17.00	22	383	-13.00	42	8	23	55	9.82
WD	8926	192.29	201.00	182	71	0.00	359	123	268	359	92.79
WS	8926	0.62	1.00	1	5568	0.00	1	0	1	1	0.48
RH	8926	79.95	88.00	103	1563	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.

2.2. Modeling by Time Series Methods and ANNs

By analyzing patterns in historical data, such as trend, seasonality and noise, one can construct regressive models for predicting future data points. TSMs in forecasting are constructed based on historical data pattern in the series. Widely used kinds of TSMs are AR, ARMA, ARIMA, etc. and their multivariate forms such as ARMAX and ARIMAX (Taşpınar et al., 2013; Ibrahim et al., 2009; Suganthi and Samuel, 2012).

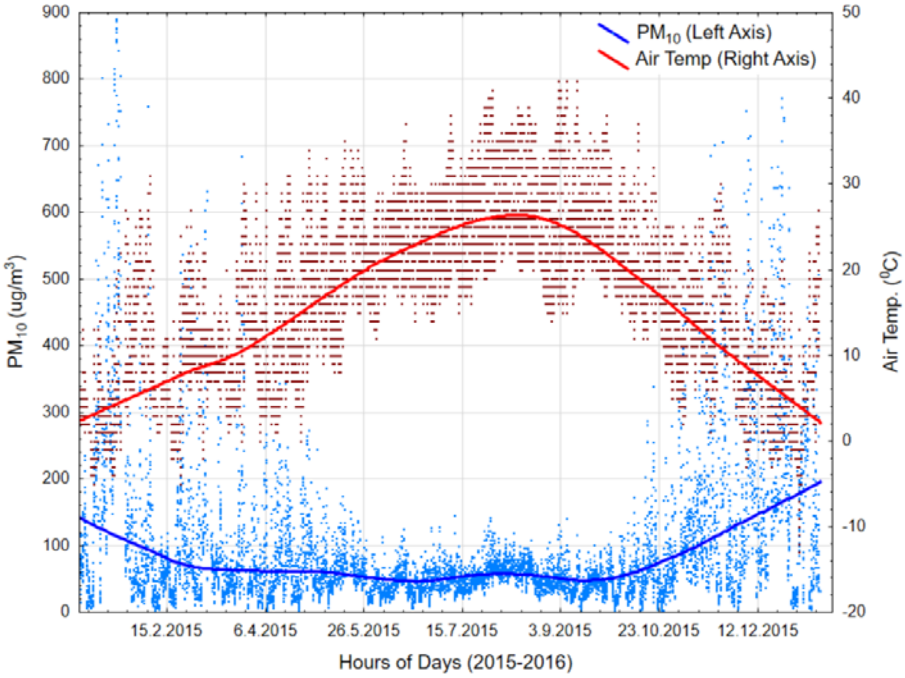

Figure 1. Hourly time series plots of PM_{10} and air temperature for 2015-2016 period.

In order to model one-hour ahead PM_{10} level, ARIMA and ARIMAX models with meteorological factors (AT, WS, WD, RH) were applied on hourly dataset. Based on these variables, the models ARIMA(p,d,q) and ARIMA[X](p,d,q) were examined. The non-negative integer elements p, d and q used in the non-seasonal models refer to the order of autoregressive part (AR(p)) and the order of differencing (I(d)) and moving average (MA(q)) parts of the models, and X refers to exogenous variables such as AT or WS used in this study, respectively. In the construction of models, the order of the model is selected by plotting the autocorrelation function (ACF) for determining the value of q used in MA(q) model and partial autocorrelation function (PACF) for determining the value of p used in AR(p) model. ARIMA model with a single variable and ARIMAX model with multi input-variable can be represented by the following equations, respectively:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q} \tag{1}$$

$$\hat{y}_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \ldots + \beta_k X_{k,t} + \frac{(1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)}{(1-\varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)} \varepsilon_t \tag{2}$$

where y_t is the t-th observation of the dependent variable, $X_{1,t}$, $X_{2,t}$, …, $X_{k,t}$ are the corresponding observations of the explanatory variables, β_0 is a constant, β_1, β_2, .. , β_k are the parameters of the regression part, and B is the backshift operator ($By_t = y_{t-1}$, $B^2 y_t = y_{t-2}$), ε_t is error residuals (~$N(0,\sigma^2)$), $\varnothing_1$, $\varnothing_2$, …, $\varnothing_p$, and θ_1, θ_2,..., θ_q are the weights for the non-seasonal autoregressive and moving average terms, respectively. In order to test the lack of fit of time series models, the Ljung-Box test was applied in model diagnostic and the most suitable model was selected according to normalized Bayesian information criteria (NBIC) (Salcedo et al., 1999; Ljung and Box, 1978).

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 (Rumelhart et al., 1986). 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. Generally, 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).

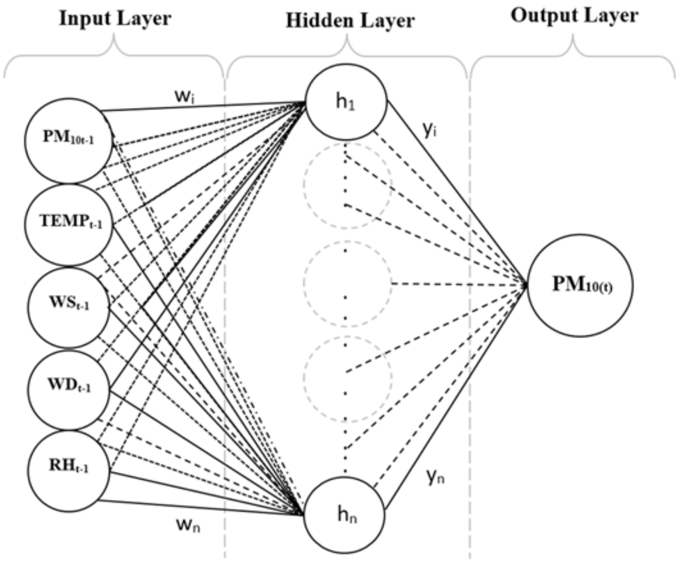

Figure 2. General structure and inputs of ANN model used in hourly PM_{10} modeling

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 or hyperbolic tangent function, respectively, as follows:

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

$$tanh(x) = \frac{1-e^{-2X}}{1+e^{-2X}} \tag{5}$$

Generally, feedforward MLP networks are trained using error back propagation (BP) algorithm (Lahmiri, 2011), which covers heuristic and numerical optimization algorithms. Heuristic techniques include gradient descent and the resilient algorithm (Dong and Zhou, 2008). So, some parameters such as learning rate, learning momentum, hidden layer neuron count etc. have been determined before training stage and then ANN model should be constructed. 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.

2.3. Data feature extraction and pre-processing

Time series dataset covering the variables PM_{10}, AT, WD, 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. Thus, the average valid data was about 91% of the entire dataset. When inputting to ANN models, the blank inputs can be skipped, however, TSMs need fully-filled input data. Hence, to execute TSMs on entire dataset, all the blank cells after missing value analysis were filled by the mean of the actual variable. The parameter WD is also converted to wind direction index (WDI) to avoid the discontinuity according to the following expression:

$$WDI = 1 + \sin\left(\mathrm{WD} + \frac{\pi}{4}\right) \tag{6}$$

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:

$$y^{'} = 0.05 + \frac{\left(y - y_{\min}\right)}{\left(y_{\max} - y_{\min}\right)} * 0.95 \tag{7}$$

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

Results and Discussion

Time Series Models and Performance Evaluation

Time series model for predicting one-hour ahead PM_{10} level is somewhat difficult comparing to ANN models. Because, tested TSMs are all hourly based which is difficult to handle in determining input lags of external variables. In fact, this problem is valid for ANN models, however, training an ANN model is much more fast and easy over a huge dataset like this.

In order to construct TSMs using ARIMA and ARIMAX methods, firstly ACF and PACF graphs were plotted for at least twenty lags of PM_{10} data. These plots were shown in Fig. 3. ARIMA model that is based on only PM_{10} data is firstly constructed. Since, the data used is based on hourly values, the periodicity is set to 24 in this case.

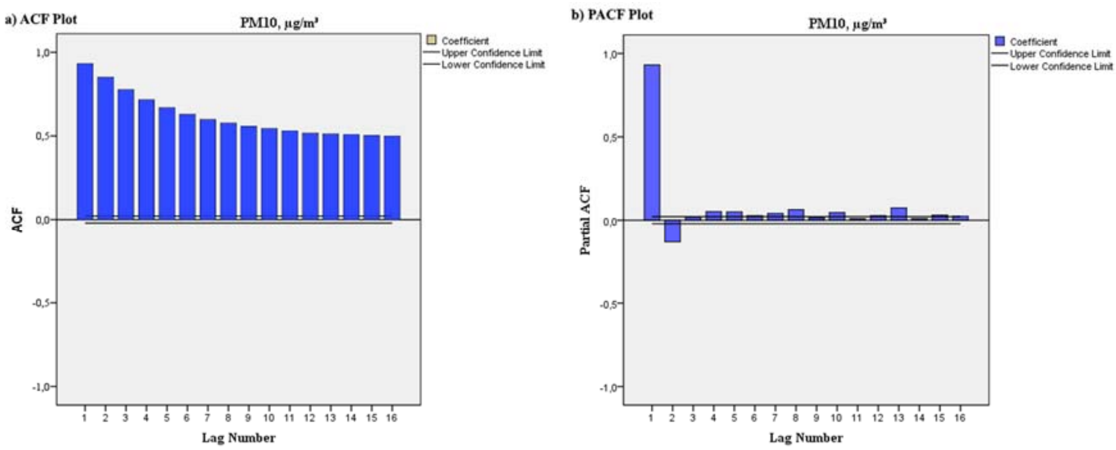

Figure 3. a) ACF and b) PACF plots for hourly PM_{10} data.

In ACF plot given in Fig. 3(a) an exponential decay with many lags over indicates moving average part in the data. PACF plot shows a significant lag at first which is an indication of AR process. Furthermore, the data is nonstationary considering high order lags in ACF plot. Thus, a non-seasonal differencing can be applied, setting parameter d to 1. So, ACF and PCAF plots for one lag non-seasonal differenced data was given in Fig. 4.

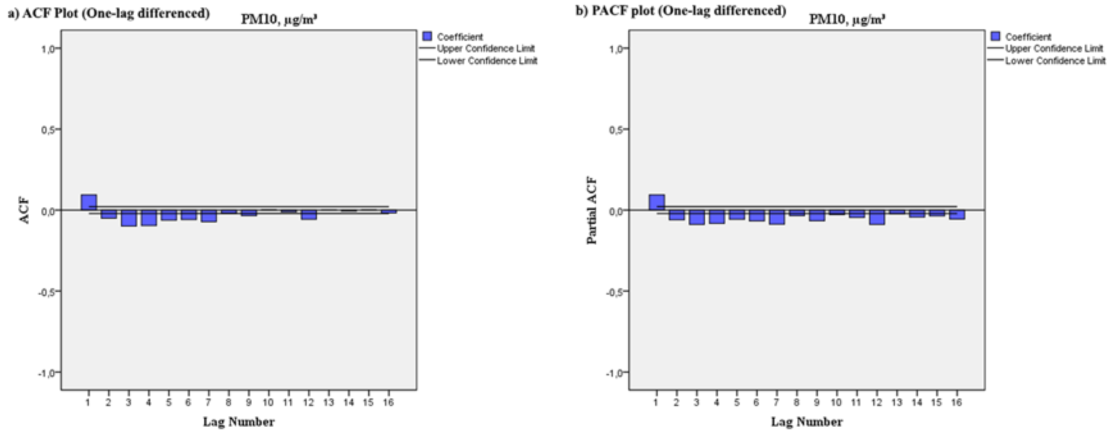

Figure 4. a) ACF and b) PACF plots for one-lag differenced hourly PM_{10} data.

One-lag differenced data shows a stationary character with time. Thus ARIMA model should include I(1) term. However, the degree of AR(p) and MA(q) processes are difficult to determine as periodicity is set to 24, which means many lags may be involved in the models. Here, although ACF plot promotes a clear MA(1) process and PACF plot promotes an AR(1) process at first sight, such ARIMA(1,1,1) model, other significant but negative lags were present in both levels at higher lags. ACF plot shows some significant lags up to 12 lag and then a sharp cut-off is observed. Therefore, we employed some models varying p and q between 1 to 3 to identify the best model without unit roots, comparing their NBIC values. Table 2 shows the models tried and related model performance statistics. Consequently, a trial-and-error work changing these model parameters was resulted in determining ARIMA(3, 1, 2) model including both AR(3) and MA(2) process with the lowest NBIC of 6.607 and R^2 of 0.663. AR lags from 1 to 3 was significant whereas MA lag at level 2 as significant. The parameter estimates of ARIMA(3,1,2) models was tabulated in Table 3 and arranged model equation was then given in Eq. (8).

Table 2: Identified ARIMA models in the prediction of hourly PM_{10} levels and model statistics

Model	NBIC	Stationary-R^2	Significant Lags (at $p<0.05$)
ARIMA(1,1,1)	0.791	0.549	AR(1), MA(1)
ARIMA(1,1,2)	0.698	0.561	AR(1), MA(1)
ARIMA(2,1,1)	0.692	0.602	AR(1,2), MA(1)
ARIMA(2,1,2)	0.702	0.597	AR(1,2), MA(1)
ARIMA(3,1,1)	0.675	0.653	AR(1,2,3), MA(1)
ARIMA(3,1,2)	*0.607*	*0.663*	AR(1,2,3), MA(2)
ARIMA(3,1,3)	0.670	0.658	AR(1,2,3), MA(2)

Table 3: Parameter estimates of ARIMA(3,1,2) model.

Model Partitions		Parameter Estimate	SE	t	Significance (at $p<0.01$)
AR	Lag 1	0.218	0.010	21.289	0.000
	Lag 2	0.845	0.005	158.374	0.000
	Lag 3	-0.364	0.010	-35.627	0.000
Differencing Order		1			
MA	Lag 2	0.996	0.002	635.015	0.000

$$PM10_t = 1.218 \cdot PM10_{t-1} - 1.063 \cdot PM10_{t-2} - 0.481 \cdot PM10_{t-3} + 0.364 \cdot PM10_{t-4} + (1 + 0.996.B)\varepsilon_t \quad (8)$$

Estimates of ARIMA(3,1,2) model with upper and lower confidence limits against to original data as time series plot were given in Fig. 5, which shows that estimated points fitted the historical PM_{10}pattern very well. Also, based on this univariate ARIMA(3,1,2) model, an ARIMAX model with the inputs from AT, WS, WD and RH were constructed. These meteorological factors affect PM_{10} level in air in real-life, so this situation can be simulated by an ARIMAX model. ARIMAX models are constructed by transfer functions by calculating weights of external variables. After parameter estimation a model equation can be arranged based on significant lags of variables.

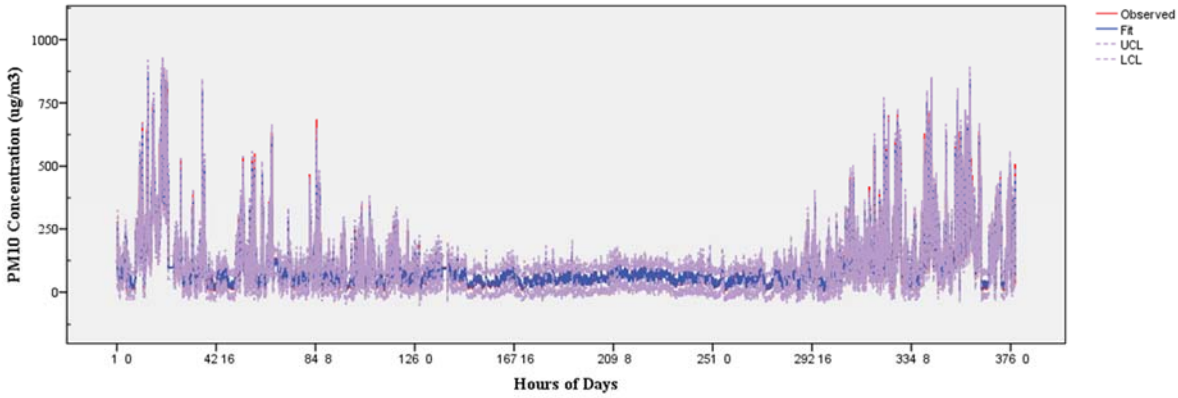

Figure 5. Line plot of the estimates of ARIMA(3,1,2) model with confidence limits against to observed data.

Based on the parameter estimates for ARIMAX(3,1,2) model including the terms from meteorological factors with R^2 of 0.667 and NBIC of 6.607, arranged ARIMAX model equation is given in Eq. (9) as follows:

$$\begin{aligned} PM10_t &= 1.218 \cdot PM10_{t-1} - 1.063 \cdot PM10_{t-2} - 0.481 \cdot PM10_{t-3} + 0.364 \cdot PM10_{t-4} + (1 + 0.996.B) \cdot \varepsilon_t + \\ &(AT_t - 0.768 \cdot AT_{t-1}) + (WDI_t - 0.008 \cdot WDI_{t-1}) + ((1 + 9.021 \cdot WS_t) / (1 - 1.193 \cdot WS_{t-1} - 0.552 \cdot WS_{t-2})) + \\ &(RH_t - 0.463 \cdot RH_{t-1} - 0.484 \cdot RH_{t-4}) \end{aligned} \tag{9}$$

Parameters estimates of ARIMAX model showed that $PM_{10(t)}$ was predicted by using AT at lags 0-1, WDI at lags 0-1, WS at lags 0-2 and RH at lags 0,1,4. Performance of ARIMA and ARIMAX models for one-hour ahead PM10 concentration prediction is very similar, comparing R^2 values. Thus, we selected PM_{10} ARIMAX model with external parameters from meteorological factors as benchmark for further tests.

ANN Models and Performance Evaluation

Designing of ANN models are related to selection of some parameters such as hidden layer neuron count and learning rate. Before construction of ANN model, 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 lags of all model inputs as shown in Fig. 2, which can be written as $PM10_t = f_{net}(PM10_{t-1}, AT_{t-1}, WD_{t-1}, WS_{t-1}, RH_{t-1})$ model. 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. Therefore, the ANN model were set by utilizing cascading-training technique of this library.

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, setting the validation process at every 10 training epochs. A starting learning rate of 0.45 was gradually decreased by 1.1% at every epoch during the cascading-training procedure, which was resulted in an ANN(5-13-1) model, including 13 hidden layer neuron. Several experiments with different structures were tried as mentioned here, however, ANN(5-13-1) model produced the best error measure and model accuracy with RMSE of 0.478, training R^2 of 0.857 and testing R^2 of 0.841. Also, Index-of-Agreement (IA) that measures prediction errors was calculated to test the quality of fit. IA value of 0.81 obtained with this model, which is close to 1.0, suggested 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 model was visualized in Fig. 6. The performance plot of the results of the best ANN model for the whole dataset is visualized on Fig. 3. The red line indicates an exact fit of R^2=1.0, hashed black line indicates a linear fitting line of R^2=0.857. 95% confidence band limits were also shown in blue-dotted lines, clearly indicating the most of the data points fall in the band limits.

Hourly time series dataset was plotted against to predictions of ANN(5-13-1) model in Fig. 7. All the data is well followed by ANN model, simulating historical pattern of hourly PM_{10} concentration. As it can be seen in Fig. 7 that at some extreme conditions with elevated PM_{10} levels, particularly the levels higher than 550 µg/m^3 observed during strictly calm days. Due to air circulation issues occurred in Duzce province in calm nights during winter periods, extreme conditions can be observed. Therefore, elevated PM_{10} levels mostly occurred winter times with a rate of 2% for levels higher than 500 µg/m^3. However, most frequent PM_{10} values within a range of $\mu \pm 3\sigma$ were predicted reasonably successfully. Extreme value problem for ANN model is a well-known issue, because neural networks cannot successfully evaluate less trained input values or less frequent data observed at extreme conditions. However, all the ANN models experimented in the tests were very successful in the predictions comparing to TSMs.

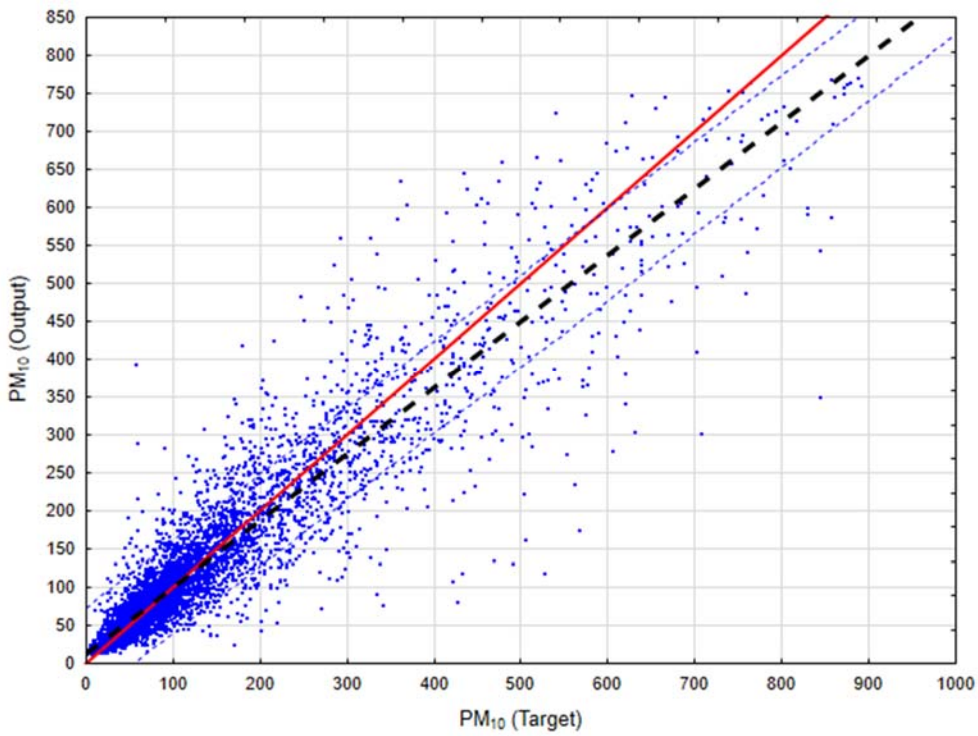

Figure 6. Performance plot for ANN(5-13-1) model and linear fitting line

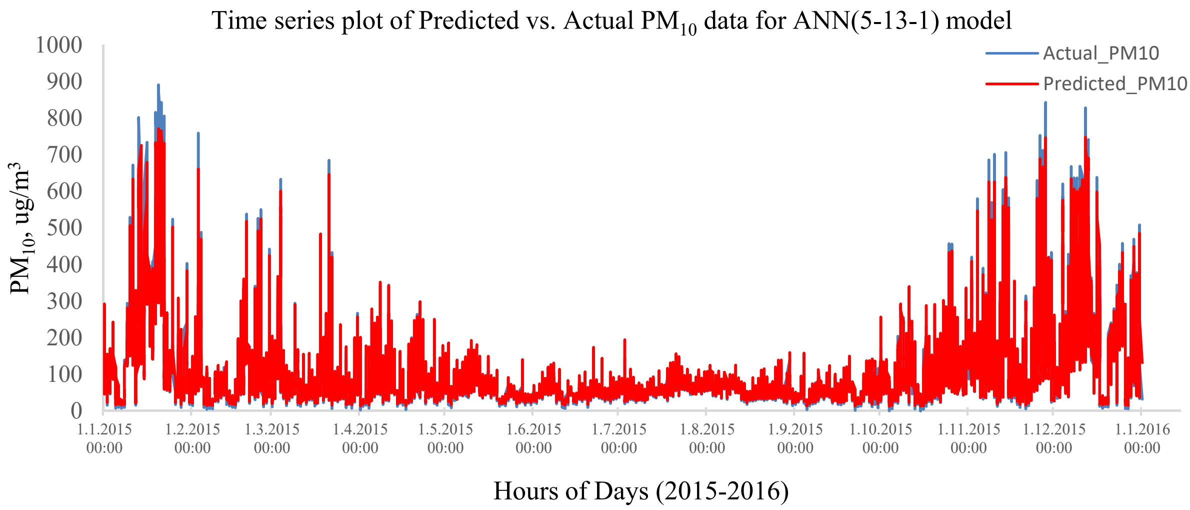

Figure 7. Hourly time series plots for the predictions of the ANN model against to actual data

Conclusion

The present study investigated short-term PM_{10} modelling based on hourly monitoring data of meteorological factors and particle concentrations. In order to obtained a reasonable model for time series data, we have used time series modeling techniques and ANN models. ARIMA and ARIMAX models were applied to data as TSMs. A univariate ARIMA(3,1,2) model were fitted to predicted hourly PM_{10} level one-hour ahead. This model included

AR(3) and MA(2) processes with one lag differentiation, with R2 of 0.663. Based on this model, a multivariate ARIMAX(3,1,2) model fitted by covering the terms from $PM_{10(t-1...t-4)}$, $AT_{(t..t-1)}$, $WDI_{(t..t-1)}$, $WS_{(t..t-2)}$ and $RH_{(t,t-1,t-4)}$ with R^2 of 0.667. Therefore, ARIMAX model is selected as benchmark to compare predictions by ANN model.

ANN models for prediction next-hour PM_{10} level were designed. The best ANN model is in the form of 5-13-1 with 13 hidden neuron in middle layer, with testing R^2 of 0.857. The input vector of ANN was $PM10_{t-1}$, AT_{t-1}, WD_{t-1}, WS_{t-1}, RH_{t-1} and the output was $PM10_t$. Cascading-learning 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 TSMs in general. On the other hand, ANN model did not tend to overpredict as time series plots indicated. Comparing to TSMs, ANN models were very flexible in handling of model and execution, because TSMs need to be updated with new model residuals before operating on new input data in any new hour. Hence, TSMs are not tolerate any interruption on time scale as they were constructed on sequenced residual data, so their handling in real-life applications are difficult. On the other hand, ANN models obtained after training step can be used on any data with proper input vector. Therefore, we concluded that a real-life application of emergency perception strategy can be designed based on ANN models for hourly PM_{10} prediction. However, to properly tackle with extreme values of air PM_{10} levels observed during pollution episodes in winter periods in particular, some other methods such as time based hybrid models or discrete ANN models for higher levels should be considered.

References

Aneja, V.P., Agarwal, A., Roelle, P.A., Phillips, S.B., Tong, Q., Watkins, N. and Yablonsky, R. (2001). Measurements and analysis of criteria pollutants in New Delhi, India. Environment International,27,35-42.

Dong X. and Zhou D. X. (2008). Learning gradients by a gradient descent algorithm. Journal of Mathematical Analysis and Applications 341(2): 1018-1027.

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.

G. E. P. Box, G. M. Jenkins, G. C. Reinsel, Time series analysis: forecasting and control, 3rd ed., Prentice Hall, New Jersey 1994.

Horowitz J., Barakat S. (1979). Statistical analysis of the maximum concentration of an air pollutant: effects of autocorrelation and non-stationarity, Atmos. Environ., 13, 811-818.

Ibrahim M. Z., Zailan R., Ismail M., Lola M. S. (2009). Forecasting and Time Series Analysis of Air Pollutants in Several Area of Malaysia, American Journal of Environmental Sciences, 5(5): 625-632.

Jian L., Zhao Y., Zhu Y. P., Zhang M. B., Bertolatti D. (2012). An application of ARIMA model to predict submicron particle concentrations from meteorological factors at a busy roadside in Hangzhou, China, Sci. Total Environ., 426, 336–345.

Kampa, M. and Castanas, E. (2008). Human health effects of air pollution. Envir. Pollution, 151, 362–367.

Kao J., Huang S. (2000). Forecasts Using Neural Network versus Box-Jenkins Methodology for Ambient Air Quality Monitoring Data, J. Air Waste Manage. Assoc., 50 (2), 219-226.

Kukkonen J., Partanena L., Karppinen A., Ruuskanen J., Junninen H., Kolehmainen M., Niska H., Dorling S., Chatterton T., Foxall R., Cawley G. (2003). 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 37, 4539–4550.

Kumar A., Goyal P. (2011). Forecasting of Daily Air Quality Index in Delhi, Sci. Total Environ., 409, 5517–5523.

Kurt A., Gulbagci B., Karaca F., Alagha O. (2008). An online air pollution forecasting system using neural networks, Environment International 34, 592–598.

Lahmiri S. (2011). A Comparative Study of Backpropagation algorithms in financial prediction, International Journal of Computer Science, Engineering and Applications (IJCSEA) 1(4).

Ljung G., Box G. (1978). On a Measure of Lack of Fit in Time Series Models, Biometrika 1978, 65, 297 303.

Modarres R., Dehkordi A. K. (2005). Daily air pollution time series analysis of Isfahan City, Int. J. Environ. Sci. Technol., 2 (3), 259-267.

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

Nissen, S. (2003) Implementation of a fast artificial neural network library (FANN). Department of Computer Science University of Copenhagen, Online: [Avail.: 01.06.2016] http://leenissen.dk/fann/wp/

P. Goyal, Chan A. T., Jaiswal N. (2006). Statistical models for the prediction of respirable suspended particulate matter in urban cities, Atmos. Environ., 40, 2068–2077.

Perez, P., Trier, A., Reyes, J. (2000) Prediction of $PM_{2.5}$ concentrations several hours in advance using neural networks in Santiago, Chile. Atmospheric Environment, 34, 1189–1196.

Pope, C.A. and Dockery, D.W. (2006). Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc., 56, 709–742.
Rumelhart D. E., G. E. Hinton and R. J. Williams (1986). Learning representations by back-propagating errors. Nature 323(6088): 533-536.
Salcedo R. L. R., Alvim Ferraz M. C. M., Alves C. A., Martins F. G. (1999). Time-series analysis of air pollution data, Atmos. Environ. 1999, 33 (15), 2361-2372.
Schlink, U., Dorling, S., Pelikan, E., Nunnari, G., Cawley, G., Junninen, H., Greig, A., Foxall, R., Eben, K., Chatterto, T., Vondracek, J., Richter, M., Dostal, M., Bertucco, L., Kolehmainen, M., Doyle, M. (2003) A rigorous intercomparison of ground-level ozone predictions. Atmospheric Environment, 37, 3237–3253.
Sfetsos, A. and Vlachogiannis, D. (2010). A new methodology development for the regulatory forecasting of PM10. Application in the Greater Athens Area, Greece. Atmospheric Environment, 44, 3159-3172.
Shang, Y., Sun, Z., Cao, J., Wang, X., Zhong, L., Bi, X., Li, H., Liu, W., Zhu, T. and Huang, W. (2013). Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. Environment International, 54, 100–111.
Suganthi L., Samuel A. A. (2012). Energy models for demand forecasting - A review, Renewable Sustainable Energy Rev. 16 (2), 1223-1240.
Taspinar F. and Bozkurt Z. (2014). Application of Artificial Neural Networks and Regression Models in The Prediction of Daily Maximum PM10 Concentration in Düzce, Turkey. Fresenius Environmental Bulletin 23(10): 2450-2459.
Taşpınar F., Çelebi N., Tutkun N. (2013). Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods, Energy Buildings, 56, 23-31.
Vahlsing, C. and Smith, K.R. (2012). Global review of national ambient air quality standards for PM10 and SO2 (24 h). Air Qual. Atmos. Health, 5, 393–399.
WHO (2006). Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment. WHO, Geneva.

