DEVELOPMENT OF FORECASTING MODELS FOR THE GLOBAL TEMPERATURE

Sarojamma B[1], Naresh M, [2]Anil Kumar K[3], and Venkatramana Reddy S[4]

***Abstract:** Temperature plays an important role in growth of plants and also animals. Temperature is growing decade to decade in many places and in some other places it is going downward, it is because of global warming. Temperature fluctuations are carried out due to abnormal cutting of trees, deforestation, picking soils from sea shores, etc. Forecasting temperature is essential not only for atmospheric scientists but also human beings and plants. In this paper two models ARMA (3, 3) and 3^{rd} degree polynomial regression are fitted for forecasting global temperature. Theil's U test is used for testing accuracy of forecasted models with naïve model and MSE is calculated for both models ARMA (3, 3) and 3^{rd} degree polynomial model. We suggest that the 3^{rd} degree polynomial regression model is the best for temperature forecast.*

Keywords: Global temperature, ARMA (3, 3), Polynomial Regression, Theil's U-test, MSE, accuracy.

1. INTRODUCTION

Temperature is essential for growth of human beings, plants, animals and without proper temperature cultivation is not carried out. If sufficient temperature exists only plants will develop well and they give fruits, otherwise plants cannot grow well, because they cannot produce photosynthesis without sunlight. Therefore, forecasting temperature plays an important role for meteorological sciences. Global temperature is changing from January to December for location to location. Temperature is raising decade to decade because of deforestation, cutting trees abnormally, picking soils from sea shores, utilization of radiators heavily and emission of green house gases through industries, etc.

Venkatramana Reddy et al [1] published a paper on "Time Series Regression models for temperature of Global, Land and Ocean". They separately fitted higher order regression models for Global, Land and Ocean temperature and by using Chow test they studied the structural break for temperature of Global, Land and Ocean. The structural break for Ocean, Land and Global is occurred for the year 2005. During the period from the years 1997 to 2011, the highest temperature recorded in 2005. Jon K. Eischeid et al [2] used several estimation techniques based on special objective analysis schemes to estimate a missing value replacement. This development of data set was useful in a variety of meteorological and hydrological research applications. They developed special interpolation schemes separately by interpolation method and Calendar month on Cross validation. There was no error found in the estimation procedure. The result gave forecast up to 10^{th} year 2034, maximum – minimum daily time series temperatures for the western United States and 2962 total daily precipitation locations. Igor V. Polyakov et al [3] in their article "Variability and trends of air temperature and pressure in maritime Arctic, 1875-2000", studied Arctic atmosphere variability

ISBN - 978-93-81583-57-9

during industrial era and assessed using sea level pressure (SLP) and spatially averaged Air temperature (SAT). P.D. Jones [4] in his article on "Urbanization effects in large scale temperature records with an emphasis on China". In this paper he pointed out how temperature effects due to urbanization. The effects of data gaps on the calculated monthly mean maximum and minimum temperatures in the continental United States. A spatial and temporal study was given by Stook Sbury et al [5].

2. METHODOLOGY

In Time series many methods are used for forecasting. Simple regression model, logistic regression model, growth curves, polynomial curves, exponential smoothing model, double exponential smoothing model, auto regression integrated moving average model, Generalized Auto regression conditional heteroscadisticity model, etc [6]. All the above models are fitted to the data so that error of model is minimum. A model which have minimum error for any data is called the best model for that data.

In this paper we fitted two Time series models for data. The first model is Auto Regressive moving average (ARMA) model and second one is polynomial time series regression model.

ARMA model: Auto regression of order 'p' combined with moving average of order 'q', we get auto regressive moving average (ARMA) model. In this model present time series value 'Y_t' is depending upon past time series values Y_{t-1}, Y_{t-2}, Y_{t-3}, ..., Y_{t-p} and also the present and past error terms e_t, e_{t-1}, e_{t-2},..., e_{t-q}. Auto regression tells about the relationship between current time series value and past time series values and moving averages studies about relation between error with time series values.

The model of ARMA (p, q) is

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

ARMA (p, q) model with back shift operator is

$$(1 - \emptyset_1 B - \cdots \emptyset_p B^p) Y_t = C + (1 - \theta_1 B - \theta_2 B^2 - \cdots \theta_q B^q)\, e_t$$

Here B is Back shift operator.

Auto Regressive moving average (ARMA) (3, 3): For this model third Auto regressive is mixed with third order moving average. The model ARMA (3, 3) is as follows.

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3}$$

(or)

$$(1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3) Y_t = C + (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3)\, e_t$$

Where B is back shift operator and

$BY_t = Y_{t-1}$, $B^2 Y_t = Y_{t-2}$, $B^3Y_t = Y_{t-3}$.

$Be_t = e_{t-1}$, $B^2 e_t = e_{t-2}$, $B^3e_t = e_{t-3}$.

Polynomial regression model: A polynomial of order p is fitted to data is as follows.

$Y_t = b_0 + b_1t + b_2t^2 + b_3t^3 + \text{---} + b_pt^p$.

Here Y_t: time series value at time 't'.

b_0, b_1, ---, b_p are regression coefficients, which are estimated using principle of least squares and t: time

Polynomial regression is fitted to data by assuming time series value is in linear form with time 't'. In this paper the time series equation, 3rd degree polynomial regression equation is

$Y_t = b_0 + b_1t + b_2t^2 + b_3t^3$

By using principle of least squares b_0, b_1, b_2 and b_3 are estimated. The fitted polynomial regression equation to data is

$$\widehat{Y_t} = \widehat{b_0} + \widehat{b_1}\, t + \widehat{b_2}t^2 + \widehat{b_3}\, t^3$$

Theil's U – Statistic: Theil's U – Statistic is used as a measure of accuracy. This statistic allows a relative comparison of formal forecasting methods with naïve approaches and also squares the errors involved, so that large errors are given much more weight than small errors. Mathematically Theil's U – Statistic is defined as

$$U = \sqrt{\frac{\sum_{t=1}^{n-1}(FPE_{t+1} + APE_{t+1})^2}{\sum_{t=1}^{n-1}\left(APE_{t+1}\right)^2}}$$

Where $FPE_{t+1} = \frac{F_{t+1} - Y_t}{Y_t}$

$APE_{t+1} = \frac{Y_{t+1} - Y_t}{Y_t}$

FPE_{t+1} is a forecast relative change, APE_{t+1} is an actual relative change.

The ranges of U-statistic is summarized as follows

U = 1: use naïve method or forecasting method, both will give you the same result.
U < 1: the forecasting method gives better result than the naïve method.
U > 1: the naïve method gives better result than forecasting method.

Mean square error criterion: Difference between observed and calculated values is residue or error. Average squared errors become mean square error or simply MSE. We calculated MSE's for both the models. A model having least MSE is better model than another model. If MSE of ARMA (3, 3) is greater than MSE of polynomial regression of order 3, then polynomial regression of order 3 is better model than ARMA(3, 3)

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2$$

Where n = No. of observations,
e_t = error $= Y_t - \widehat{Y_t}$

3. EMPIRICAL INVESTIGATIONS

The third order auto regression mixed with third order moving average is given by

$Y_t = -0.9 - 1.8425\ Y_{t-1} - 1.1519\ Y_{t-2} - 0.1473\ Y_{t-3}$

$- 2.9793\ e_{t-1} - 2.4024\ e_{t-2} - 0.9\ e_{t-3}$

The third degree polynomial regression model fitted to data is

$Y_t = -8232866 + 12365.86t - 6.160922\ t^2 + 0.00102\ t^3$

Theil's U-statistic and MSE: For testing accuracy of forecasting model, we use Theil's U-statistic and mean square error is calculated for Auto regressive moving average (3, 3) model and 3rd degree polynomial regression model, which model is better model.

Table - I

Models	*U-statistic values*	*MSE values*
ARMA(3, 3)	*1.0001*	*0.063048*
3rd degree polynomial regression	*1.0001*	*0.009409*

From the above table-I, U-statistic values for both models is 1.0001. By using naïve method ARMA (3, 3) and 3rd degree polynomial regression are good for forecasting. Mean square error value of 3rd degree polynomial regression is smaller than the MSE of auto regressive moving average (3, 3). Therefore, 3rd degree polynomial regression gives the best forecasting than the auto regressive moving average (3, 3).

The data for interpretation is global temperature data for years 1999 to 2011 and for the months January to December [7]. The des

criptive statistics of the data is listed in the Table - II.

Table-II

	Max	Min	Average	Median	S.D.	C.V.	Skewness	Kurtosis
Jan	58.784	57.47	58.12908	58.154	0.381451	0.656214	-0.27967	-0.25725
Feb	58.414	57.65	58.17338	58.19	0.24244	0.416753	-0.66616	0.218118
Mar	58.766	57.686	58.262	58.244	0.278952	0.478788	0.056837	1.08645
Apr	58.568	57.65	58.12492	58.136	0.227392	0.391212	-0.01365	1.161889
May	58.352	57.596	58.03492	58.136	0.226678	0.39059	-0.49478	-0.59911
Jun	58.28	57.812	58.02938	58.046	0.166514	0.286948	-0.04055	-1.50854
Jul	58.37	57.542	58.064	58.136	0.23891	0.41146	-0.8878	0.559912
Aug	58.352	57.704	58.09308	58.154	0.206221	0.354984	-0.42431	-0.78221
Sep	58.388	57.74	58.188	58.136	0.180599	0.310745	-0.69153	0.782206
Oct	58.496	57.542	58.15538	58.19	0.257769	0.443242	-1.28441	1.741796
Nov	58.46	57.704	58.17477	58.226	0.252519	0.434068	-0.83818	-0.28857
Dec	58.478	57.632	58.07923	58.064	0.24558	0.422837	0.03781	-0.42919

3. SUMMARY AND CONCLUSIONS

We fitted two models for forecasting temperature and, they are ARMA (3, 3) and 3rd degree polynomial regression model and are

$$Y_t = -0.9 - 1.8425\ Y_{t-1} - 1.1519\ Y_{t-2} - 0.1473\ Y_{t-3} - 2.9793\ e_{t-1} - 2.4024\ e_{t-2} - 0.9e_{t-3}$$

$$Y_t = -8232866 + 12365.86t - 6.160922\ t^2 + 0.00102\ t^3$$ respectively.

For testing accuracy of forecasting models with naïve models we are using Theil's U-statistic. The U-value for both models is 1.0001 and is nearly equal to '1'. Usage of forecast models ARMA (3, 3), 3rd degree polynomial regression and naïve method gives the same forecast values.

Mean square error of 3rd degree polynomial regression model has 0.00949, which is less than mean square error of auto regressive moving average (3, 3) value 0.063048. Therefore, 3rd degree polynomial regression model is the best model when compared with auto regressive moving average (3, 3) model.

4. REFERENCES

1. S. Venkatramana Reddy, G. Karthick Kumar Reddy, M. Naresh, S. C. Thasleema, B. Sarojamma, R. Rajanikanth and T.K. Ramkumar: International Journal of Environmental Research and Development. Vol. 2, No.2, pp.109-121, 2012.
2. Jon K. Eischeid, Phil A. Pasteris, Henry F. Diaz: Journal of Applied Meteorological, Vol. 39, pp. 1580-1591, 2000.
3. Polyakov, Jgor V., Roman V. Benkryev, Genrikh V. Alekseev, Uma S. Bhatt, Roger L. colony, Mark A. Johnson, Alexander P. Masshtas, david Walsh: Journal of Climate, Vol. 16, pp. 2067-2077, 2003.
4. P.D. Jones, D.H. Lister and Q.Li: Journal of Geophysical Research, Vol. 113, D16122, 2008, doi:10.1029/2008JD009916.

5. D.E. Stooksbury, C.D. Idso and K.G. Hubbard, Journal of Climate, Vol.12, pp 1524-1533, 1999.
6. Makridakis, S.C. Wheelwrignt, Rob J. Hyndman, Forecasting methods and applications, John Wiley & Sons, 3/e, New York, 2005.
7. National Aeronautics and Space Administration Goddard Institute for Space Studies (NASA GISS), http://data.giss.nasa.gov/gistemp/tabledata_v3/GLB.Ts.txt

****

----

[1]Author 1 Sarojamma B, Department of Statistics, Sri Venkateswa University, Tirupati-517502, India.
saroja14397@rediffmail.com
[2]Author: Naresh M, Department of Statistics, Sri Venkateswara University, Tirupati-517502, India.
naresh.ygr@gmail.com
[3]Author: Anil Kumar K, Department of Statistics, Oxford College of Science, Bangalore, India.
anilstats@rediffmail.com
[4]Author: Venkatramana Reddy S, Department of Physics, Sri Venkateswara University, Tirupati-517502, India.
** Author for correspondence*
drsvreddy123@gmail.com