Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 11, pp.41643-41646, November, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

PREDICTION OF SURFACE OZONE USING ARIMA MODEL

*Arputharaj, S. and Samuel Selvaraj, R.

Department of Physics, Presidency College, Chennai

ARTICLE INFO

Article History:
Received 25th August, 2016
Received in revised form
22nd September, 2016
Accepted 29th October, 2016
Published online 30th November, 2016

Key words:

ARIMA,
Chennai,
Ground level ozone,
Forecasting.

ABSTRACT

Analysis of the ground level ozone concentration is vital for the purpose of forecasting and in identifying the changes and impacts that are very crucial for an agro-based economy like the city of Chennai. The paper examines Ozone concentration and meteorology in Chennai at Koyambedu which is a major big Bus terminus. This will be very useful for further research and future researchers. A linear ARIMA model was used for the forecast of Ozone over the study area. The performance evaluation was also done to check the best fit model for forecast.

Citation: Arputharaj, S. and Samuel Selvaraj, R. 2016. "Prediction of surface ozone using arima model", *International Journal of Current Research*, 8, (11), 41643-41646.

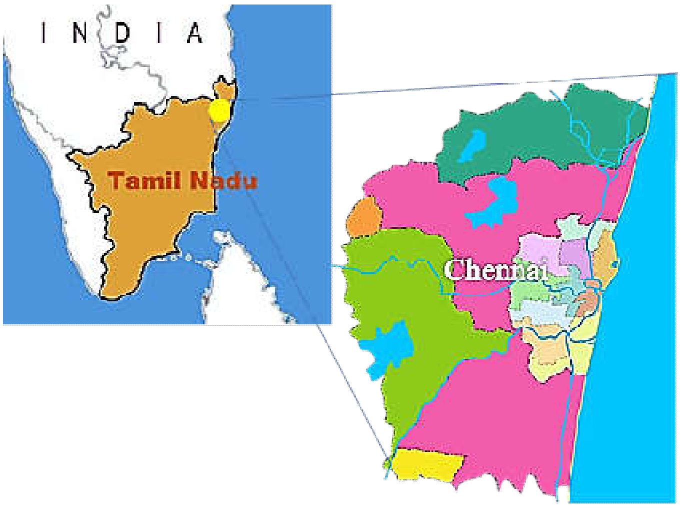
INTRODUCTION

Gases in the atmosphere with concentration less than 1ppmv are called trace gasses. These trace gases have a major impact on the environment in spite of their relatively low concentration. Few of these gases affect the climate trough atmospheric green house effect and few other are toxic and cause damage to plant and animal life. One such important gas is Ozone. Ozone in the stratosphere protects the earth from the UV radiation. In the troposphere they act as green house gas and helps in climate formation (Crutzen et al., 1998). They also take part in the physicochemical processes. Ozone has strong oxidising property. Increased concentration of the surface ozone or ground level ozone will cause severe damage to humans, animals, materials and vegetation. Ozone will undergo photochemical reaction with the atmospheric pollutants like hydrocarbon compounds, nitrogen oxides, etc and affect the human health (Kalabokas et al., 2000). They also affect the leaf and plant growth. Thus it is vital to predict the ozone concentration at a place and take necessary precaution to prevent adverse effects. Therefore, the development of effective prediction models of ozone concentrations in urban areas is important. Management of control and public warning strategies for ozone levels (particularly densely populate area) requires accurate forecasts of the concentration of ambient ozone (Prybutok *et al.*, 2000), to check whether or not a threshold is exceeded. Such information could be exploited by environmental and medical authorities to announce public health warning. Surface ozone was measured throughout Tamil Nadu during the year 2011 and it was found that Kanniyakumari district had the highest daily average of 17.8 ppbv (Samuel Selvaraj *et al.*, 2013), and the previous studies show that the maximum value of surface ozone recorded was 56ppbv during April 2011 in Kanyakumari district. Moreover the surface ozone levels studied in Chennai during 2004- 2005 it was found that the hourly values varied from 1 ppbv to 50.27 ppbv (Pulikesi *et al.*, 2006). The ground-level O_3 concentration in Chennai varied between 2 ppbv and 53 ppb in the urban area .All these studies have indicated that the effects of O3 on vegetation were quite severe in India and other parts of Asia (Emberson *et al.*, 2001). A limited number of studies have analyzed the seasonal variation in surface ozone and its precursors in urban areas of India (Naja *et al.*, 2002). To the best of the author's knowledge, no measurements have been carried out over Chennai metropolitan area in recent years. Thus this study measures the surface ozone concentration in the urban city of Chennai.

**Corresponding author: Arputharaj, S.,*
Department of Physics, Presidency College, Chennai

Study area

Chennai is situated on the south east coast of India and north east coast of Tamil Nadu. This area is one of the most highly populated urban sites. Chennai lies on the thermal equator and is also a coastal. The latitude and longitude of the center of the city are E80° 14’51” and N13° 03’ 40” (Monsingh *et al*., 2009 and Regional Meteorological Cetre). The geographical location of the experimental site is shown in Fig.1 and it is located in south Chennai. The different sources of air pollution are classified under the following categories Transport, Industries, Residential in Chennai City. This Urban City can be divided into four areas, North, Central, South and West. The Northern part is primarily an industrial area comprising of petrochemical industries in the Manali area and other general industries in Ambattur. Chennai has many industrial areas. This study was conducted at Koyembedu which houses Chennai’s moffussil Bus terminus and 100’s of Buses and other vehicles ply daily and hence the vehicular emission is very high. This site is surrounded by the number of Industrial areas located within a short radius.



Data

The measurement of surface ozone were carried out in the area which has selected to represent the typical residential with high commercial and traffic influenced. Using Aeroqual 200 series Ozone data has measured. Surface ozone measurements were carried out daily and ten measurements were made on all days between 08.00 hrs and 17.00 hrs (IST) during the period from June 2011 to September 2012.

MATERIALS AND METHODS

In this paper we present ARIMA (Auto Regressive Integrated Moving Average) for studying and forecasting the ground level ozone in Chennai based on the past observations.

ARIMA model

The ARIMA model is an extension of the ARMA model in the sense that by including auto-regression and moving average it has an extra function for differencing the time series. If a dataset exhibits long-term variations such as trends, seasonality and cyclic components, differencing a dataset in ARIMA allows the model to deal with them. A brief procedure for the Box and Jenkins (ARIMA) model is discussed below.

Autoregressive process

Most time series consist of elements that are serially dependent in the sense that one can estimate a coefficient or a set of coefficients that describe consecutive elements of the series from specific, time-lagged (previous) elements. Each observation of the time series is made up of random error components (random shock, ε) and a linear combination of prior observations.

Moving average process

Independent from the autoregressive process, each element in the series can also be affected by the past errors (or random shock) that cannot be accounted for by the autoregressive component. Each observation of the time series is made up of a random error component (random shock, ε) and a linear combination of prior random shocks.

General form of non-seasonal and seasonal ARIMA

ARIMA models are sometimes called Box-Jenkins models. An ARIMA model is a combination of an autoregressive (AR) process and a moving average (MA) process applied to a non-stationary data series. As such, in the general non-seasonal ARIMA-(p,d,q) model, AR-(p) refers to order of the autoregressive part, I-(d) refers to degree of differencing involved and MA-(q) refers to order of the moving average part.

The equation for the simplest ARIMA-(p,d,q) model is:

$$[1-\varphi 1(B)-\varphi 2(B)2-\ldots-\varphi p(B)p]*Xt$$
$$=c+[1-\theta 1(B)-\theta 2(B)2-\ldots-\theta p(B)p]*et$$

or in a general form:

$$\varphi 1(B)Xt=c+\theta 1(B)et$$

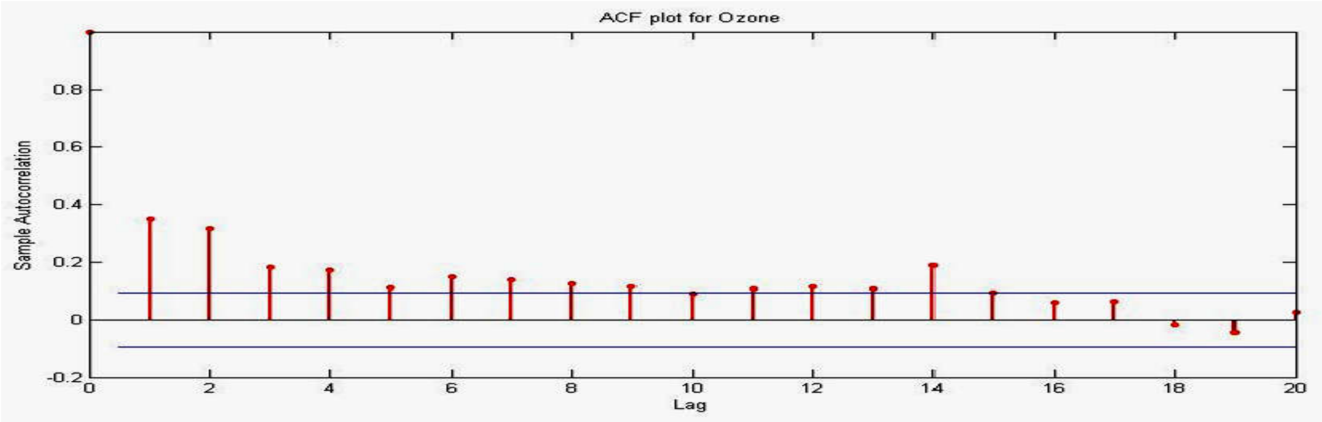
where, φi refers to ith term autoregressive parameter, θi refers to ith term moving average parameter, c means constant, e means error at time t, Bp refers to pth order backward shift operator, and Xt refers to time series value at time t. Four phases are involved in identifying patterns of time series data using ARIMA. They are: model identification, parameter estimation, diagnostic checking and forecasting. The first step is to determine if the time series is stationary and if there is any significant seasonality that needs to be modeled. A total of 446 observations for ozone were examined to check for the most appropriate class of ARIMA processes through selecting the order of the consecutive differencing required to make the series stationary. We identified the stationary component of a data set by performing the Ljung and Box test. The major tools used in the identification phase are plots of the series, correlograms (plots of autocorrelation and partial autocorrelation verses lag) of the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The ACF and the PACF are the most important elements of time series analysis and forecasting. The ACF measures the amount of linear dependence between observations in a time series that are separated by a lag k. The PACF plot helps to determine how many autoregressive terms are necessary to reveal one or more of the following characteristics: time lags where high correlations appear, seasonality of the series, and trend either in the mean level or in the variance of the series.

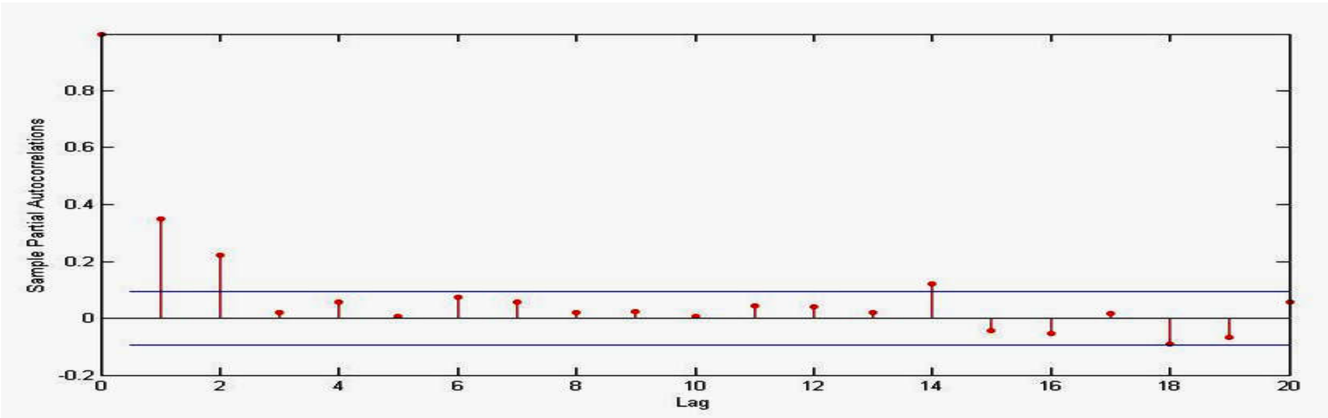
RESULTS AND DISCUSSION

The ARIMA model used for predicting the ozone was (1,0,2). The prediction error can be analyzed using performance evolution methods. The correlogram for the stationary time series will decay soon, if the time series is not stationary it will decay slowly. figure 2 shows the ACF correlogram for a ozone data of the study region. The order of the MA and the AR parameters are then decided based on the ACF and PACF plot of the data. Figure 4 shows the ACF and PACF plot of the data. From the correlogram the best model for ARIMA was chosen and the model with the decided AR and MA parameters were fed in to SPSS software.

A zero value in NRMSE indicates a perfect prediction while a value larger than 1 indicates that the prediction are no better than using the average value of the time series ($\bar{x}$). The absolute error measure, mean absolute error (MAE), is given as:

$$MAE = \frac{\sum_{i=N}^{N} \| x_i - \hat{x}_i \|}{N}$$





Performance evolution

At the forecasting stage, the estimated parameters were tested for their validity using error statistics such as mean square error (MSE), normalized mean square error (NMSE) and mean absolute error (MAE) criteria. The formula used for the evaluation of errors (MSE) and (NMSE) are given below:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2$$

where $\hat{x}_i$ is the predicted value of x_i and $\bar{x}$ is the average value of the time series. The normalized mean square error is given below:

Table 1. Ozone Time Series- comparing the prediction performances of the model

Time series	Prediction method	Prediction Error		
		MAE	MSE	NMSE
Ozone	ARIMA(1,0,2)	0.1754	0.0895	0.1275

$$NMSE = \frac{\sum_{I=1}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

Where y_i is the real value and $\hat{y}_i$ is the predicted value N is the length of the observed data. For the case of perfect prediction, the value of MSE is zero.

Conclusion

The ozone time series data was modeled using the ARIMA model; the best fit model was selected using the XLstat software. The error analysis also shows that the fit model is good. The ARIMA model though very simple it is one of the most popular methods in forecasting the rainfall of region.

REFRENCES

Box, G. E. P., and G. M. Jenkins. 1970. *Time-series analysis: Forecasting and control*. Rev. ed. San Francisco: Holden-Day.

Crutzen, P. J. 1998. How the atmosphere keeps itself clean and how this is affected by human activities?; *Pure Appl.Chem.* 70 1319–1326.

Emberson, L. D., *et al.* 2001. "Impacts of air pollutants on vegetation in developing countries." *Water, Air, and Soil Pollution* 130.1-4, 107-118.

Kalabokas, P. D., Viras, L. G., Bartzis, J. and Repapis, C. C. 2000. Mediterranean rural ozone characteristics around the urban area of Athens; *Atmos. Environ.* 34 5199–5208

Naja, M., and S. Lal. 2002. "Surface ozone and precursor gases at Gadanki (13.5 N, 79.2 E), a tropical rural site in India." *Journal of Geophysical Research: Atmospheres* 107. D14.

Prybutok, R., Junsub, Y., Mitchell, D. 2000. Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations. *Euro. J. Operat. Res.,* 122, 31-40.

Pulikesi, M. Baskaralingam, P. Rayudu, V.N. Elango, D. Ramamurthi, and V. Sivanesan, S. 2006. Surface ozone measurements at urban coastal site Chennai, in India. *J. Hazardous Materials*, B137, 1554–1559.

Regional Meteorological Cetre, Mausam Bhavan, Nungambakkam, Chennai.

Samuel Selvaraj, R., Padma, K., Miltoin Boaz, B. 2013. Seasonal variation of surface ozone and its association with meteorological parameters, UV radiation, rainfall, cloud cover, over Chennai, India. *Current science*, vol. 105, no.5., 676-684.

Urban Factors and The Intensity of Heat Island in the city of Chennai. Monsingh D. Devadas, Lilly Rose A, The seventh International Conference on Urban Climate,une - 3 July 2009, Yokohama, Japan.

****