Holt-Winter Approach To Forecasts Air Quality Index of Industrial Areas of Chennai City In India

Imran Nadeem[1], P.S. Sheik Uduman[2]
[1, 2] Dept of Mathematics
[1, 2] B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai-600048, India

Abstract- *The Air Quality Index (AQI) is a prominent technique for raising public awareness and involvement in initiatives aimed at reducing and managing vehicular traffic in urban areas. In this regard, the Holt-winter approach is employed for forecasting the AQI of the industrial sites of Chennai city. A total of three univariate models have been developed for forecasting the AQI of Kathivakkam, Manali and Thiruvottiyur. The evaluation of the model statistics , RMSE and MAPE values indicates that a significant level of real-time forecasting of the pollutants can be attained by employing the Holt-winter approach. Moreover, the actual monthly AQI of all the sites have been compared with the AQI as prescribed by the central pollution control board of India and are found to be satisfactory for all the sites.*

Keywords- Air pollution, Holt-winter models, Real-time forecasting, Time series analysis

I. INTRODUCTION

Air pollution is a major environmental problem that affects almost every country on the planet. The alarming rate of degradation in air quality over the last few decades has compelled most countries around the world to enact stringent rules and regulations for monitoring and restricting the emission levels of various pollutants. The urban areas, in particular, need more attention due to the continuous increase in industrial and vehicular emissions, which leads to an increase in the number of people affected (Afroz et al. 2003). The Air Quality Index (AQI) is a numerical value that is used to know the air pollution of a particular area. The AQI is a possible air quality indicator for city dwellers that can assist them in determining the quality of their surroundings ambient air. Government institutions use the air quality index (AQI) to aware the public of how contaminated the air is now or may become in the future. The AQI was created by the Central Pollution Control Board in collaboration with IIT-Kanpur and an expert group comprised of medical professionals, air-quality experts, and other stakeholders. The particulate matter PM_{10}, $PM_{2.5}$, Ozone (O_3), Sulphur dioxide (SO_2), Nitrogen dioxide (NO_2), Carbon monoxide (CO), Lead (Pb) and Ammonia (NH_3) are the key pollutants used to calculate the AQI of a region. In India, all eight pollutants are not monitored at all the surveillance centres but the three major pollutants Particulate Matter $(RSPM)$, Sulphur dioxide (SO_2) and Nitrogen dioxide (NO_2) are monitored at all the surveillance centres. Overall AQI is calculated only if data are available for a minimum of three pollutants out of which one should necessarily be either PM2.5 or PM10. Else, data are considered insufficient for calculating AQI.

Several methods have been employed to analyse and forecast time series data, such as autoregressive model, autoregressive moving average model, autoregressive conditional heteroscedasticity model, autoregressive integrated moving average model (ARIMA), Holt Exponential Smoothing Model and so on (Box et al 1994). Exponential smoothing is one of the commonly employed methods for forecasting. In the 1950s (Brown, 1959, Holt, 1957) work paved the way for the development of exponential smoothing forecasting methods. The exponential smoothing methods are one of the efficient methods to forecast from time series. They are generally applied in business to forecast demand for inventories (Gardner, 1985). Three fundamental variations of Simple exponential smoothing (SES) are mostly used: simple exponential smoothing (Brown, 1959); trend-corrected exponential smoothing (Holt, 1957); and Holt–Winters' method (winters, 1960). SES is a basic forecasting technique that unequally weights the observed time series by giving recent observations a higher weight than distant observations. One or more smoothing parameters are used to determine how much weight is assigned to each observation, resulting in unequal weighting. (Li et al 2008). The optimal value of the smoothing constant has a significant impact on the accuracy of the SES process. SES is the simplest technique of this kind, is suitable for a series that goes randomly above and below a continuous mean (stationary series) without having a trend and seasonal patterns (Yorucu, 2003). However, the Holt-winter model, an extension of SES is a widely used method if a series shows trend and seasonality. Its popularity stems from its simplicity, computational performance, ease of adjusting responsiveness to changes in the process being forecasted and fair accuracy (Montgomery, 1990). In this article we compared different exponential smoothing methods practised to forecast the actual AQI like single Exponential Smoothing, double exponential smoothing (Holt-winter), triple exponential smoothing additive and multiplicative and assess the performance.

Data Collection and Study Sites

Chennai is located at 13.0827° N, 80.2707° E and is the capital of the Indian state of Tamil Nadu. Chennai is located along the southern coast of India and has a tropical savanna climate with dry summers and winters (Koppen climate classification). The monitoring sites in Chennai collect Air quality index data twice or thrice a week. Out of the total eight monitoring sites in Chennai, the AQI data of all the three industrial locations of Chennai namely, Kathivakkam, Manali and Thiruvottiyur are collected from January 2017 to December 2020. The collected AQI data is converted into a monthly average and then based upon a suitable Holt-Winters model, the forecasting from January 2021 to December 2021 is done for all three locations.

II. METHODOLOGY

The basic principle of the simple exponential smoothing (SES) model is that the level of time series should vary about a constant rate or change slowly over time (Li et al 2008). The Holt-Winters method, also known as double exponential smoothing, is a trended and seasonal time series extension of exponential smoothing. The Holt-Winters smoothing approach is commonly utilized for forecasting data with seasonality, evolving patterns, and seasonal correlation (Gelper et al 2010).

Simple exponential smoothing is given by: $Y_t = \alpha x_t + (1-\alpha)Y_{t-1}$

Additive Holt Winter is given by

$$Y_t = L_t + B_t + S_t$$

where

$$L_t = \alpha(Y_t - S_{t-m}) + (1-\alpha)(L_{t-1} + B_{t-1})$$
$$B_t = \beta(L_t - L_{t-1}) + (1-\beta)B_{t-1}$$
$$S_t = \gamma(Y_t - L_t) + (1-\gamma)S_{t-m}$$

Multiplicative Holt Winter is given by

$$Y_t = L_t \times B_t \times S_t$$

where

$$L_t = \alpha\left(\frac{Y_t}{S_{t-m}}\right) + (1-\alpha)(L_{t-1} + B_{t-1})$$

$$B_t = \beta(L_t - L_{t-1}) + (1-\beta)B_{t-1}$$

$$S_t = \gamma\left(\frac{Y_t}{L_t}\right) + (1-\gamma)S_{t-m}$$

L_t is level, B_t is the trend and S_t is seasonality.

Statistical metrics to measure forecasting error

Following the specification of models for different sites, the efficiency of the models is evaluated by comparing the forecasting obtained from models with the historical data of time series.

There is no agreement among researchers about which metric is the best for evaluating the best forecasting from a given time series. The key concern in determining the efficiency of a forecast is accuracy, which is the criterion that decides the best forecasting model. The forecaster objective is to reduce error as much as possible (Ryu and Sanchez, 2003). The commonly used metrics to indicate the forecasting of a model are MSE (Mean squared error), RMSE (Root mean squared error), or MAPE (Mean absolute percentage error)

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \times 100$$

The Coefficient of Determination, also known as R-Squared, is a number between 0 and 1 that indicates how well our regression line matches our results. R-Squared value closed to 1 considered to be a good value for forecasting the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\left(Y_i - \overline{Y}\right)^2}$$

Where Y_i represent the actual value, $\hat{Y}_i$ represent the predicted value of Y and $\overline{Y}$ is the Mean value of Y.

III. RESULTS AND DISCUSSION

The values of the Coefficient of determination (R^2), RMSE and MAPE from Table 1 suggests that fitted models have acquired a good degree of forecasting accuracy. The solid line in all three figures indicates the observed AQI data while the dotted line shows the fitted data. In each of the figures, the bold solid line represents the forecasted AQI values from January 2021 to December 2021.

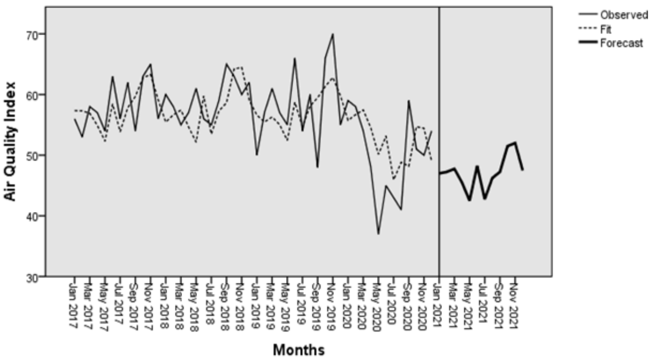


Figure 1: Kathivakkam Winter s' Additive method

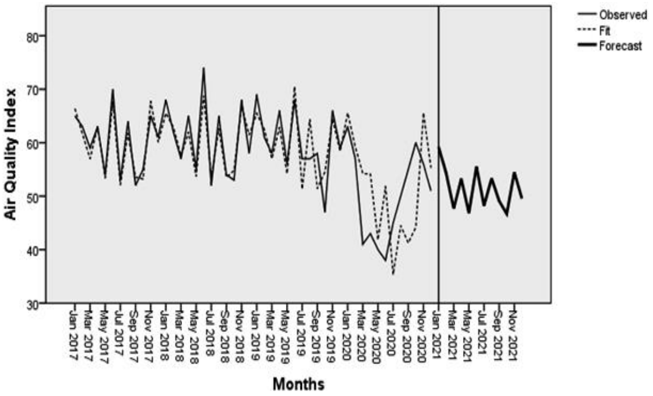


Figure 2:Manila Simple Seasonal method

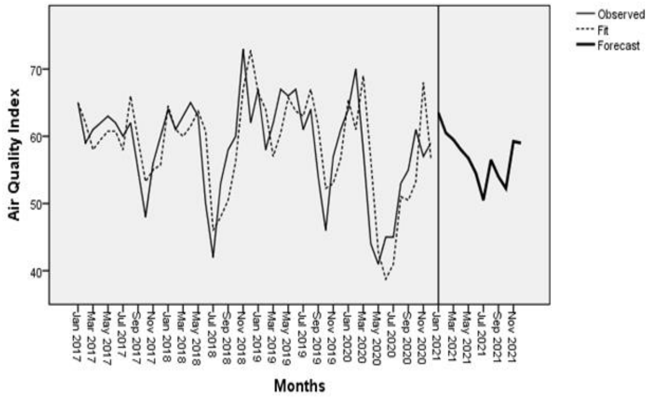


Figure 3: Thiruvottiyur Winters' Multiplicative method

Table 1 based on R^2, RMSE and MAPE values shows the best model among the different version of exponential smoothing models for all three industrial sites.

Table 1: showing the fit model for all the three industrial locations

Locations	Best model	Coefficient of determination (R^2)	RMSE	MAPE
Kathivakkam	Winter s' Additive Holt-winter	86.34	5.23	7.78
Manali	Simple Seasonal Holt-winter	85.55	5.52	7.82
Thiruvottiyur	Winters' Multiplicative Holt-winter	85.20	5.78	7.48

Moreover, the monthly average of AQI for all three sites is compared in Figure 4. As the monthly average of AQI from all the sites during the period of Jan 2017 to Dec 2020 lies between 50 to 100. So the AQI level is satisfactory when compared with the Central pollution control board standard.

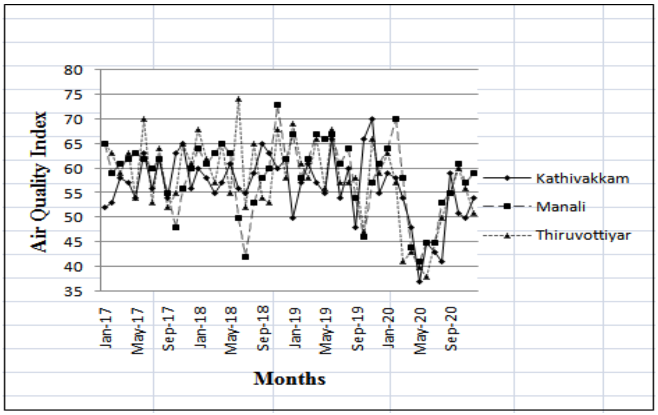
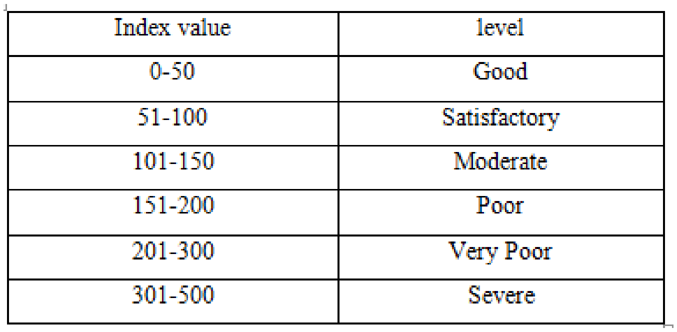


Figure 4: showing the comparison of AQI for all three sites from Jan. 2017 to Dec.2020

All the sites Kathivakkam, Manali and Thiruvottiyur have a minimum monthly average of AQI as 37, 38, and 41 respectively. It is found that the minimum monthly average of AQI for all three sites is during the lockdown period of Covid-19. While the maximum monthly average of AQI from all the sites Kathivakkam, Manali and Thiruvottiyur are 70, 73 and 74 respectively. The monthly average of AQI during the period of Jan 2017 to Dec 2020 for three sites is 56.5, 58.8 and 58.1 respectively. Figure 5 Manali has the highest AQI among all three sites. Further, the AQI of all the three sites is compared with the air quality index defined by the Indian Government to know the level of pollution at all the sites.

Table 2: showing air quality index defined by the Government of India



Index value	level
0-50	Good
51-100	Satisfactory
101-150	Moderate
151-200	Poor
201-300	Very Poor
301-500	Severe

Figure 4 suggests that the monthly average of AQI for all the sites Kathivakkam, Manali and Thiruvottiyur during the period of Jan 2017 to Dec 2020 lies between 50 to 100. This indicates that the AQI level is satisfactory when compared with the Central pollution control board standard of India (Table 2). The AQI level suggests that sensitive people may suffer from minor breathing discomfort.

IV. CONCLUSION

The present study has employed the Holt-winter approach to forecasting the AQI of industrial sites of Chennai city in India with a good degree of forecasting accuracy. A total of three models are selected, one for each for the industrial sites. These models are extremely useful for forecasting air quality since the forecasting from all three models is almost error-free. The level of accuracy indicates that the current forecasting method is quite efficient. The study shows that AQI is satisfactory at all three industrial sites for sensitive people. The actual AQI level on comparing with limit as suggested by Central pollution control board index value and forecasting accuracy of Holt-winter approach of three sites provide an integrative approach for setting a suitable method to deal with the deteriorating level of air quality in industrials areas of Chennai City.

REFERENCES

[1] Afroz, R. Hassan, MN. Ibrahim, NA. Review of air pollution and health impacts in Malaysia. Environmental Research 2003; 92(2): 71-77.

[2] Box, G.E.P., Jenkins, G.M., and Reinsel, G.C.(1994), Time Series Analysis: Forecasting and Control, 3rd edition, Prentice-Hall: Englewood Cliffs, New Jersey.

[3] Brown, R. G. (1959). Statistical forecasting for inventory control. New York7 McGraw Hill.

[4] Gardner, E. S. (1985). Exponential smoothing: The state of the art. Journal of Forecasting, 4, 1– 28.

[5] Holt, C. C. (1957). Forecasting Trends and Seasonal by Exponentially Weighted Averages, ONR Memorandum No. 52, Carnegie Institute of Technology, Pittsburgh, USA (published in International Journal of Forecasting 2004, 20, 5–13).

[6] Li, Z. P. Yu, H. Liu, Y. C. Liu, F. Q., An Improved Adaptive Exponential Smoothing Model for Short Term Travel Time Forecasting of Urban Arterial Street, Acta automatica sinica, Vol. 34, No. 11, 1404–1409, 2008.

[7] Gelper, S. Fried, R. Croux, CH., Robust Forecasting with Exponential and Holt-Winters Smoothing, Journal of forecasting, Vol. 29, No. 3,285–300, 2010.

[8] Montgomery, D. C., Johnson, L. A., Gardiner, J. S., Forecasting and Time Series Analysis, McGraw-Hill, Inc., 1990, ISBN 0-07-042858-1.

[9] Ryu, K., Sanchez, A., The Evaluation of Forecasting Methods at an Institutional Foodservice Dining Facility, The Journal of Hospitality Financial Management, Vol. 11, No. 1, 2003.

[10] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management Science, 6, 324–342.

[11] Yorucu, V., The Analysis of Forecasting Performance by Using Time Series Data for Two Mediterranean Island, Review of Social, Economic & Business Studies, Vol. 2, 175–196, 2003.

[12] Özgür, C. and Tosun, E., 2017. Prediction of density and kinematic viscosity of biodiesel by artificial neural networks. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 39(10), pp.985-991.