Asian Journal of Water, Environment and Pollution, Vol. 14, No. 4 (2017), pp. 109–113.
DOI 10.3233/AJW-170041

Forecasting of PM_{10} Using Autoregressive Models and Exponential Smoothing Technique

Vibha Yadav* and Satyendra Nath
Department of Environmental Science and NRM, College of Forestry, Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, Uttar Pradesh-211007 India
✉ yvibha3@gmail.com

Received June 1, 2017; revised and accepted September 11, 2017

Abstract: Particulate matter with 10 μm or less in diameter (PM_{10}) have adverse effects on environment and human health. To reduce PM_{10} emissions in India, it is essential to have models that accurately estimate and predict PM_{10} concentrations for reporting and monitoring purposes. In this paper Exponential Smoothing Technique and Autoregressive (AR) models are developed to forecast 1-month ahead value of PM_{10} for Allahabad city which is novelty of this study. AR (1) and AR (5) models are developed using Burge and Yule Walker methods. The mean absolute percentage error (MAPE) for Burge method in AR (1) and AR (5) are 14.23% and 10.20%. The MAPE for Yule Walker method in AR (1) and AR (5) are 32.72% and 31.13%. The MAPE in Exponential Smoothing Technique is 5.81% which shows it forecasts better than AR model based on Burge and Yule Walker methods. It is found that Burge Method in AR (5) has less MAPE than Yule Walker Method. Therefore Exponential Smoothing Technique can be used to forecast PM_{10} for cities in India, showing it is beneficial for giving prior information for human health.

Key words: PM_{10}, forecasting, AR models, exponential smoothing.

Introduction

In recent times it seems that air pollution has been a critical problem in major cities. Overexploitation of natural resources leads to unsustainability and highly affecting the air quality. The air quality can be determined by numerous defined air quality indicators such as particulate matter (PM_{10}) carbon monoxide (CO), sulphur dioxide (SO_2) and nitrogen dioxide (NO_2). Increase in concentration of particular indicator from a defined standard value may impart damage of an organ in human beings (Filho and Fernandes, 2013).

Classification of particles can be done on the basis of size as may be coarse or fine. Fine particles are $PM_{2.5}$ having diameter of 2.5 μm or less whereas coarse particles are PM_{10} having diameter of 10 μm or less. Particulate matters are those which can be originated from natural sources like volcanic eruptions and forest fires beside anthropogenic sources like combustion of fossil fuels, vehicular emissions, industrial emissions, sewage disposal, waste dumping etc. Particulate suspended in air less than permissible limit are directly inhaled by humans and causes many pulmonary disorder, cardiac problems, skin infections which ultimately may result to be fatal (Whalley and Zandi, 2016). Forecasting of PM_{10} is required to ensure the permissible prescribed concentration of particulate matter estimated by government agencies. To attempt the approaches of forecasting and modelling of air pollutant, a large number of methods have been developed. Syafei et al. (2015) developed prediction model of air pollutants level using linear model with

*Corresponding Author

component analysis. To obtain the subsets of predictor variables, Independent Component Analysis (ICA) and Principal Component Analysis (PCA) were used. Stoimenova (2016) constructed the stochastic modelling of problematic air pollution with particulate matter in City of Pernik, Bulgaria. Alwee (2013) forecasted crime rates using Hybrid Support Vector Regression and Autoregressive Integrated Moving Average Models improved by particle swarm optimization for property climate rates with economic indicators. Salini and Perez (2015) studied the dynamic behaviour of Fine Particulate Matter in Santiago, Chile and developed a forecasting model based on real data exhibit deterministic chaotic behaviour.

Kaushik and Melwan (2007) used seasonal Autoregressive Integrated Moving Average (ARIMA) approach, implemented by Box–Jenkins to forecast the level of ambient air quality parameters at ITO intersections in Delhi, India. Kumari et al. (2013) have been compared two different autoregressive models to check the forecasting ability using air pollution concentration present in R.K Puram areas of New Delhi. Liu and Li (2015) forecasted the trend of $PM_{2.5}$ at Guangzhou by application of comprehensive forecasting model such as Autoregressive Integrated Moving Average Model (ARIMA), Artificial Neural Networks (ANNs) model and Exponential Smoothing Method (ESM). Chaudhuri et al. (2011) developed the Autoregressive and Radial Basis Function Network (RBNF) models to forecast the concentration of some important atmospheric pollutant in Kolkata. Robles-Diaz et al. (2008) demonstrated the potential of forecasting the air quality at Temuco, Chile and developed a hybrid ARIMA and Artificial Neural Networks model. In this study Autoregressive (AR) models and exponential smoothing technique are developed to forecast 1-month ahead PM_{10} values for Allahabad city which is the novelty of this research.

Methodology

Study Area and Data Collection

Monthly time series data of PM_{10} during 2013-2015 for Bharat Yantra Nigam, Allahabad, U.P India was taken from Uttar Pradesh Pollution Control Board (UPCB), Lucknow. The values of PM_{10} are shown in Table 1.

Table 1: Measured PM_{10}

Months	*PM_{10} for year 2013*	*PM_{10} for year 2014*	*PM_{10} for year 2015*
Jan	247	274	216.4
Feb	252	314	236
March	205	246	203
Apr	198	198	234
May	176	240	233
June	217	209	254
July	186	246	239
Aug	206	242	235
Sep	218	214	228
Oct	269	219	233
Nov	236	231	230
Dec	256	206	260
Average	22.2	236	233.5
Max	269	314	260
Min	176	198	203

Exponential Smoothing Technique

In simple exponential smoothing, the smoothed values is forecasted. When time series is fairly stable and has no significant trend then exponential smoothing technique is used. The different types of common smoothing are moving average, weighted moving average, centered moving average and exponential smoothing. It is given by following equations:

$$F_{t+1} = \alpha y_t + (1 - \alpha) F_t \tag{1}$$

where F_{t+1} is forecast value for period $t + 1$, y_t actual value for period t, F_t forecast value for period t, α is smoothing constant (0 to 1). In this study α is taken as 0.2.

Autoregressive Model

In statistics application, an autoregressive (AR) model is a random process representation to describe time varying process. The AR model discusses about linear dependency of output on its own previous values on a stochastic term (an imperfectly predictable term); thus the model is in the form of a stochastic difference equation. When AR model is combined with moving-average (MA) model it is called autoregressive moving-average (ARMA). An autoregressive model of order p is denoted by $AR(p)$. It is defined as

$$X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t \tag{2}$$

where $\phi_i \ldots \phi_p$ are the parameters of the model, c is a constant and ε_t is a white noise. This can be written by using the backshift operator B.

$$X_t = c + \sum_{i=1}^{p} \phi_i B_i X_t + \varepsilon_t \tag{3}$$

so that, moving the summation term to the left side and using polynomial notation, we have

$$\phi(B)\, X_t = c + \varepsilon_t \tag{4}$$

The AR models can be implemented by Burge method (BM) and Yule Walker method (YWM). The syntax in matlab codes are as follows:

[ar_coeffs noise variance]=arburg(y, order) for BM
[ar_coeffs noise variance]=aryule (y, order) for YWM
where y is data samples.

The error in AR model forecasting are evaluated by mean absolute percentage error (MAPE), root mean square error (RMSE) and mean bias error (MBE) as follows:

$$\text{MAPE} = \left(\frac{1}{n} \sum_{i=1}^{n} \left| \frac{AP_{i(AR)} - AP_{i(actual)}}{AP_{i(actual)}} \right| \right) \times 100 \tag{5}$$

$$\text{RMES} = \sqrt{\sum_{i=1}^{n} \frac{\left(AP_{i(AR)} - AP_{i(actual)}\right)^2}{n}} \tag{6}$$

$$\text{MBE} = \sum_{i=1}^{n} \frac{\left(AP_{i(AR)} - AP_{i(actual)}\right)}{n} \tag{7}$$

where $AP_{i(AR)}$ forecasted value of air pollutant by AR is model, $AP_{i(actual)}$ is actual value of air pollutant and n is number of air pollutant data samples.

Results and Discussions

The exponential smoothing, AR (1) and AR (5) models are developed in Matlab (R2011a) using PM_{10} values shown in Table 1. The PM_{10} value of year 2015 are used for validation of AR models and exponential smoothing forecasting model. The AR models incorporates Burge and Yule Walker method. The forecasting value of 1-month ahead PM_{10} value using Burge and Yule Walker method in AR models are shown in Table 2 and with exponential smoothing method is shown in Table 3. The forecasting accuracy of AR models are calculated with MAPE given by Lewis (1982). The MAPE ≤ 10% indicates high forecasting accuracy, 10% ≤ MAPE ≤ 20% indicates good forecasting, 20% ≤ MAPE ≤ 50% indicates reasonable forecasting and MAPE ≥ 50% indicates inaccurate forecasting. The MAPE, RMSE and MBE are used to test forecasting accuracy of AR models and are shown in Tables 4 and 5.

Table 2: Burge and Yule Walker forecasting

Months	*Measured*	*Burge method*		*Yule Walker method*	
	PM_{10} for year 2015	*AR (1)*	*AR (5)*	*AR (1)*	*AR (5)*
Jan	216.4	204.94	209.02	196.74	194.52
Feb	236	203.89	210.80	187.94	188.33
March	203	202.85	214.72	179.58	182.56
Apr	234	201.82	210.20	171.64	174.59
May	233	200.80	210.91	164.09	167.42
June	254	199.79	210.87	156.92	161.01
July	239	198.79	211.73	150.10	154.92
Aug	235	197.81	210.80	143.63	148.93
Sep	228	196.83	210.92	137.47	143.21
Oct	233	195.86	210.78	131.63	137.79
Nov	230	194.90	210.92	126.07	132.61
Dec	260	193.95	210.68	120.79	127.65

Table 3: Exponential smoothing forecasting

Months	*Measured PM_{10} for year 2015*	*Forecasted PM_{10} for year 2015*	*MAPE (%)*	*RMSE*	*MBE*	*Accuracy*
Jan	216.4	225.65	5.81	16.21	-7.79	High forecasting
Feb	236	225.65				
March	203	225.65				
Apr	234	225.65				
May	233	225.65				
June	254	225.65				
July	239	225.65				
Aug	235	225.65				
Sep	228	225.65				
Oct	233	225.65				
Nov	230	225.65				
Dec	260	225.65				
Average	233.5	225.65				
Max	260	225.65				
Min	203	225.65				

Table 4: Burge method based AR models

Model	*MAPE(%)*	*RMSE*	*MBE*	*Accuracy*
AR (1)	14.23	37.76	-34.09	Good Forecasting
AR(5)	10.20	26.85	-22.41	Good Forecasting

Table 5: Yule Walker method based AR models

Model	*MAPE(%)*	*RMSE*	*MBE*	*Accuracy*
AR (1)	32.78	84.75	-77.89	Reasonable forecasting
AR(5)	31.13	80.29	-73.98	Reasonable forecasting

It is found that exponential smoothing method has least MAPE of 5.81% showing high forecasting accuracy of PM_{10} as per Lewis.

Conclusions

A novel exponential smoothing and AR models based on Burge and Yule Walker methods are proposed that is capable of exploiting the strengths of traditional time series approaches for air quality forecasting. The MAPE of Burge based AR (1) and AR (5) are 14.23% and 10.20% whereas MAPE of Yule Walker based AR (1) and AR (5) are 32.78% and 31.13% showing AR model based Burge method give better forecasting of PM_{10} than Yule Walker methods. The MAPE of exponential smoothing based PM_{10} forecasting is 5.81%, showing high forecasting accuracy. The developed model in this study can be used to forecast future value of PM_{10} for Allahabad city which is useful for environmental engineers to analyze and to characterize concentration measurements of air pollutants with time.

Future research work is focused on forecasting of air pollutants of different sites in India using ARMA and ANN models.

Acknowledgement

The authors acknowledge the Uttar Pradesh Pollution Control Board (UPCB) Lucknow, India for providing air pollutant data for this research.

References

Alwee, R., Shams Uddin, H. and R. Sallehuddin (2013). Hybrid Support Vector Regression and Autoregressive

Integrated Moving Average Models Improved by Particle Swarm Optimization for Property Crime Rates Forecasting with Economic Indicators. *The Scientific World Journal*, Article ID 951475, 1-11.

Chaudhuri, S., Das, D., Middey, A. and S. Goswami (2011). Forecasting the concentration of atmospheric pollutants: Skill assessment of Autoregressive and Radial Basis Function Network Models. *International Journal of Environmental Protection*, **1(5):** 41-47.

Filho, A. and M. Fernandes (2013). Time series forecasting of pollutant concentration levels using Particle Swarm Optimization and Artificial Neural Networks. *Quim. Nova*, **36(6):** 783-789.

http://www.uppcb.com/ambient_quality.htm.

https://en.wikipedia.org/wiki/Autoregressive_model.

http://www.statoek.wiso.unigoettingen.de/veranstaltungen/graduateseminar/SmoothingMethods_Narodzonek-Karpowska.pdf.

Kaushik, I. and R. Melwan (2007). Time series analysis of ambient air quality at ITO intersection in Delhi (India). *Journal of Environmental Research and Development*, **2(2):** 1-5.

Kumari, S., Jain, V.K. and Isha (2013). Autoregression model for the prediction of ambient air pollutant concentration in Delhi. *International Journal of Environmental Science: Development and Monitoring*, **4(2):** ISSN 2231-1289.

Lewis, C.D. (1982). International and business forecasting methods. Butter-worths, London.

Liu, J. and L. Li (2015). Application Study of Comprehensive Forecasting Model Based on Entropy Weighting Method on Trend of PM2.5 Concentration in Guangzhou, China. *International Journal of Environmental Research and Public Health*, **12:** 7085-7099.

Robles-Diaz, A.L., Bravo, O.C.J., Fu, S.J., Reed, D.G., Chow, C.J., Watson, G.J. and A.J. Herrera-Moncada (2008). A hybrid ARIMA and Artificial Neural Networks model to forecast the particulate matter in urban areas. *Atmospheric Environment*, **42:** 8331-8340.

Salini, A. and P. Perez (2015). A Study of the Dynamic Behaviour of Fine Particulate Matter in Santiago, Chile. *Journal of Aerosol and Air Quality Research*, **15:** 154-165.

Stoimenova, P. (2016). Stochastic Modeling of Problematic Air Pollution with Particulate Matter in the City of Pernik, Bulgaria. *Ecologia Balkanica*, **8(2):** 33-41.

Syafei, D., Fujiwara, A. and J. Zhang (2015). Prediction Model of Air Pollutant Levels Using Linear Model with Component Analysis. *International Journal of Environmental Science and Development*, **6(7):** 519-525.

Whalley, J. and S. ZandI (2016). Particulate matter sampling techniques and data modelling methods. Air quality measurement and modelling.