SindhUniv. Res. Jour. (Sci. Ser.) Vol. 50 (004) 621- 624 (2018)

http://doi.org/10.26692/sujo/2018.12.00100

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Forecasting of Metrological Parameters by Decomposition Method: Case of Karachi Pakistan

R. SOOMRO, M. I. RAJPUT*

Government Degree Boys College Gulistan-e-Jouhar Karachi

Received 21st October 2017 and Revised 19th September 2018

Abstract: Forecasting is an important subject for different applications in Metrology and Environment fields. The most simple and basic approach for analyzing time series is decomposition. This study is based on the trend component for significant parameters which were obtained through regression models fitted to the polynomial functions of power k for time t. In this regard monthly data of six metrological parameters precipitation amount (R_f), minimum temperature (T_{min}), maximum temperature (T_{max}), relative humidity (R_h), wind speed (W_s) and atmospheric pressure (A_p) for the period of 25 years (1990-2014 or 300 data points) of Karachi city, Pakistan were considered for forecast. The forecast values will be obtained by multiplying the values of these four components (trend, seasonal, irregular and cyclic) and were compared to the original values to justify the model. The reliability of forecasted values will be checked by using goodness of fit test by means of chi squire statistics χ^2, which indicates that all the parameters follow the test except precipitation. After justification we found that our models are accurate and reliable.

Keywords: Climate variables, time series, decomposition, goodness of fit test.

1. INTRODUCTION

The meteorological parameters in a time series are often sequentially dependent. The goal of time series model is the nature of this dependence, which in turn allows predicting or forecasting. The challenges have grown over time, the original use of time series analysis was primarily as an aid to forecasting. Augustine (Asuquo *et al*, 2013) have worked and modelled correlation and regression equations for the Akwa Ibom state for period of 10 years (2003-2012). The selected parameters were rainfall and relative humidity. This study concluded that there is positive correlation between the parameters, rainfall and relative humidity. (Antonia and Paulo 2009), have worked on the time series analysis of Italy. They have calculated seasonal and annual trend analysis for variable precipitation, for data of 1918-1999. They have concluded that there was negative but significant trend on annual and seasonal basis except summer period.(Ebenzer *et al*, 2016), they have used seasonal ARIMA model and decomposition of time series technique for the forecasting the rainfall. Unit root test identification was checked by Augmented Dickey Fuller test and ACF and PACF were also measured. They have founded that more rainfall producing months are September and October and best performance criterion is AIC criteria and seasonal ARIMA model predicted good values and these models were helpful in the long run for accurate forecasting and educational entertainments.(Habiyaremye and Jiwen, 2010), have carried out the statistical analysis on parameters rainfall and temperature and concluded that temperature alone is not sufficient to explain the amount of rainfall i-e rainfall and temperature are independent. They have used monthly data of 5 years (2008-2013). The selected variables were macroeconomic variable (CPI) and growth rate (LSM). The data analyzed through two time series methods, decomposition and ARIMA models. For checking accuracy of the results they have used mean absolute deviation (MAD) and sum of squire error (SSE). They have found that ARIMA could give better results.(Nigarand, 2015), used seasonal ARIMA and decomposition of time series methods. They selected variable temperature at minimum and maximum levels. Unit root identification was checked by ADF test. ACF and PACF were also measured. For diagnosing the forecasting model, stability test and serial correlation test was applied. It is suggested that seasonal ARIMA model is better for the time dependent data.(Akpinar *et al.*,2010), have worked on modelling of metrological data for Turkey. They have used monthly data of 10 years (1994-2003) for 13 cities. The data was analyzed through regression methods. They have also found polynomial and trigonometric functions of time 't'. They have concluded that models were adequate and can be used for determination of specific parameter values and predicting the missing values. (Muslehuddin *et al*, 2005), the probability of Sindh monsoon rainfall was analyzed with relation to the some important global and regional parameters for 30 years (1961-1990) on

Corresponding Author Ramzan Soomro Email: ramzsom@yahoo.com

*Department of Statistics, University of Sindh Jamshoro,

monthly data, which showed that the 60% rain of annual occurred in monsoon season. Analysis carried out the indices concluded that the precipitation plays an important role in development of Pakistan. (Prema and Rao, 2015), have used time series techniques for cities of India with short term prediction for the daily data of 5 months (Feb.2013–June. 2013). The techniques, Decomposition of time series, ARIMA and BPNN models were also developed. It is concluded that decomposition is a best model comparing to other models.

In view of these researchers, this work has been proposed for Karachi city containing the monthly average data of six metrological parameters precipitation amount(R_f), minimum temperature (T_{min}), maximum temperature (T_{max}), relative humidity (R_h), wind speed (W_s) and atmospheric pressure (A_p) for period 25 years (1990-2014). Time series modeling will be used to investigate changes over time related to the past time periods. The time series models are decomposed into four components, linear trend, seasonal variation, cyclic variation and irregular or accidental variation. In this study we will analyze the trend for metrological parameters through regression models fitted to the polynomial functions of power k at time t and found significant parameters precipitation and wind speed. For finding the trend component of other non-significant parameters we will use five years moving average. The forecast values will be obtained by multiplying the values of these four components and were compared to the original values to justify the models, goodness of fit test will be used to check that how much our models are reliable for forecasting the amount of different meteorological parameters. After justification we found that our forecast values obtained through these models are near to the original values. In last we will forecast the monthly values of allMetrological parameters for year 2015. Time series models and forecasting have many applications in hydrology and environment.

2. DATA AND METHODOLOGY

Monthly data of different meteorological parameters precipitation, minimum & maximum temperature, relative humidity, wind speed and atmospheric pressure for 25 years (1990-2014) for Karachi city were obtained from Computerized Data Processing Centre (CDPC) of Pakistan Meteorological Department (PMD) Karachi.

Decomposition of all six parameters:

Decomposition methods identify four separate components of the basic underlying pattern that characterize meteorological and Environmental series. These are the seasonal, trend, cycle and irregular components. The seasonal component relates to periodic fluctuation of common length and proportional depth that are caused by natural phenomena. The trend component, which represents the long run behavior of the data, can be increasing, decreasing or unchanged. Often the trend can be approximated by straight line, but an exponential, S-curve or other long term pattern may exist in certain situations. The cyclic component represents the ups and downs caused by natural phenomena. The cyclic often follows the pattern of a wave, passing from a large to a small value and back again to a large value. The irregular component represents the unsystematic variation occurring completely in and unpredictable manner.

2. METHODOLOGY

The approaches to decompose a time series, have aim to isolate each component of the time series as accurately as possible. The basic concept in such separation is empirical and consists of removing first seasonality, then trend, then cyclic and finally irregular component. The procedure is as under:

Table.1: Descriptive Statistics for different parameters for 1990-2014

Statistics	R_f	T_{min}	T_{max}	R_h	W_s	A_p
Minimum	0.00	9.00	24.40	24.00	1.50	996.40
Maximum	270.40	29.80	37.00	78.00	14.10	1018.00
Range	270.40	20.80	12.60	54.00	12.60	21.60
Mean	17.51	21.42	32.37	48.03	7.60	1007.28
Std. dev	38.94	5.89	3.14	14.11	2.53	6.04
Total Obs:	300	300	300	300	300	300

Source: Authors calculations

Methodology is distributed into these main parts

By taking 12 months moving average (mma) of any time series (Y_t) we acquired TC component, by taking ratio of Y_t and TC we get SI component. Again taking 5 mma of component SI, we achieved seasonal component S. By taking ratio of the component SI and S we get component I. The trend component T can beobtained through regression model fitted to the polynomial function of power kat time t and found significant parameters precipitation and wind speed, thus we use five years moving average for all other nonsignificant parameters. Then by taking ratio of TC and trend component T, we obtained component C. In this way we decompose the time series into four components s (seasonal), T (trend), C (cyclic) and I (irregular). The general mathematical representation of the decomposition approach is:

$$\boldsymbol{Y_t = f(S, T, C, I)}$$

Where
Y_t= Time series value (actual data) at period t,
S = Seasonal factor
T = Trend factor,C = Cyclic factor, I= irregular factor
Two functional forms of time series used to relate these four components are generally applied for different series i-e multiplicative and additive models, which are

$$Y_t = S * T * C * I \quad \text{(Multiplicative rule)}$$
$$Y_t = S + T + C + I \quad \text{(Additive rule)}$$

As mostly used model is multiplication which gives more convenient values, here we also use multiplication model to forecast short term values of all meteorological parameters for city Karachi. The trend component for significant parameters may be obtained through regression models fitted to the polynomial functions of power k for time t, otherwise five years moving average i-e

$$X_t = c + t + t^2 + \cdots t^k$$

Where X_t is any parameter and t is time component (month)

To find out the reliability and sufficiency of the forecast values, i-e analytical checking, goodness of fit test is used for analyzing the residuals obtained from the forecast models through chi squire test χ^2at 5 % level of significance.

$$\chi^2 = \sum\left[\frac{(Oi - Ei)^2}{Ei}\right] \text{ withd.f, v = k-1}$$

Where O_i = Original values
E_i =Forecasted value obtained through decomposition of time series.

Forecasting by Re composition:

For short term forecasting we use re composition technique for next six months, we use average of all components for each month and by using multiplicative rule we obtained the forecast values for each parameter.

3. RESULTS

Analyzing the trend of the meteorological parameters, precipitation, minimum &maximum temperature, relative humidity, wind speed atmospheric pressure, decomposition of the time series were performed to forecast the values of different meteorological parameters. Measuring the reliability of the meteorological parameters goodness of fit test is applied, the values are listed in table.2, which shows that all the parameters follow the test except precipitation. The range of precipitation is (270.4), which is very high shown in table.1, having smaller mean (17.51) and high standard deviation(38.94), thus it is very typical to forecast the exact value of precipitation, i-e in table.2 goodness of fit test does not follow this series, this is due to the nonoccurrence of the precipitation in the region, but in other five parameters, range of T_{min} is 20.80 with SD(5.89), range of T_{max} is (12.60) with SD (3.14), range of R_h is (54.0) with SD(14.11), range of Ws is (12.6) with SD(2.53), range of Ap is (21.6) with SD(6.04), all are sufficient values shown in table.1and goodness of fit test follow these series shown in **(Table.2).**

Figure.1: Comparison of the original and forecasted values for year 2015.

Table.2: Goodness of fit test for the different parameters for year 2015

Parameter	Observed	Expected	Weighted	Tabulated	Conclusion
R_f	53.20	209.89	153.78	36.42	$\chi^2_{cal} > \chi^2_{tab}$
T_{min}	269.10	258.60	0.90	36.42	$\chi^2_{cal} < \chi^2_{tab}$
T_{max}	394.70	389.02	0.29	36.42	$\chi^2_{cal} < \chi^2_{tab}$
R_h	575.00	574.66	5.58	36.42	$\chi^2_{cal} < \chi^2_{tab}$
W_s	120.50	92.53	9.32	36.42	$\chi^2_{cal} < \chi^2_{tab}$
A_p	12095.10	12087.45	0.01	36.42	$\chi^2_{cal} < \chi^2_{tab}$

Source: Authors calculations

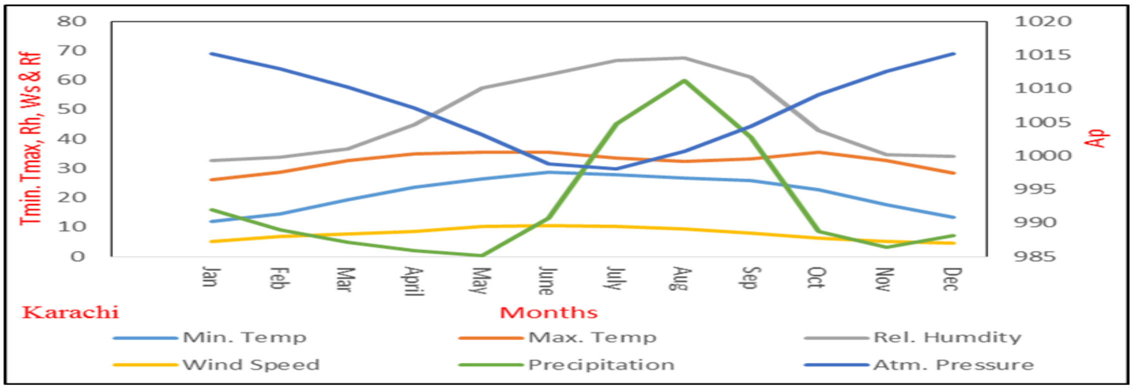

Figure.2: Forecasted values for different parameters for year 2015.

4. CONCLUSIONS

In this paper, we attempted to forecast monthly values of the different meteorological parameters (fig.2)by decomposition of time series. The forecasted values are compared to the original values (fig.1) through goodness of fit test and found no deviation between predicted values from the actual for all parameters (table.2) except precipitation, thus the predicted models are reliable. But only in precipitation there are some deviations, in some cases it is due to delay in the actual commencement cases of the monsoon and zero values in the data. It is suggested that, zero entries should be replaced with non-zero entries (minimum values i-e 0.01 or 0.001) to accommodate data.

5. AKNOWLEDGMENT

The authors are very grateful to Mr. Nadeem Faisal, Deputy Director, Metrological department and CDPC Karachi for his value able advices and providing meteorological data.

The authors are also grateful to Dr. Saghir Pervaiz Ghauri, research department, state Bank of Pakistan Karachi for his value able suggestions and support.

REFERENCES:

Augustine A. U., (2013), “Rainfall and relative humidity occurrence patterns in Uyo Metropolis, Akwa Ibom state, South-South Nigeria” IOSR Journal of Engineering, Vol.3 Issue.8, 27-31.

Antonia L. and P. Villani, (2009), “Trend analysis of annual and seasonal rainfall time series in the Mediterranean area”, International Journal of Climatology, Royal metrological Society, 1-9.

Akpinar S. and E. K. Akpinar, (2010), “Modelling weather data for the east Anatolia region of Turkey”, J. of urban and Environmental Engineering, 4(1), 9-22.

Ebenezer A., (2016), “SARIMA modelling and forecasting of monthly rainfall in the BrongAhafo region of Ghana”, World environment, Vol. 6(1), 1-9.

Habiyaremye G. and G. Jiwen, (2010), “Statistical analysis of categorical climatic variables, Case of temperature and rainfall”, 2nd conference on Environmental Science and Information Application Technology, 412-415.

Mohammad M., (2005), “Sindh summer (June-Sep) monsoon rainfall prediction”, Pakistan Journal of Meteorology, Vol.2, Issue.4, 91-108.

Nigar S. and Md. M. Hasan, (2015), “Forecasting Temperature in the coastal area of Bay of Bengal. An application of Box-Jenkins seasonal ARIMA model”, Civil and Environmental Res, Vol.7, No.8, 149 –160.

Prema, V. and U. Rao, (2015), “Time series decomposition model for accurate wind speed forecast”, Wind, water and Solar Journal, No. 2 (8),1-11.