STATISTICAL TREND AND FORECAST MODELING OF AIR POLLUTANTS IN BHUBANESWAR

Manoranjan_Sen[1], Gowthaman_T[2], Kunja_Bihari_Panda[3], Paramananda_Pradhan[4]

[1]Dept. of Agril. Statistics, C.A, OUAT, Bhubaneswar, Odisha-751003, India, email:manoranjanstat@gmail.com
[2]Dept. of Agril. Statistics, C.A, OUAT, Bhubaneswar, Odisha-751003, India, email: agrigowtham77@gmail.com
[3]Department of Statistics, Utkal University, Bhubaneswar, Odisha-751004, India, email: Kbpanda62@yahoo.co.in
[4]Dept. of Agril. Statistics, C.A, OUAT, Bhubaneswar, Odisha-751003, India, email: Pradhan1pnp@gmail.com

Abstract. *In this study, attention has been paid on the statistical analysis of trends and developing models for forecasting of the atmospheric pollutants and meteorological parameters of Bhubaneswar, Odisha. Further, using autoregressive integrated moving average model, future values of air pollutants levels are predicted. The monthly data on seven parameters SO_2, NO_2, O_3, $PM_{2.5}$, PM_{10}, average temperature and relative humidity, for the period 2005 to 2019 were collected from AQI station. Mann-Kendall and Sen's slope estimator tests are used for non- seasonal data for the period 2005 – 2019, trend results and power of the slopes are estimated. Atmospheric pollutants, $PM_{2.5}$ and meteorological parameter, average temperature show non-significant increasing trend, whereas other parameters such as SO_2, NO_2, O_3, PM_{10}, and relative humidity shows an increasing trend over the past few years. For this analysis, different autoregressive integrated moving average models are compared with goodness of fit statistics. ARIMA (2, 1, 2) is found as the best-suited model for forecasting of different pollutants in Bhubaneshwar. PM_{10} and O_3 show a rising trend with predicted approximate annual concentration of 92.530$\mu g/m^3$ and 26.46 $\mu g/m^3$; PM 2.5, SO_2 and NO_2 show a reducing trend with an approximate annual concentration of 40.93$\mu g/m^3$, 2.21$\mu g/m^3$ and 38.63$\mu g/m^3$, respectively, by the year 2025. The meteorological parameters that are average temperature and relative humidity exhibit a rising trend in the annual concentration of Bhubaneshwar. Hence, ARIMA (2, 1, 2) is a suitable model for forecasting the atmospheric pollutants and meteorological parameters of air pollution of Bhubaneshwar, for future planning and policy-making to reduce the pollution of the city.*

Keywords: *Air pollutants; Forecast; Mann-Kendall; Sen's slope estimator; Autoregressive integrated moving average (ARIMA); Prediction.*

1. INTRODUCTION

Air pollution has become an important factor in environmental degradation. Air pollution is a total of pollutants that freely exit in the air, and on coming in contact with human beings and plants, can cause harmful effects (Sharma *et al.*, 2018). Urban and peri-urban air pollution is one of the important environmental concerns throughout the world. Atmospheric particulates and gaseous pollutants, pose severe health effects both for humans and plant species (Gupta and Kulshrestha, 2016; Maatoug, 2010).Monitoring for urban AQI mostly comprises trackingof pollutants like carbon monoxide (CO), particulate matters (PM), Ozone (O3), lead (Pb)particles, Sulphurdioxide (SO2), nitrogen dioxide (NO2) and ammonia(NH3) (Azmi *et al.*, 2010; Gurjar*et al.*, 2008).With available pollutants data under AQI, trend analysis and forecasting are possible through various statistical modeling techniques. Can(2017) used graphical and statistical approaches for time-series analysis of air pollutants, Rani *et al.* (2018) used past air pollution index data for trend analysis using XLSTAT Trend estimation highly depends upon the characteristics of data andthus are considered as a complex approach (Kisi and Ay, 2013). In the proposed study, non-parametrictests are applied for statistical analysis. Parametric approaches are considered more precise than nonparametric tests but come with a limitation of normally distributed independent data whereas non-parametric tests have no such constraints (Watthanacheewakul, 2011). The proposed study uses the non-parametric Mann-Kendall (M-K) test in addition with Sen's-slope estimator approach for trend estimations of different pollutants and autoregressive integrated moving average (ARIMA) approach for modeling the pollutants forecast. M-K and Sen's - slope estimator tests are well-established tests for estimating the rising or reducing trends for the non-parametric data. ARIMA modeling is a generalized approach in which the models are fit on the time-series data to predict the future values (Brocklebank et al., 2018; Eymen and Köylü, 2018).The proposed study first used M-K test along with Sen's slope estimator tests to assess trend existence in the pollutants time series data and afterward ARIMA modeling is done to forecast the pollutants value with precision. Different pollutants considered in the study of statistical analysis are PM 2.5, and PM 10, O_3,NO2, and SO2.Mann-Kendall and Sen's slope estimator tests are used for non- seasonal data and ARIMA modeling is done for the yearly forecast of the pollutants consider. This study has been carried out for the city Bhubaneshwar, India in 2020 and considers the past data of year 2005 to 2019 for statistical trend assessment of air pollution.

2. MATERIALS AND METHODS

The past pollutants data for the study is retrieved from State Pollution Control Board (SPCB), Bhubaneshwar, India. The air pollutants which are selected in this study are the PM particles PM10 and PM 2.5, O_3, SO_2 and NO_2. The pollutants data retrieved for the study is from January 2005 to December 2019.The meteorological parameters (average temperature and relative humidity) are collected from the India Meteorological Department (IMD). Excel-XLSTAT version 2019.1 is used as statistical software for M-K test, Sen's slope estimator, and ARIMA modeling.

2.1. Mann-Kendall test

The Mann–Kendall (MK) test is a non-parametric trend analysis for identifying the increasing and decreasing pattern in time series of the data. It compares the relative magnitudes of sample data rather than the data values themselves (Gilbert 1987). The MK test is first implemented using the null hypothesis (H_0)of no trend testing, that is, the observations Xi are randomly ordered in time, against the alternative hypothesis (H_1), where there is an increasing or decreasing monotonic trend. The data values evaluated as ordered time series are compared with all subsequent data values. If a data value from a later period is higher than a data value from an earlier period, the statistic S is incremented by 1. On the other hand, if the data value from a later period is lower than a data value sampled earlier, S is decremented by 1.

The net result of all these increments and decrements yields the final value of S (Shahid 2011; Shrestha *et al.*, 1999; Yue *et al.*, 2002; Domonkos*et al*., 2003). The MK test statistic S is computed as:

$$S = \sum_{k=1}^{n=1} \sum_{j=k+1}^{n} sgn\,(x_j - x_k) \tag{1}$$

$$Sgn(x_j - x_k) = \begin{cases} -1 & if\ (x_j - x_k) < 0 \\ 0 & if\ (x_j - x_k) = 0 \\ +1 & if\ (x_j - x_k) > 0 \end{cases} \tag{2}$$

where x_j and x_k are the annual values in different years jand k, j>k, respectively. If n<10 then the value of |S| is compared directly with the theoretical distribution of S that is derived by the Mann–Kendall test (Gilbert 1987). The two-tailed test issued. At some probability level, H_0 is rejected in favour of H_1 if the absolute value of S equals or exceeds a specified value Sα/2, where Sα/2 is the smallest Shaving the probability less than α/2. A positive (negative) value of S indicates an upward (downward) trend (Salmi *et al*., 2002; Luo *et at.,* 2008).

For n≥10, the statistic S is approximately normally distributed with the mean and variance as follows:

$$E(S) = 0$$

$$Var\ (S) = \frac{1}{18}[n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t_p - 1)\ (2t_p + 2)\ ] \tag{3}$$

q is the number of tied groups and t_p is the number of data values in the p^{th} group. The standard test statistic Z is computed as:

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(s)}} & ; if\ S > 0 \\ 0 & ; if\ S = 0 \\ \frac{S+1}{\sqrt{Var(s)}} & ; if\ S < 0 \end{cases} \tag{4}$$

The presence of a statistically significant trend is evaluated using the Z value. A positive (negative) value of Z indicates an upward (downward) trend. To test for either an upward or downward monotonic trend (a two-tailed test) at α level of significance, H0 is rejected if $|Z| > Z_{1-\frac{\alpha}{2}}$, where $Z_{1-\frac{\alpha}{2}}$ is obtained from the standard normal cumulative distribution tables. The Kendall's τ values are calculated as Eq. 5.

$$\tau = 2\frac{S^*}{z(z-1)} \tag{5}$$

In which S* denotes Kendall's sum, computed as S* = A-B where A represents the number of chances when the difference of x_b to x_a is greater than zero and B represents the number of chances when the difference of x_b to x_a is less than zero (Chattopadhyay *et al.*, 2012; Xu *et al.*, 2004).

2.2, Sen's slope estimator test

This test also termed as Theil–Sen slope test is a widely used statistical tool for non-parametric data to estimate the power of trend, detected through the M-Ktest (Caloiero et al., 2017; Eymen and Köylü, 2018). Developed by Theil, 1950, and Sen, 1968, this is a median-based tool that evaluates the slope of the trend through a linear model. If there are m number of pollutant data points in a time-series (X_1, X_2, X_3……X_m) and X_a and X_b are the pollutant values at time instance a and b such that b > a, then variance of the residual is computed as Eq. 6 and 7.

$$T_i = \frac{x_b - x_a}{b-a}\ for\ i = 1,2,3, \dots \dots \dots m \tag{6}$$

The median of all $T_i$values, denoted as T_{med}, is the Sen's slope estimator and is calculated as equation 7. The sign of T_{med} reveals the upward or down ward trend of the data and its numeral denote the trend steepness.

$$T_{med} = \begin{cases} T_{\frac{m+1}{2}} & if\ m = odd \\ \frac{T_{\frac{m}{2}} + T_{\frac{(m+2)}{2}}}{2} & if\ m = even \end{cases} \tag{7}$$

The trend prediction of the pollutants through the M-K test depends upon the significance level α, and there is a possibility of the existence of trends with other significant levels. So through Sen's slope estimator, the changing rates can be assessed for the pollutants which shows no trend in M-K Test.

2.3. Autoregressive integrated moving average (ARIMA)

The ARIMA model, developed for prediction and estimation of future values in univariate time-series data, was introduced by Box and Jenkins (1976). ARIMA includes a combination of several timeseries techniques to give a better representation and analysis of time-series data. Auto regression (AR), differencing order integration (I) and moving average (MA) collectively makes ARIMA (p, d, q) model in which p is the order of auto regression model, d is for differencing order integration, and q is the moving average model order. In the first step of multimethodology, time-series data are checked whether itis stationary or not. Dickey-Fuller (D-F) test is used in the paper to check the data (Dickey and Fuller,1979). If the data is stationary, the model moves in the second step else the data is made stationary by difference. In the next step, p, d, q possible values are estimated using correlogram of autocorrelation and partial autocorrelation functions (ACF and PACF). In next stage, for determining the adequacy of the model, the values of Akaike information criteria (AIC), and other error estimation measures are assessed over the best-suited goodness of fit statistics to select appropriate ARIMA model order. For the idea of order determination of the ARIMA model in the provided study, various goodness of fit statistics criteria observed which, other than AIC, includes sum of squared errors (SST), root mean squared deviation (RMSD), W-N Variance, mean absolute percentage deviation (MAPD) and final prediction error (FPE). With the chosen model, the last step involves estimation of forecasted values for the provided time-series data. A generalized expression of ARIMA(p,d,q) can be given as Eq. 9.

$$\phi(\beta)\nabla^d f_t = \theta(\beta)e_t \quad (8)$$

Where, $\phi(\beta)$and $\Theta(\beta)$ represent the polynomial of degree p and q respectively, β is a backward shift operator, ∇ is difference operator, ft is pollutants parameter at time t and et is the error term at time t.

3. RESULTS AND DISCUSSIONS

In this section of study, results estimation and analysis of its inferences are carried out for the Mann-Kendall test, Sen's slope estimator test and, ARIMA modeling of time-series pollutants data of AQI sampling station of Bhubaneshwar. The analysis in annual scale indicated that the SO_2 has a decreasing trend, the p-value (0.014) is below the significance level, with a negative value of Kendall's tau (-0.144), so the null hypothesis is rejected confirming the alternate hypothesis of acceptance of trend in the time series data (Table 1; Fig 1). The trend of SO_2 on a monthly scale, on the other hand, shows an increasing trend from November and December (Table 2). The MK trend analysis of nitrogen dioxide (NO_2) shows an increasing trend, with a positive value of Kendall's tau (0.506), the p-value is below the significance level, so the null hypothesis is rejected then accepted the alternative hypothesis for the time series data (Table 1; Fig 2). It is interesting to note that the Nitrogen dioxide (NO_2) shows a significant increasing trend in the monthly scale as well (Table 2). The tropospheric ozone (O_3) is recognized as one of the major air pollutants affecting the climate of the earth. The result reveals that the tropospheric O_3 show an increasing trend over Bhubaneshwar with positive Kendall's value 0.236 (Table 1; Fig. 3). The trend of Tropospheric Ozone (O_3) shows an increasing pattern in September and November (Table 2).

The particulates of size 2.5-µm diameter ($PM_{2.5}$) shows no trend over provided years, as p-values are 0.194, that is more to 0.05, H_o is accepted (Table 1; Fig. 5). In the case of a monthly scale, an increasing trend is observed in February and November. The particle of size 10-µm diameter (PM_{10}) shows an increasing trend with Kendall's tau value 0.254, the p-value is below the significance level (Table 1; Fig.5). The monthly trend of PM_{10} concentration, on the other hand, shows an increasing trend from January to June and September, November (Table 2).

Climate change and its inter-annual and intra-seasonal variability are the major global concern in recent times. The temperature plays a deciding role in understanding the climate change brought about by urbanization and industrialization. The MK trend of average surface temperature shows no trend over Bhubaneshwar from 2005 to 2019 with Kendall's tau value 0.085, the p-value is above 0.05, H_o is accepted (Table 1; Fig. 6). However, the monthly trend of average temperature shows an increasing pattern from June to August and no trend observed in the rest of the months (table 2). The MK trend of average relative humidity indicates an increasing trend over Bhubaneshwar with Kendall's Tau value 0.112 (Table 1; Fig. 7). The MK trend in the monthly scale depicts a decreasing nature in January and no trend is found in the rest of the months (Table 2).

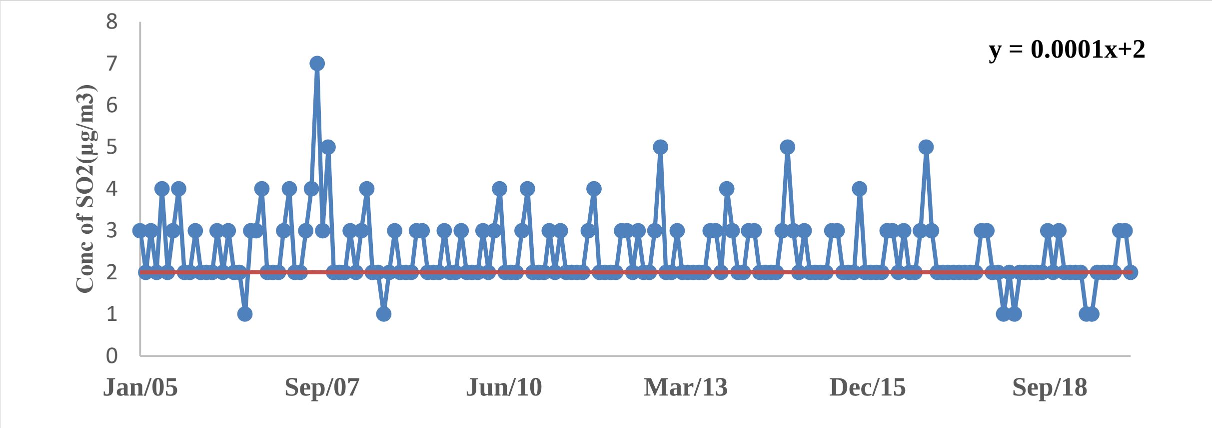
Table 1: The result of the Mann–Kendall trend test for atmospheric pollutants and meteorological parameters over Bhubaneshwar from 2005 to 2019

Pollutants	Kendall's tau (τ)	p-value*	S value	Result	Sen's slope T_{med}
SO_2(µg/m³)	-0.144	0.014*	-1731.00	Trend exits	-0.011
NO_2(µg/m³)	0.506	<0.0001*	7934.00	Trend exits	8.061
O_3(µg/m³)	0.236	0.006*	563.00	Trend exits	0.013
PM 2.5 (µg/m³)	0.066	0.194	1050.00	No trend	0.004
PM 10 (µg/m³)	0.254	<0.0001*	4076.00	Trend exits	4.698
Relative Humidity-%	0.112	0.029*	1766.00	Trend exits	0.107
Average temperature-^{0}C	0.085	0.099	1333.00	No trend	0.002

*p value<0.05 indicates rejection of the null hypothesis of no trend and thus, revealing the existence of the trend

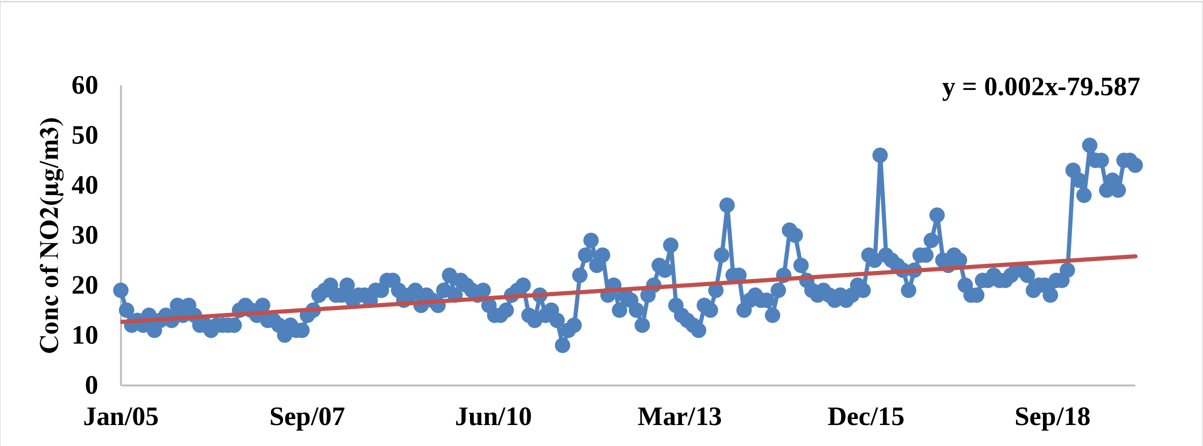
Source: Realised by authors

The values of Sen's slope estimator T_{med} is also provided in Table 4.2, and the outcomes of Sen's slope test validates the M-K test results. T_{med} values are also calculated for those pollutants data in which no trend exists. This is for the reason that the hypothesis in the M-K test is established over a significant level α and there is a possibility of the presence of a trend, and thus the trend slope possibility, beyond α, In the proposed study, α is kept at 5% for the results. The Sen's slope results presented in Table 4.2 confirms the results of M-K test and shows the similar slope orientations. The T_{med}value of SO_2 shows a negative slope for the trend (-0.011). M-K test for the $PM_{2.5}$and Average temperature shows no trend, and the Sen's slope estimator values predicted a positive slope with value 0.004 and 0.002, respectively. The T_{med}value of NO_2, O_3 and PM_{10} shows a positive slope confirming an increasing trend in the data of the air pollutants. The T_{med} value for relative humidity shows a positive slope for the trend (0.107).



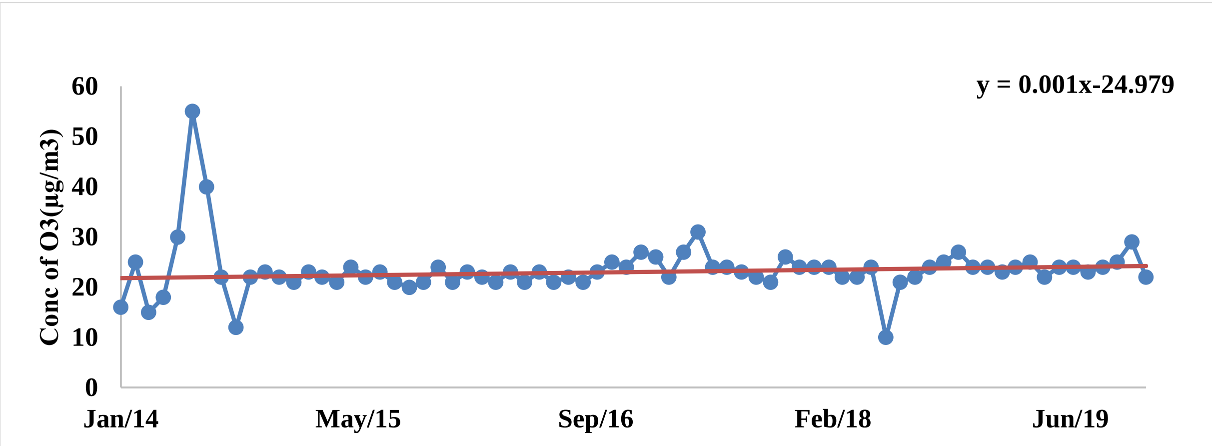
Source: Realised by authors

Fig. 1 Trends of SO_2 concentrations (µg/m^3) over Bhubaneshwar during the period from 2005 to 2019



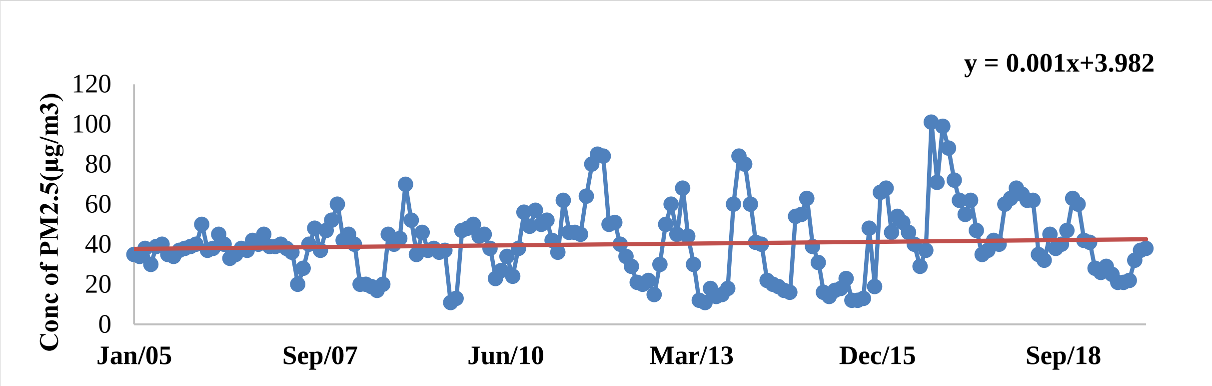
Source: Realised by authors

Fig. 2 Trends of NO_2 concentrations (µg/m^3) over Bhubaneshwar during the period from 2005 to 2019



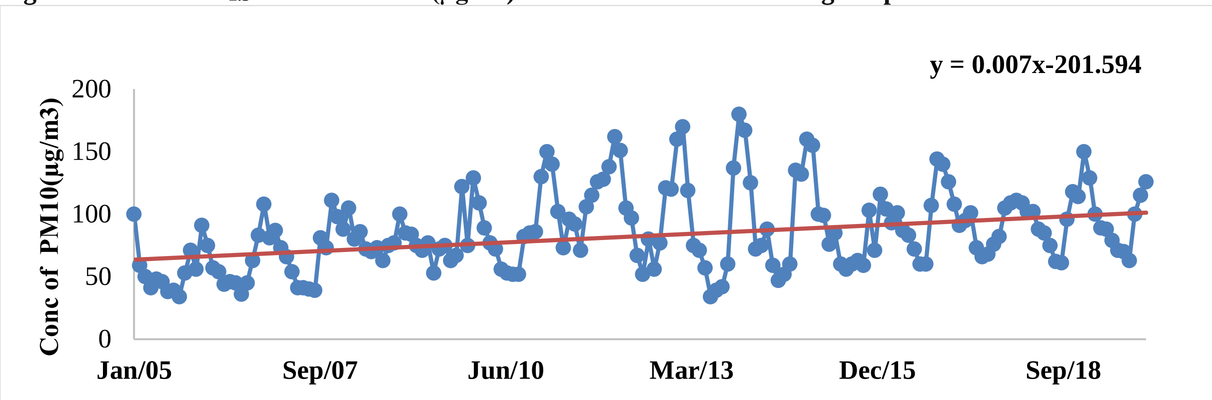
Source: Realised by authors

Fig. 3 Trends of O_3 concentrations (µg/m^3) over Bhubaneshwar during the period from 2005 to 2019



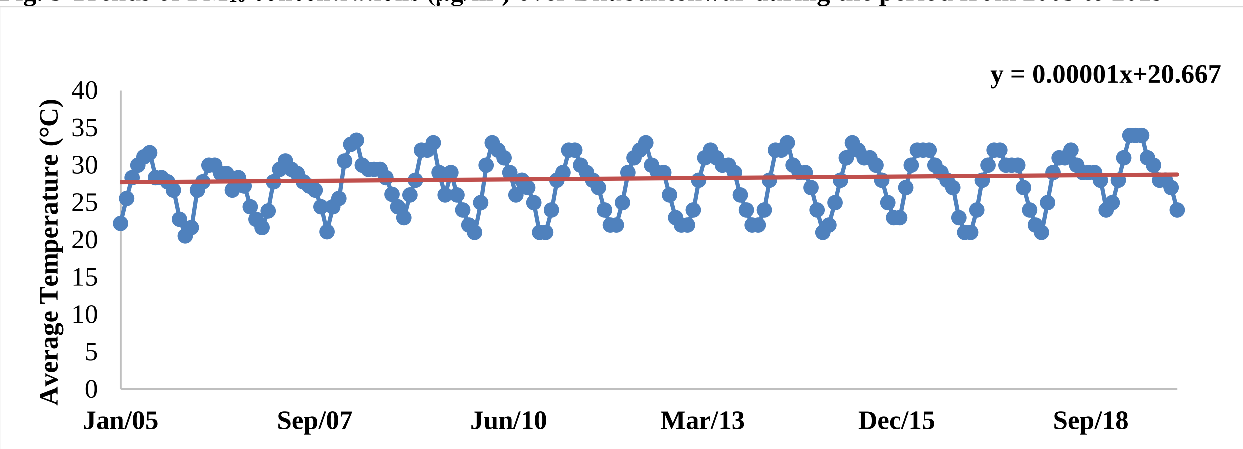
Source: Realised by authors

Fig. 4 Trends of $PM_{2.5}$ concentrations (µg/m^3) over Bhubaneshwar during the period from2005 to 2019



Source: Realised by authors

Fig. 5 Trends of PM_{10} concentrations (µg/m^3) over Bhubaneshwar during the period from 2005 to 2019



Source: Realised by authors

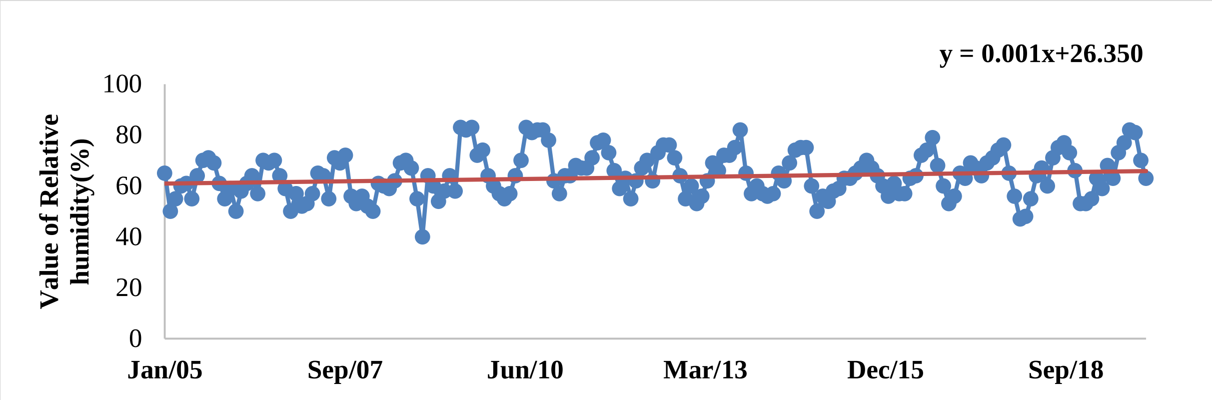
Fig. 6 Trends of average temperature (^{0}c) over Bhubaneshwar during the period from 2005 to 2019

Table 2: Result of the Mann-Kendall trend test for pollutants&meteorological parameters (monthly scale)

Oct	Nov	Dec
0.267	0.476	0.635
0.248 21.00 No trend	0.026 42.00 Trend exits	0.006 46.00 Trend exits
0.604	0.572	0.560
0.003 61.00 Trend exits	0.004 58.00 Trend exits	0.005 58.00 Trend exits
0.501	1.000	0.234
0.243 7.00 No trend	0.009 15.00 Trend exits	0.681 3.00 No trend
0.038	0.425	0.271
0.882 4.00 No trend	0.033 44.00 Trend exits	0.180 28.00 No trend
0.364	0.429	0.314
0.067 38.00 No trend	0.029 45.00 Trend exits	0.113 33.00 No trend
0.267	0.163	0.235
0.203 26.00 No trend	0.447 16.00 No trend	0.263 23.00 No trend
0.320	0.243	0.117
0.112 33.00 No trend	0.233 25.00 No trend	0.584 12.00 No trend

Pollutants	Trend tests	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep
SO_2(µg/m^3)	Kendall's τ value	-0.153	0.052	-0.678	0.307	-0.064	-0.127	0.339	-0.282	-0.116
	P-value*	0.552	0.908	0.003	0.179	0.817	0.616	0.148	0.211	0.632
	S value	-8.00	2.00	-52.00	25.00	-5.00	-8.00	26.00	-22.00	-9.00
	Result	No trend	No trend	Trend exits	No trend	No trend	No trend	No trend	No trend	No trend
NO_2(µg/m^3)	Kendall's τ value	0.638	0.580	0.625	0.657	0.579	0.612	0.525	0.586	0.566
	P-value*	0.001	0.003	0.002	0.001	0.004	0.002	0.009	0.003	0.005
	S value	65.00	60.00	64.00	66.00	59.00	63.00	53.00	60.00	58.00
	Result	Trend exits	Trend exits	Trend exits	Trend exits	Trend exits	Trend exits	Trend exits	Trend exits	Trend exits
O_3(µg/m^3)	Kendall's τ value	0.690	-0.138	0.745	0.600	-0.276	-0.138	-0.215	0.545	0.966
	P-value*	0.08	0.848	0.070	0.133	0.566	0.848	0.697	0.217	0.013
	S value	10.00	-2.00	10.00	9.00	-4.00	-2.00	-3.00	7.00	14.00
	Result	No trend	No trend	No trend	No trend	No trend	No trend	No trend	No trend	Trend exits
$PM_{2.5}$(µg/m^3)	Kendall's τ value	0.237	0.471	0.243	0.058	-0.087	-0.010	0.010	-0.174	-0.010
	P-value*	0.249	0.017	0.233	0.804	0.691	1.000	1.000	0.398	1.000
	S value	24.00	49.00	25.00	6.00	-9.00	-1.00	1.000	-18.00	-1.000
	Result	No trend	Trend exits	No trend	No trend	No trend	No trend	No trend	No trend	No trend
PM_{10} (µg/m$^{3)}$	Kendall's τ value	0.390	0.478	0.333	0.510	0.580	0.574	0.364	0.290	0.414
	P-value*	0.048	0.015	0.092	0.010	0.003	0.003	0.067	0.150	0.040
	S value	41.00	50.00	35.00	53.00	60.00	60.00	38.00	30.00	42.00
	Result	Trend exits	Trend exits	No trend	Trend exits	Trend exits	Trend exits	No trend	No trend	Trend exits
Average Temperature(^{0}C)	Kendall's τ value	-0.010	0.001	0.262	0.213	0.323	0.481	0.770	0.460	0.270
	P-value*	1.00	1.00	0.216	0.313	0.131	0.019	<0.001	0.025	0.201
	S value	-1.00	O.01	25.00	21.00	29.00	47.00	71.00	45.00	26.00
	Result	No trend	No trend	No trend	No trend	No trend	Trend exits	Trend exits	Trend exits	No trend
Relative humidity (%)	Kendall's τ value	-0.531	0.301	0.363	0.089	0.347	0.234	0.176	0.126	0.215
	P-value*	0.009	0.136	0.072	0.688	0.089	0.252	0.397	0.551	0.295
	S value	-53.00	31.00	37.00	9.00	35.00	24.00	18.00	13.00	22.00
	Result	Trend exits	No trend	No trend	No trend	No trend	No trend	No trend	No trend	No trend

*p value<0.05 indicates rejection of the null hypothesis of no trend and thus, revealing the existence of the trend



Source: Realised by authors

Fig. 7 Trends of relative humidity (%) over Bhubaneshwar during the period from 2005 to 2019

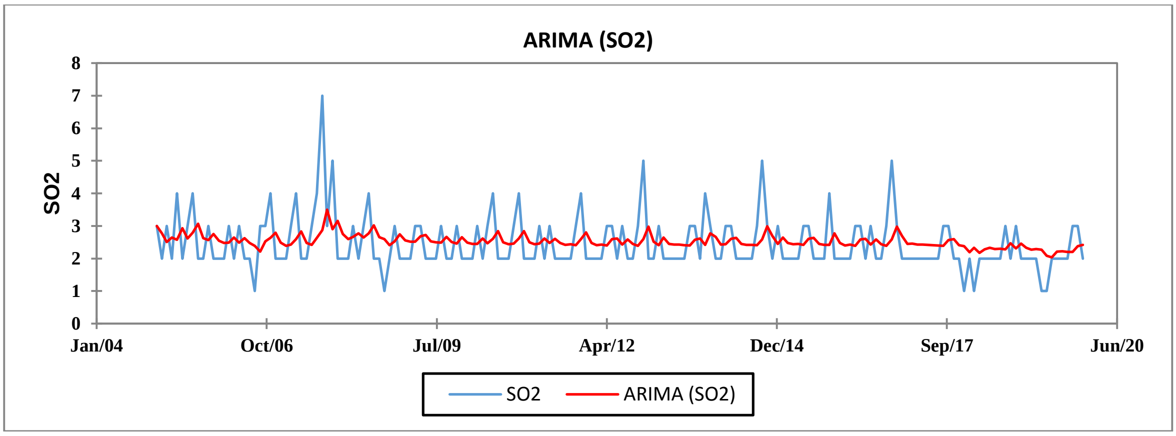
The statistically significant relationship between the meteorological parameters and pollutants is observed for Bhubaneshwar where highlighted correlations are significant at $p<0.05000$. PM_{10}, SO_2, NO_2, and O_3 are positively correlated with each other, whereas NO_2, SO_2, $PM_{2.5}$, $PM_{10}$are negatively correlated with temperature. $PM_{2.5}$ and PM_{10} are highly correlated with each other (Table 3).

Table 3: Correlation matrix of all pollutants and meteorological parameters over Bhubaneshwar (Marked correlations are significant at p<0.05000)

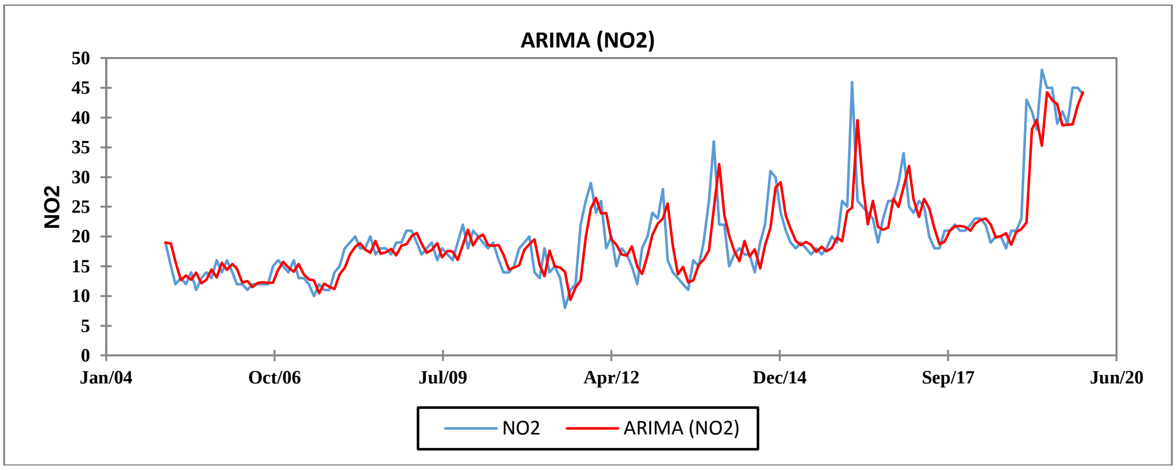
	SO_2	NO_2	PM2.5	PM_{10}	Temperature	Humidity	O_3
SO_2	**1**						
NO_2	**-0.13477**	**1**					
PM2.5	**0.125816**	**0.224156**	**1**				
PM_{10}	**-0.00231**	**0.492156**	**0.616585**	**1**			
Temperature	**-0.06602**	**-0.07757**	**-0.52222**	**-0.51422**	**1**		
Humidity	**0.076917**	**-0.07267**	**-0.30742**	**-0.36213**	**0.28607**	**1**	
O_3	**-0.16191**	**0.529737**	**0.086872**	**0.180262**	**0.124029**	**0.007094**	**1**

Source: Realised by authors

ARIMA time-series model is fitted on the available pollutant data. Using the D-F test, the data is checked with null hypothesis H0 which shows the existence of unit rooting time-series data and alternative hypothesis H1 which shows no unit root confirming a stationary time-series data. Computed p-value in D-F test for the time-series data of each of the pollutants in different season comes lower to a significance level of 0.05, due to which H0 is rejected, and data confirms to be stationary and suitable for applying ARIMA (p, d, q) model. Three models are selected for the purpose; ARIMA (2, 1, 1), ARIMA (2, 1, 2) and ARIMA (2, 0, 2) at 95% level of confidence. All three models are selected and checked over the goodness of fit statistics for choosing the best-suited model. Table 4 gives comparisons of the goodness of fit statistics for different pollutant time-series data. From the results of Table 4 and statistical analysis of the residual plots, the ARIMA (2, 1, 2) model has the least error estimation values in the goodness of fit statistics and thus appears to be best suited for the forecasts of the pollutants value of Bhubaneshwar. Figs. 8, 9, 10, 11, 12 ,13, 14, shows the actual pollutant values and the forecasted pollutant values with ARIMA (2, 1, 2) model at a 95% confidence interval for the year 2005 to 2019. In Figures, ARIMA (2, 1, 2) plot is compared with the original data points of different pollutions where the blue line represents the monthly mean concentrations of observed pollutants value and the red line represents the ARIMA (2, 1, 2) model values. ARIMA (2, 1, 2) model can be used for predicting and forecasting future pollutants values which can aid the decision-makers for planning steps to mitigate the pollutants which shows a rise in trend and are above the permissible standard limits.From the result of Table 5 shows that there is a reducing trend in the annual concentration of SO_2, NO_2 and $PM_{2.5}$, whereas O_3 and PM_{10} shows a rising trend in the annual concentrations of Bhubaneshwar. Also, Table 4.6 illustrates much higher concentrations of PM 2.5 and PM_{10}, to that of the permissible standard annual concentration limits, that is, 40 $\mu g/m^3$of $PM_{2.5}$ and $60\mu g/m^3$ of PM_{10}. $SO_2$and $NO_2$are through under the permissible standard annual concentration limits $50\mu g/m^3$and $40\mu g/m^3$ respectively and are showing a decreasing trend, but still are very close to the permissible limits and the results from Fig.3 shows that NO_2 have frequently crossed the acceptable limits during the year 2005 to 2019. Average Temperature and Relative Humidity shows a rising trend in the annual concentration of Bhubaneshwar.

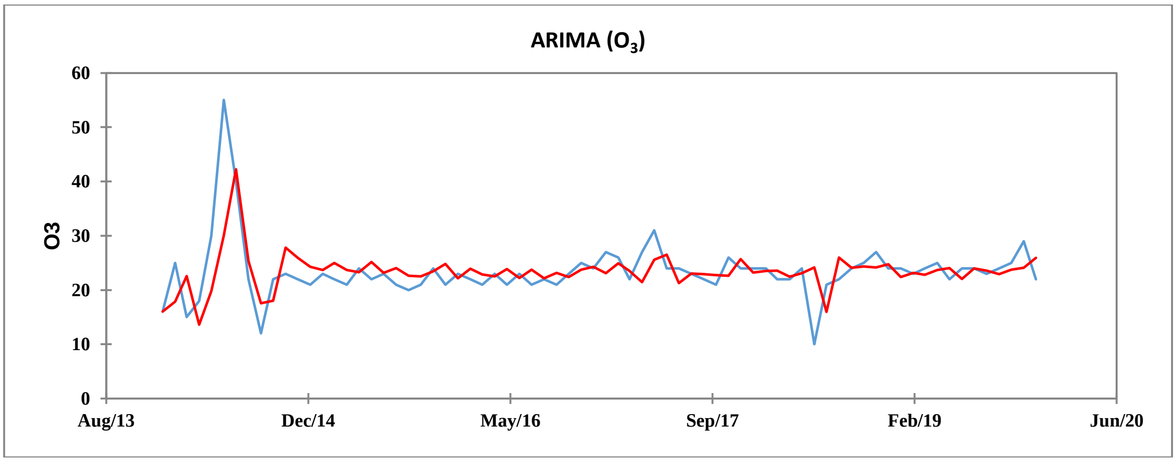


Source: Realised by authors



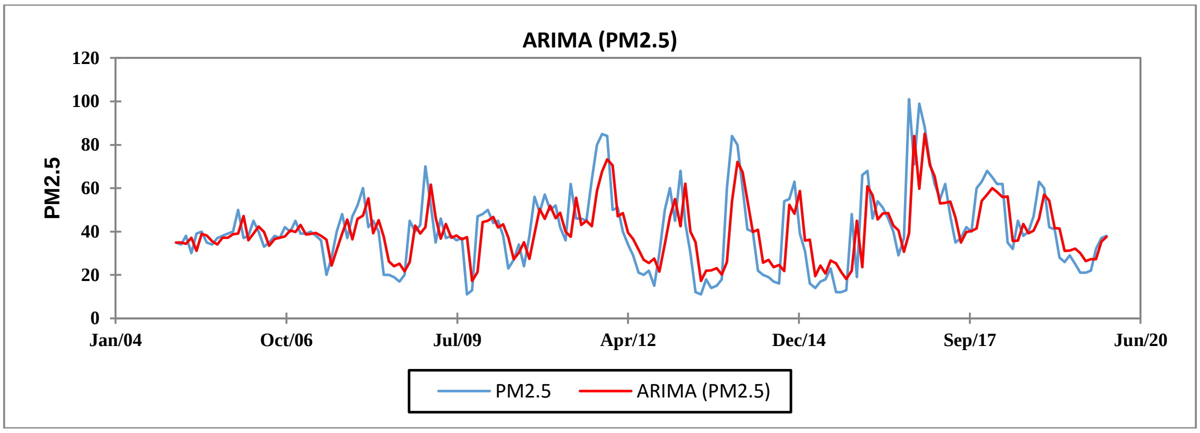
Source: Realised by authors

Fig.9 The diagram shows the observed and ARIMA (2,1,2) model forecast of NO_2 concentration($\mu g/m^3$)



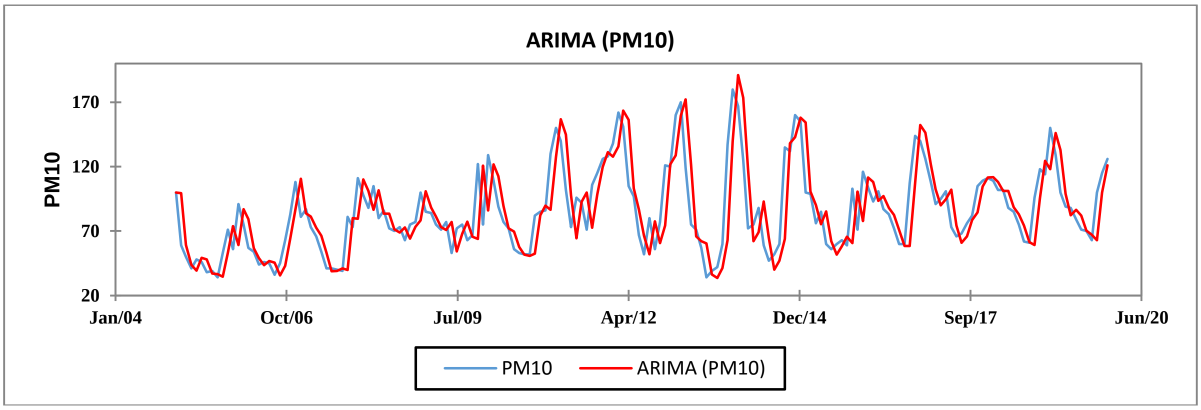
Source: Realised by authors

Fig.10 The diagram shows the observed and ARIMA (2,1,2) model forecast of O_3 concentration($\mu g/m^3$)



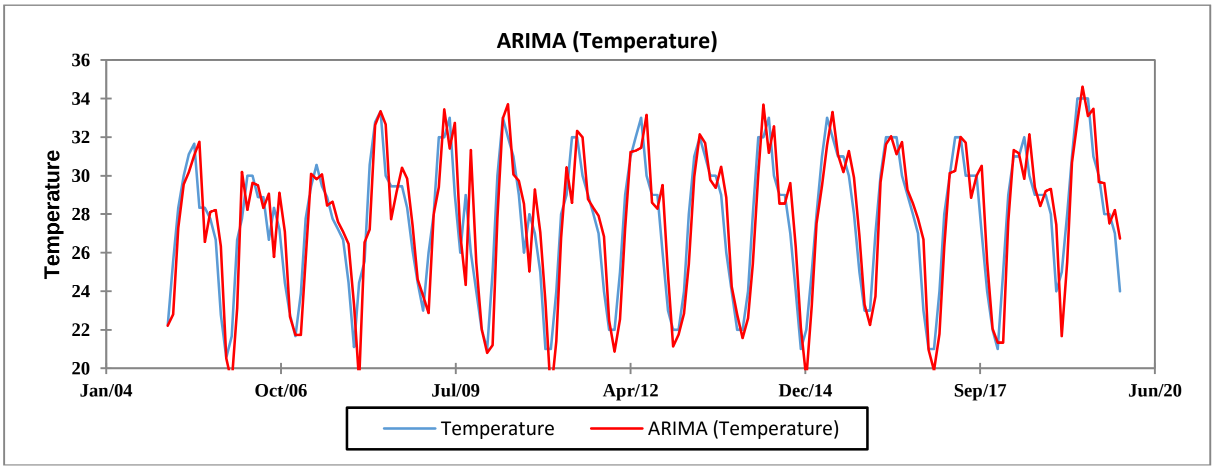
Source: Realised by authors

Fig.11 The diagram shows the observed and ARIMA (2,1,2) model forecast of $PM_{2.5}$ concentration($\mu g/m^3$)



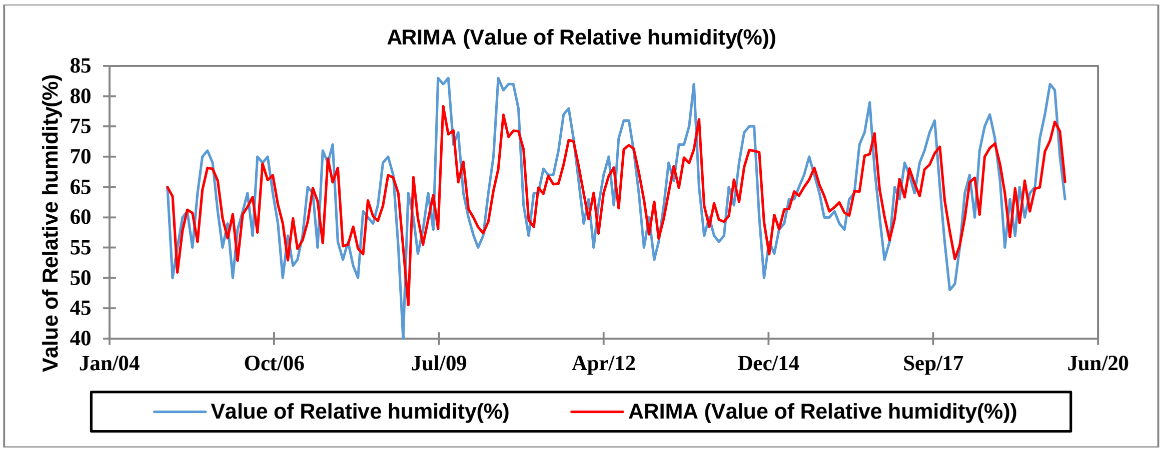
Source: Realised by authors

Fig.12 The diagram shows the observed and ARIMA (2,1,2) model forecast of PM_{10} concentration($\mu g/m^3$)



Source: Realised by authors

Fig.13 The diagram shows the observed and ARIMA (2,1,2) model forecast of average temperature(^{0}C)



Source: Realised by authors

Fig.14 The diagram shows the observed and ARIMA (2,1,2) model forecast of relative humidity (%)

Table 4: The results of different models at 95 % confidence level for forecasting the spatial distribution of meteorological parameters and pollutants

Seasons	Goodness of fit statistics	ARIMA (2,1,1) 95%	ARIMA (2,1,2) 95%	ARIMA (2,0,2) 95%
SO_2(μg/m^3)	SSE	119.162	119.200	138.288
	MAPE	28.3131	28.452	28.936
	WN variance	0.665	0.665	0.768
	FPE	0.680	0.680	0.785
	AIC	445.855	440.916	476.803
	AICC	446.084	443.263	477.147
	BIC	458.604	453.853	492.7678
NO_2(μg/m^3)	SSE	2952.916	2902.736	3258.996
	MAPE	12.923	12.929	18.105
	WN variance	0.664	0.665	0.768
	FPE	16.869	16.582	18.512
	AIC	1018.148	1017.017	1046.402
	AICC	1018.377	1017.364	1046.747
	BIC	1030.897	1032.954	1062.367
O_3(μg/m^3)	SSE	1493.261	1465.24	2385.666
	MAPE	12.833	12.686	15.780
	WN variance	21.031	20.932	33.134
	FPE	22.251	22.163	35.027
	AIC	429.934	425.85	469.045
	AICC	430.540	426.798	469.955
	BIC	438.985	437.103	480.429
$PM_{2.5}$ (μg/m^3)	SSE	33105.59	28280.29	32620.21
	MAPE	24.893	26.027	23.8596
	WN variance	0.665	0.665	0.768
	FPE	189.127	161.560	185.295
	AIC	1450.387	1427.747	1459.148
	AICC	1450.617	1428.094	1459.493
	BIC	1463.136	1443.684	1475.113
PM_{10} (μg/m^3)	SSE	77112.08	88108.4	89243.79
	MAPE	18.527	19.208	18.592
	WN variance	0.6622	0.6659	0.768
	FPE	440.529	503.3494	506.940
	AIC	1607.265	1590.669	1640.874
	AICC	1607.754	1592.016	1641.219
	BIC	1626.39	1610.606	1656.839
Average temperature (^{0}C)	SSE	638.630	614.0329	812.4368
	MAPE	5.477	5.392	6.4636
	WN variance	0.6622	0.6659	0.7682
	FPE	3.648	3.507	4.6149
	AIC	748.136	739.199	798.922

	AICC	748.625	739.546	799.267
	BIC	767.261	755.136	814.887
Relative Humidity (%)	SSE	6804.156	6794.745	8386.719
	MAPE	7.8003	7.7243	8.5956
	WN variance	0.6622	0.6659	0.768
	FPE	38.871	38.817	47.6399
	AIC	1173.16	1171.642	1217.013
	AICC	1173.648	1171.989	1217.358
	BIC	1192.284	1187.579	1232.977

Source: Realised by authors

Table 5: Annual concentration of pollutants forecasted by ARIMA (2, 1, 2) of Bhubaneshwar up to 2025

Year	**SO_2(µg/m³) Forecast value**	**NO_2(µg/m³) Forecast value**	**O_3(µg/m³) Forecast value**	**$PM_{2.5}$(µg/m³) Forecast value**	**PM_{10}(µg/m³) Forecast value**	**Average temperature(^{0}C) Forecast value**	**Relative Humidity (%) Forecast value**
2014	2.560	20.489	20.973	41.472	96.212	27.553	64.526
2015	2.515	20.078	22.305	29.450	89.734	27.886	62.751
2016	2.524	25.068	22.142	48.717	90.663	27.721	63.298
2017	2.470	23.688	24.596	53.168	93.136	27.542	65.090
2018	2.319	21.146	22.547	48.717	92.411	28.031	64.039
2019	2.247	39.027	24.062	34.640	95.907	28.721	65.320
2020	2.240	40.168	26.496	40.375	96.168	27.999	65.042
2021	2.240	38.716	26.457	40.925	92.617	28.048	66.231
2022	2.230	38.642	26.467	40.931	92.532	28.036	66.435
2023	2.220	38.639	26.467	40.931	92.530	28.034	66.452
2024	2.220	38.639	26.469	40.930	92.530	28.034	66.454
2025	2.215	38.639	26.469	40.930	92.530	28.030	66.455

Source: Realised by authors

The ARIMA (2, 1, 2) model prediction of NO_2 shows a better condition and the pollutants in future years are predicted satisfactorily below to the annual permissible limits of 50µg/m³in Bhubaneshwar (National Air Quality Index, 2014; Permissible level for pollutants, 2017). The result of the study helps to assess the conditions of different air pollutants in Bhubaneshwar in recent past years. It is inferred from the results of M-K and Sen's slope estimator tests presented in Table 1, that more control measures are required for pollutants especially for particulate matter 10, O_3 and Nitrogen dioxide. Result reveals that PM_{10} and O_3 are increasing in past years of Bhubaneswar and thus better policies are required such as improved road traffic conditions, limiting vehicular pollutions by better vehicle types these are the main sources of PM_{10} and NO_2 (CAI-Asia Factsheet, 2010; Lenschow*et al.*, 2001). After the introduction of BSES IV environment standard vehicles, the Indian government has somewhat limited the growth of traffic-related NO_2 and PM_{10} emissions (Bansal and Bandivadekar, 2013; Hilboll et al., 2017)but still, the positive trend in results indicates the need of better strategies for countering such pollutants.

SO_2 and $PM_{2.5}$, though show a decreasing tendency in previous years but the low magnitude of their slopes indicates that these pollutants also required specific measures for systematic controlling.

Inferences from the results of the ARIMA model gives an estimate that PM_{10} and O_3 are a bigger concern in the coming years and will require specific measures to control its emissions. The study summarizes that PM_{10} with an increasing trend and higher concentrations, and $PM_{2.5}$ with the decreasing trend but a still higher concentration is the primary concern in Bhubaneshwar.

4. CONCLUSION

The study presented in the paper provides a statistical analysis of trends in the atmospheric pollutants of the city Bhubaneshwar, and further, a forecasting model is formulated to predict different pollutants concentrations in the forthcoming years. M-K and Sen's slope estimator tests are applied to past pollutants data retrieved from Central Laboratory and Regional Office, State Pollution Control Board, Bhubaneswar, Odisha. The meteorological parameters (average temperature and relative humidity) are collected from the India Meteorological Department (IMD) and ARIMA (p, d, q) model is applied for predictive analysis. Results of M-K test shows the existence of a trend in some of pollutants data in different seasons and the outcomes of Sen's slope estimator test defined power of the trends. ARIMA (2, 1, 2) model resulted in being best suited for predicting the future pollutant levels by comparing the goodness of fit statistics. The forecast result shows that there is a reducing trend in the annual concentration of SO_2, NO_2 and $PM_{2.5}$, whereas O_3 and PM_{10} show a rising trend in the annual concentrations of Bhubaneswar. PM_{10} has a higher concentration which is above the standard annual concentration limits. The concentration of $SO_2$and NO_2in the air are under the permissible standard annual concentration limits 50µg/m³and 40µg/m³ respectively and are showing a decreasing trend, but still are very close to the permissible limits. PM_{10} and O_3 exhibit a rising trend with predicted approximate annual concentration of 92.530µg/m³ and 26.46µg/m³; $PM_{2.5}$, SO_2 and NO_2show a reducing trend with an approximate annual concentration of 40.93µg/m³, 2.21µg/m³ and 38.63µg/m³, respectively, by the year 2025. The meteorological parameters, average temperature and relative humidity show a rising trend in the annual concentration of Bhubaneshwar.

ACKNOWLEDGEMENT

Authors are grateful to Odisha University of Agriculture and Technology (OUAT) for facilitating the necessary support to carry out the research. Authors are also thankful to Central Laboratory and Reginal Office, State Pollution Control Board, Bhubaneshwarand Indian Meteorological Department, for making the meteorological data available for study.

ABBREVIATIONS

%	-percentage
$\mu g/m^3$	-Microgram per cubic meter
ACF	-Autocorrelation function
AIC	-Akaike's information criterion
API	-Air Pollution Index
AQI	-Air Quality Index
AR	-Auto-regression
ARIMA	-Autoregressive Integrated Moving Average
d	-Degree of differencing
D-F	-Dickey-Fuller
e_t	-Error term at time t
FPE	-Final Prediction Error
MAPE	-Mean Absolute Percentage Error
NO_2	-Nitrogen dioxide
O_3	-Ozone
PM	-Particle Matter
RMSE	-Root Mean Square Error
SO_2	-Sulphur dioxide
SSE	-Sum of Squares Error
α	-Significance level
τ	-Mann-Kendall's tau value

5. REFERENCES

[1] Alam, M.J.B., Rahman, M.H. and Jaigirdar, M.A.(1999). *Ambient Air Quality at Roadside in Dhaka City*, Pollution Research. 18 (2): 115-119.

[2] Azmi, S.Z., Latif, M.T., Ismail, A.S., Juneng, L., Jemain, A.A., (2010). *Trend and status of air quality at three different monitoring stations in the Klang Valley, Malaysia*. Air Qual. Atmos. Health., 3(1): 53- 64 (12 pages).

[3] Bansal, G., Bandivadekar, A., (2013). *Overview of India's vehicle emissions control program*. ICCT, Beijing, Berlin, Brussels, San Francisco, Washington.

[4] Box, G.E., Jenkins, G.M., (1976). *Time series analysis: forecasting and control.* Revised Ed. Holden-Day.

[5] Brocklebank, J.C., Dickey, D.A., Choi, B., (2018). *SAS for forecasting time series.* SAS institute.

[6] Caloiero, T., Coscarelli, R., Ferrari, E., Sirangelo, B., (2017). *Trend analysis of monthly mean values and extreme indices of daily temperature in a region of southern Italy*. Int. J. Climatol., 37(S1): 284-297 (14 pages).

[7] Can, A., (2017). *Time Series Analysis of air pollutants for Karabük province*. In ITM web of conferences Vol. 9, p. 02002. EDP Sciences. AMCSE, Rome, Italy 5-7 October.

[8] Chattopadhyay, G., Chakraborthy, P., Chattopadhyay, S., (2012). *Mann–Kendall trend analysis of tropospheric ozone and its modeling using ARIMA.* Theor. Appl. Climatol., 110(3): 321-328 (8 pages).

[9] Chaudhuri, S. and Dutta, D. (2014). *Mann- Kendall trend of pollutants, temperature and humidity over an urban station of India with forecast verification using different ARIMA models.* Environmental Monitoring and Assessment, 186, 4719–4742.

[10] Dickey, D.A., Fuller, W.A., 1(979). *Distribution of the estimators for autoregressive time series with a unit root*. J. Am. Stat. Assoc., 74(366a): 427-431 (5 pages)

[11] Drápela, K.; Drápelová, I., (2011). *Application of Mann-Kendall testand the Sen's slope estimates for trend detection in deposition datafromBílýKříž* (Beskydy Mts., the Czech Republic) 1997-2010.Beskydy, 4(2): 133-146

[12] Gupta, P. G., Kumar, B., Singh, S. and Kulshrestha, U. C. (2016). *Deposition and impact of Urban Atmospheric Dust on two medicinal plants during different seasons in NCR Delhi.* Aerosol and Air Quality Research, 16: 2920- 2932.

[13] Gurjar, B.R.; Butler, T.M.; Lawrence, M.G.; Lelieveld, J., (2008). *Evaluation of emissions and air quality in megacities.* Atmos. Environ., 42(7): 1593-1606

[14] Jaiswal, A., Samuel, C. and Kadabgon, V.M. (2018). *Statistical trend analysis and forecast modeling od air pollutants*. Global Journal of Environmental Science and Management, 4(4): 427-438.

[15] Jaruskova, D.; Liska, I., (2011). *Statistical analysis of trends inorganic pollution and pollution by nutrients at selected Danuberiver stations*. J. Environ. Monit., 13(5): 1435-1445

[16] Kisi, O., Ay, M., (2014). *Comparison of Mann–Kendall and innovative trend method for water quality parameters of the Kizilirmak river*, Turkey. J. Hydrol., 513: 362-375 (14 page).

[17] Lenschow, P., Abraham, H.J., Kutzner, K., Lutz, M., Preuß, J.D., Reichenbächer, W., (2001). *Some ideas about the sources of PM10.* Atmos. Environ., 35: S23-S33 (11 pages).

[18] Mondal, A., Kundu, S. and Mukhopadhyay, A. (2012). *Rainfall trend analysis by Mann-Kendall test: A case Study of North-Eastern Part of Cuttack District, Orissa.* International Journal of Geology, Earth and Environmental Sciences. 2(1): 70-78.

[19] National Air Quality Index, (2014). CPCB, *ministry of environment, forests & climate Change*, GoI.

[20] Permissible level for pollutants, (2017). *Ministry of environment, forest and climate change*, GoI.

[21] Rahman, A.M., Sultana, N. and Yunsheng, L., (2016). *Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman's rho tests and ARIMA model.* Meteorology and Atmospheric Physics, 016-0479-4.

[22] Rani, N.L.A., Azid, A., Khalit, S.I., Juahir, H., Samsudin, M.S., (2018). *Air pollution index trend analysis in Malaysia, 2010-15.* Pol. J. Environ. Stud., 27(2): 801–807 (7 pages).

[23] Ravichandran, C., Chandrasekharan, G. E., and Srikanth, S. (1998). *A Short Report on Ambient Air Quality inside and Outside Tiruchirapalli City*. Indian Journal of Environmental Protection. 18(6): 440-442.

[24] Watthanacheewakul, L., (2011). *Comparisons of power of parametric and nonparametric test for testing means of several Weibull populations*. In Proceedings of the International Multi-Conference of Engineers and Computer Scientist. Hong Kong, 16-18 March