International Journal of Scientific Research in Mathematical and Statistical Sciences
Volume-5, Issue-3, pp.40-47, June (2018)

Research Paper

E-ISSN: 2348-4519

Some Interval Time Series Models for Temperature Data in India

A.Srinivasulu[1], B.Sarojamma[2*], P.Ramakrishna Reddy[3], S.Venkataramana reddy[4]

[1] Department of Statistics, S.V.University, Tirupati-517502, A.P, India
[2] Department of Statistics, S.V.University, Tirupati-517502, A.P, India
[3] Department of Statistics, S.D.H.R Degree & P.G college, Tirupati, A.P, India
[4] Department of Physics, S.V.University, Tirupati-517502, A.P, India

[*] *Corresponding Author: saroja14397@rediffmail.com, Tel.: +91 8639652810*

Available online at: www.isroset.org

Received: 06/May/2018, Revised: 19/May/2018, Accepted: 13/Jun/2018, Online: 30/Jun/2018

***Abstract*:** The temperature is fluctuated by special climate changes during seasons in India. In this paper, we were taking temperature data from 2000 to 2016. Maximum and minimum temperature values for season wise i.e., Jan-Mar(Spring), Apr-June(Summer), July-Sep(Autumn), Oct-Dec(Winter) has to be taken throughout India. The organization of the work is divided into two parts, first part contains data from 2000 to 2010 as test group and second part contains from 2011 to 2016 as main group. For this maximum and minimum temperature seasonal data we apply nine models. Among the nine, seven are ARIMA models, the 8th one is Adaptive smoothing model and the last one is non linear model. In this paper, two measures of accuracy are used. They are Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The nine models are empirically tested using Maximum and Minimum temperature data of season wise in India.

***Keywords*:** Seasonal temperature data, Adaptive smoothing model, ARIMA models, MAE, RMSE

I. INTRODUCTON

The temperature can be fluctuated by seasons in India because of Himalayas on one side, and the other sides Bay of Bengal, Arabian Sea and Indian Ocean. Due to these, the climate change in India is differ with the other countries regarding the temperature. For the last two decades the temperature variation is more due to global warming i.e., Emission of green house gases such as Carbon Dioxide, Methane and Nitrous oxide. Burning of fossil fuels like coal, oil and natural gas for energy, cut down and burn forests, dangerous pesticides used in agriculture, deforestation and farming. Data analysis has been carried for season wise in India using interval time series temperature data from 2000 to 2016. Maximum and minimum temperature values for season wise i.e., Jan-Mar, Apr-June, July-Sep, Oct-Dec has to be taken throughout India. By conducting three interval time series models to this data. These three models are compared using R^2 criteria for the best model for predicting the temperature values.

Form the above paragraph of the introduction, we are provided the introduction of the paper. The rest of the sections is as follows. Section II contains review of literature given by different authors in their respective papers in bibliography we are given the list of papers used in the literature. Section III contains nine models as ARIMA(1,0,0) -First order Auto Regressive model, ARIMA (0,1,0)-Random Walk Model, ARIMA(1,1,0) - Differenced first order Auto Regression model, ARIMA(0,1,1) without constant-Simple Exponential Smoothing, ARIMA(0,1,1) with constant-Simple Exponential Smoothing with growth, ARIMA (0,2,1) or (0,2,2)-Linear Exponential Smoothing, ARIMA (1,1,2) without constant-Damped-Trend linear Exponential smoothing, Adaptive Smoothing Model and The Non linear model. we are compare these nine models using error criteria like MAE and RMSE. Section IV contains empirical discussion by taking temperature data from 2000 to 2016. Maximum and minimum temperature values for season wise i.e., Jan-Mar, Apr-June, July-Sep, Oct-Dec has to be taken throughout India. The organization of the work is divided into two parts, first part contains data from 2000 to 2010 as test group and second part contains from 2011 to 2016 as main group. The comparison of the data is by using MAE & RMSE for interval data. Section V contains brief summery and conclusions of the above sections.

II. RELATED WORK

The prime indicator of global warming is global mean temperature. Time series of global temperature show a well known rise since the early 20th century and most notably since the late 1970s. the impacts are: shrinking mountain glaciers, accelerating ice loss from ice sheets in Greenland and Antarctica, shrinking Arctic sea ice extent, sea level rise,

and a number of well-documented biospheric changes like earlier bud burst and blossoming time is spring. Much of the variability during that time span can be related to three known causes of short-term temperature variations: El Nino/southern oscillation, volcanic eruptions, and solar variations including the solar cycle. This complicates both comparison and trend analysis of the temperature records. Since independent measures of these variations are available, their influence can to a large extent be removed, leading to adjusted, less noisy global temperature data sets. Therefore we will remove the influence of these factors on the temperature data sets, not only to isolate the longer-term changes, but also to identify whether different data sets show meaningful differences in their response to these factors. The influence of exogenous factors will be approximated by multiple regression of temperature against ENSO, volcanic influence, total solar irradiance (TSI) and a linear time trend to approximate the global warming that has occurred during 32 years subject to analysis [1].

Global surface temperatures continue to rise. In most surface temperature data sets, the years 2014, 2015, and 2016 set new global records since the start of regular measurements. Never before have three record years occurred in a row. Global-mean surface temperature(GMST) is the most important indicator of global climate change, because (i) it is directly related to the planetary energy balance and increases quasi-linearly with cumulative greenhouse gas emission (ii) GMST is directly related to most climate impacts and risks. In this, the authors deal with the former only, i.e. with analysis of possible trend changes in the observational data [2].

Now a days the forecasting of agricultural commodity future prices is very essential. Because now a days the growth of population is high in the countries like India, China etc. In recent years, agriculture commodity futures markets in populated countries have witnessed massive growth with an increasing product variety and deepering liquidity pools. The agricultural commodity futures markets in India are playing an increasing important role in serving the global financial market and the national economy [3].

Agricultural commodity futures price forecasting is considered as a challenging task. Due to the fact that the prices are highly volatile, complex and dynamic and is thus of great interests to finance researchers, market practitioners and policy makers. An extensive investigation reveals that it is not difficult to find futures prices including stock index futures, gold futures, and metal futures. But an important point note from the past studies is their preoccupation with point forecasting rather than interval one. An interval forecasting of futures prices has the advantage of taking into account the variability and/or uncertainty so as to reduce the amount of random variation relative to that found in classic single valued futures prices time series. Interval analysis and forecasting has attracted particular attention in various fields particularly in finance market and energy market. The Interval – valued time series (ITS) forecasting method is a potential tool and will lead to a reduction in risk when making power system planning and operational decisions.

A variety of Interval- valued time series forecasting methods has been developed. They are Traditional statistical techniques, including interval Exponential smoothing methods, Vector Auto Regressive (VAR) model and Vector error correction model (VECM). The traditional statistical techniques can provide good predictions only when ITS under study are linear and stationary. But it may not possible all times. Sometimes it appears as nonlinear and non stationary due to intrinsic complexity and volatility of ITS. In order to overcome the limitations of traditional statistical techniques, machine learning techniques have recently attracted many attentions. The Interval multilayer perceptions, Multi output Supports Vector Regression (MSVR) are very useful in nonlinear modeling capability for ITS in real world.

Actually in real world, ITS appears linear (or) non linear pattern and usually contains both patterns. This difficult forecasting can be partially solved by using combined linear and non linear model. It is very difficult in practice to construct a single model which is the best in all situations. Such as a hybrid Auto Regressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) model for time series forecasting to take advantage in linear and non linear modeling respectively.

Now a days over confidence is one of the most prevalent judgment biases. Several studies show that over confidence can lead to sub optimal decisions of investors, managers or politicians. It is the object of an active field of research over the last two decades. Over confidence indeed a bias or rather than an ecological and statistical illusion and therefore only an apparent anomaly that just seems to exist, but it is not real. There are different types of overconfidence i) overestimation of one's actual performance ii) over placement of one's performance relative to others called the better than average effect iii) excessive precision in one's belief, called miscalibration.

It is important to distinguish between these main manifestations of overconfidence, in particular between (relative) performance based and miscalibration based measures, because empirical studies find them typically to be hardly related. The degree of miscalibration can be measured in various ways, such as binary choice questions or interval estimates. For difficult reasons, the analysis is restricted to miscalibration in interval estimates; it is most closely related to the facet of overconfidence that is modeled in economics, finance and management. Interval estimates are less studied than two-choice questions. Newly designed method of measuring "true overconfidence" is naturally related to interval estimates [4].

The widely used Generalized Additive Models (GAM) method is a flexible and effective technique for conducting nonlinear regression analysis in time series studies of health effects of air pollution. The GAM is being applied when the estimated regression coefficients are small and there exist confounding factors that are modeled using at least two non parametric smooth functions [5].

III. METHODOLOGY

We fitted seven Auto Regressed Integrated Moving Average (ARIMA) models and the other two are

ARIMA (1,0,0) - First order Auto Regressive model

ARIMA (0,1,0) - Random Walk model

ARIMA (1,1,0) - Differenced first order Auto Regression model

ARIMA (0,1,1) without constant - Simple Exponential Smoothing

ARIMA (0,1,1) with constant - Simple Exponential Smoothing with growth

ARIMA (0, 2, 1) OR (0, 2, 2) - Linear Exponential Smoothing

ARIMA (1,1,2) without constant - Damped-Trend linear Exponential smoothing.

Adaptive Smoothing Model and The Non linear model is fitted for the temperature data of India.

The forms of general Auto Regressive Integrated Moving Average (ARIMA) Models are in the following form.

ARIMA (p, d, q):

p is the number of auto regressive terms,

d is the number of non seasonal differences needed for stationarity, and

q is the number of lagged forecast errors in the prediction equation

The Nine models are as follows.

(a) ARIMA(1,0,0) = First-order Auto Regressive model:

The forecasting equation in this case is

$$Y_t = c + \phi_1 Y_{t-1} + e_t$$

or $$\hat{Y}_t = \mu + \phi_1 Y_{t-1}$$

This is also called AR (1) model. Observation Y_t depends on Y_{t-1} and the value of the auto regression coefficient ϕ_1 is restricted to lie between -1 and +1.

If ϕ_1 is positive and less than 1 in magnitude (it must be less than 1 in magnitude if Y is stationary), the model describes mean-reverting behavior in which next period's value should be predicted to be ϕ_1 times as far away from the mean as this period's value. If ϕ_1 is negative, it predicts mean-reverting behavior with alternation of signs, i.e., it also predicts that Y will be below the mean next period if it is above the mean this period.

(a) ARIMA(0,1,0) = Random Walk:

If the series Y is not stationary, the simplest possible model for it is a random walk model, which can be considered as a limiting case of an AR(1) model in which the autoregressive coefficient is equal to 1, i.e., a series with infinitely slow mean reversion. The prediction equation for this model can be written as:

$$\hat{Y}_t - Y_{t-1} = \mu \quad \text{or} \quad \hat{Y}_t = \mu + Y_{t-1}$$

Where the constant term is the average period-to-period change in Y. This model could be fitted as a no-intercept regression model in which the first difference of Y is the dependent variable. Since it includes a non seasonal difference and a constant term, it is classified as an "ARIMA (0,1,0) model with constant." The random-walk-without-drift model would be an ARIMA (0,1,0) model without constant

(b) ARIMA(1,1,0) = Differenced first-order Auto Regressive model:

If the errors of a random walk model are auto correlated, perhaps the problem can be fixed by adding one lag of the dependent variable to the prediction equation i.e., by regressing the first difference of Y on itself lagged by one period. This would yield the following prediction equation:

$$\hat{Y}_t - Y_{t-1} = \mu + \phi_1 (Y_{t-1} - Y_{t-2})$$

$$\hat{Y}_t - Y_{t-1} = \mu$$

This can be rearranged to

$$Y_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2})$$

This is a first-order autoregressive model with one order of non seasonal differencing and a constant term i.e., an ARIMA (1, 1, 0) model.

(c) ARIMA(0,1,1) without constant = Simple Exponential Smoothing:

Another strategy for correcting auto correlated errors in a random walk model is suggested by the simple exponential smoothing model. The random walk model does not perform as well as a moving average of past values. In other words, rather than taking the most recent observation as the forecast of the next observation, it is better to use an average of the last few observations in order to filter out the noise and more accurately estimate the local mean. The simple exponential smoothing model uses an exponentially weighted moving average of past values to achieve this effect. The prediction equation for the simple exponential smoothing model can be written in a number of mathematically equivalent forms, one of which is called "error correction" form, in which the previous forecast is adjusted in the direction of the error it made:

$$\hat{Y}_t = Y_{t-1} + \alpha e_{t-1}$$

Because $e_{t-1} = Y_{t-1} - \hat{Y}_{t-1}$ by definition, this can be rewritten as:

$$\hat{Y}_t = Y_{t-1} - (1-\alpha)e_{t-1}$$

$$\hat{Y}_t = Y_{t-1} - \phi_1 e_{t-1}$$

Which is an ARIMA (0, 1, 1) without constant forecasting equation with $\theta_1 = 1-\alpha$. This means that we can fit a simple exponential smoothing by specifying it as an ARIMA (0, 1, 1) model without constant, and the estimated MA (1) coefficient corresponds to 1 - alpha in the SES formula.

(d) ARIMA(0,1,1) with constant = Simple Exponential Smoothing with growth:

By implementing the SES model as an ARIMA model. First of all, the estimated MA (1) coefficient is allowed to be negative: this corresponds to a smoothing factor larger than 1 in an SES model, which is usually not allowed by the SES model-fitting procedure. Second, we have the option of including a constant term in the ARIMA model if you wish, in order to estimate an average non-zero trend. The ARIMA (0,1,1) model with constant has the prediction equation:

$$\hat{Y}_t = \mu + Y_{t-1} - \theta_1 e_{t-1}$$

The one-period-ahead forecasts from this model are qualitatively similar to those of the SES model, except that the trajectory of the long-term forecasts is typically a sloping line (whose slope is equal to mu) rather than a horizontal line.

(e) ARIMA(0,2,1) or (0,2,2) without constant = Linear Exponential Smoothing:

Linear exponential smoothing models are ARIMA models which use two non seasonal differences in conjunction with MA terms. The second difference of a series Y is not simply the difference between Y and itself lagged by two periods, but rather it is the first difference of the first difference--i.e., the change-in-the-change of Y at period t. Thus, the second difference of Y at period t is equal to $(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - {}_2Y_{t-1} + Y_{t-2}$. A second difference of a discrete function is analogous to a second derivative of a continuous function: it measures the "acceleration" or "curvature" in the function at a given point in time.

The ARIMA (0,2,2) model without constant predicts that the second difference of the series equals a linear function of the last two forecast errors:

$$\hat{Y}_t - 2Y_{t-1} + Y_{t-2} = -\theta_1 e_{t-1} - \theta_2 e_{t-2}$$

This can be rearranged as

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

Where θ_1 and θ_2 are the MA (1) and MA (2) coefficients. This is a general linear exponential smoothing model, essentially the same as Holt's model, and Brown's model is a special case. It uses exponentially weighted moving averages to estimate both a local level and a local trend in the series. The long-term forecasts from this model converge to a straight line whose slope depends on the average trend observed towards the end of the series.

(f) ARIMA(1,1,2) without constant = Damped-Trend linear Exponential Smoothing:

$$\hat{Y}_t = Y_{t-1} + \phi_1(Y_{t-1} - Y_{t-2}) - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

This model is illustrated for accompanying slides on ARIMA models. It extrapolates the local trend at the end of the series but flattens it out at longer forecast horizons to introduce a note of conservatism, a practice that has empirical support. It is generally advisable to stick to models in which at least one of p and q is no larger than 1.

(g) Adaptive Smoothing Model

The single exponential smoothing forecasting model requires the specifications of an α value and it has been shown that the mean absolute percentage error (MAPE) and Mean Square Error (MSE) measures depends on this choice. Adaptive Response Rate Single Exponential Smoothing (ARRSES) may have an advantage over Single Exponential Smoothing (SES), it allows the value of α to be modified in a controlled manner, as changes in the pattern of data occur. This characteristic seems attractive when hundreds (or) thousands of items require forecasting.

The basic equation for forecasting with the method of ARRSES is similar to equation

$$F_{t+1} = \alpha_t Y_t + (1-\alpha_t)F_t$$
$$\alpha_{t+1} = |A_t / M_t|$$
$$A_t = \beta e_t + (1-\beta)A_{t-1}$$
$$M_t = \beta|e_t| + (1-\beta)M_{t-1}$$
$$e_t = Y_t - F_t .$$

here β is a parameter between 0 and 1 and mod || denotes absolute values.

In the equation A_t denotes a smoothed estimate of forecast error and is calculated as a weighted average of A_{t-1} and the last forecasting error $€_t$.

M_t denotes a smoothed estimate of the absolute forecast error; being calculated as a weighted average of M_{t-1} and the last absolute forecasting error $|€_t|$.

A_t and M_t gives single exponential smoothing estimates themselves.

$\alpha_{t+1} = |A_t / M_t|$ indicates that the value of α_t to be used for forecasting period (t+2) is defined as an absolute value of the ratio of A_t and M_t. instead of α_{t+1}, we could have used α_t in the above equation . We prefer α_{t+1} because ARRSES is often too responsive to changes, thus using α_{t+1}, we introduce a small log of one period, which allows the system to "settle" a little and forecast in a more conservative manner.

(h) Non Linear Model

A non linear model fitted to the data is

$$y_t = y_{t-1} e^{(\beta_0 + \beta_1 t)} e_t \qquad (1)$$

Where y_t is the time series observation at time point t.

y_{t-1} is the time series observations at time period t-1.

β_0 is constant

β_1 is the constant for time.

e_t is the error term.

Take log on both sides for equation (1)

$$\log y_t = \log y_{t-1} + \beta_0 + \beta_1 t$$
$$Y_t = \gamma_0 + \beta_1 t \qquad (2)$$

Where $Y_t = \log y_t$

$$\gamma_0 = \log y_{t-1} + \beta_0$$

$\beta_1 = \beta_1$.

t = time

Eq (2) is straight line

$$Y_t = \hat{\gamma}_0 + \beta_1 t$$
$$where\ \hat{\gamma}_0 = \log y_{t-1} + \beta_0$$
$$\Rightarrow \hat{\beta}_0 = \hat{\gamma}_0 - \log y_{t-1}$$

The fitted equation is $y_t = y_{t-1} e^{(\beta_0 + \beta_1 t)} e_t$

IV. RESULTS AND DISCUSSIONS

For temperature data of India, we are fitted all nine models for minimum and maximum temperature data to the years 2000 to 2010 are as follows.

Table -1

SNO	Model used for minimum and maximum temperature data	The equation for the model
1	ARIMA(1,0,0)	$\hat{Y}_t = \mu + \phi_1 Y_{t-1}$
2	ARIMA(0,1,0)	$\hat{Y}_t = \mu + Y_{t-1}$
3	ARIMA(1,1,0)	$Y_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2})$
4	ARIMA(0,1,1)without constant	$\hat{Y}_t = Y_{t-1} - \phi_1 e_{t-1}$
5	ARIMA(0,1,1) with constant	$\hat{Y}_t = \mu + Y_{t-1} - \theta_1 e_{t-1}$
6	ARIMA(0,2,1)or(0,2,2) without constant	$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
7	ARIMA(1,1,2) without constant	$\hat{Y}_t = Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}) - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
8	Adoptive smoothing model	$F_{t+1} = \alpha_t Y_t + (1-\alpha_t)F_t$, $\alpha_{t+1} = \lvert A_t / M_t \rvert$, $A_t = \beta e_t + (1-\beta)A_{t-1}$, $M_t = \beta \lvert e_t \rvert + (1-\beta)M_{t-1}$, $e_t = Y_t - F_t$
9	Non linear model	$y_t = y_{t-1} e^{(\beta_0 + \beta_1 t)} e_t$

The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) for minimum temperature values and maximum temperature values for nine models are shown in table-2.

Table 2

S.NO	Model	Minimum		Maximum	
		MAE	*RMSE*	*MAE*	*RMSE*
1	ARIMA(1,0,0)	3.298	3.522	2.617	2.835
2	ARIMA(0,1,0)	4.499	4.832	3.135	4.008
3	ARIMA(1,1,0)	4.514	4.890	3.073	3.787
4	ARIMA(0,1,1)without constant	3.282	3.799	2.260	2.855
5	ARIMA(0,1,1) with	3.295	4.026	2.211	2.903

	constant				
6	ARIMA(0,2,1)or(0,2,2) without constant	2.058	2.721	2.502	3.235
7	ARIMA(1,1,2) without constant	2.777	3.360	1.930	2.479
8	Adoptive smoothing model	3.514	6.024	3.296	4.083
9	Non linear model	4.015	4.098	2.851	3.345

In the above table, the minimum Mean Absolute Error (MAE) for minimum temperature data is 2.213 to the model ARIMA(0,2,1)or(0,2,2) without constant and for maximum temperature data is 2.175 for the model ARIMA(1,1,2) without constant. The minimum Root Mean Square Error (RMSE) for minimum temperature data is 2.804 to the model ARIMA(0,2,1)or(0,2,2) without constant and for maximum temperature data is 2.835 to the model ARIMA(1,0,0).

Comparison of temperature fitted model of maximum data with minimum data using MAE and RMSE as given in table-3

Table - 3

SNO	Model	MAE	RMSE
1	ARIMA(1,0,0)	+	+
2	ARIMA(0,1,0)	+	+
3	ARIMA(1,1,0)	+	+
4	ARIMA(0,1,1)without constant	+	+
5	ARIMA(0,1,1) with constant	+	+
6	ARIMA(0,2,1)or(0,2,2) without constant	-	-
7	ARIMA(1,1,2) without constant	+	+
8	Adoptive smoothing model	+	+
9	Non linear model	+	+

Table 3 represents the positive and negative signs. Here the comparison is made-up of MAE values of maximum temperature data with minimum temperature data. The RMSE values of maximum temperature data with minimum temperature data. Here the positive (+) sign indicates increasing trend and negative (–) sign indicates decreasing trend.

The test group is from 2011 to 2016. The equations for maximum and minimum Temperature data are fitted for nine models as given in table

Table 4

SNO	Model used for minimum and maximum temperature data	The equation for the model
1	ARIMA(1,0,0)	$\hat{Y}_t = \mu + \phi_1 Y_{t-1}$
2	ARIMA(0,1,0)	$\hat{Y}_t = \mu + Y_{t-1}$
3	ARIMA(1,1,0)	$Y_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2})$
4	ARIMA(0,1,1)without constant	$\hat{Y}_t = Y_{t-1} - \phi_1 e_{t-1}$
5	ARIMA(0,1,1) with constant	$\hat{Y}_t = \mu + Y_{t-1} - \theta_1 e_{t-1}$
6	ARIMA(0,2,1)or(0,2,2) without constant	$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
7	ARIMA(1,1,2) without constant	$\hat{Y}_t = Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}) - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
8	Adoptive smoothing model	$F_{t+1} = \alpha_t Y_t + (1-\alpha_t) F_t$, $\alpha_{t+1} = \lvert A_t / M_t \rvert$, $A_t = \beta e_t + (1-\beta) A_{t-1}$, $M_t = \beta \lvert e_t \rvert + (1-\beta) M_{t-1}$, $e_t = Y_t - F_t$
9	Non linear model	$y_t = y_{t-1} e^{(\beta_0 + \beta_1 t)} e_t$

The Mean Absolute Error(MAE) and Root Mean Square Error(RMSE) for minimum and maximum temperature values for nine models during 2011 to 2016 is shown in table-5.

Table 5

SNO	Model	Minimum		Maximum	
		MAE	*RMSE*	*MAE*	*RMSE*
1	ARIMA(1,0,0)	3.291	3.741	2.540	2.940
2	ARIMA(0,1,0)	4.538	5.014	3.079	4.041
3	ARIMA(1,1,0)	4.543	5.143	2.985	3.921
4	ARIMA(0,1,1)without constant	3.839	4.374	2.658	3.258
5	ARIMA(0,1,1) with constant	3.848	4.496	2.664	3.346
6	ARIMA(0,2,1)or(0,2,2) without constant	2.120	2.862	2.603	3.376
7	ARIMA(1,1,2) without constant	2.844	3.609	2.259	3.302
8	Adoptive smoothing model	2.862	3.353	2.280	3.076
9	Non linear model	4.069	4.136	2.814	3.280

From the table-5, the minimum Mean Absolute Error (MAE) for minimum temperature data is 2.120 to the model ARIMA(0,2,1)or(0,2,2) without constant and for maximum temperature data is 2.259 for the model ARIMA(1,1,2) without constant. The minimum Root Mean Square Error (RMSE) for minimum temperature data is 2.862 to the model ARIMA(0,2,1)or(0,2,2) without constant and for maximum temperature data is 2.940 to the model ARIMA(1,0,0).

Comparisons of MAE and RMSE for the temperature of maximum data with minimum data for 2011-2016 are given in table-6.

Table 6

S.NO	Model	MAE	RMSE
1	ARIMA(1,0,0)	+	+
2	ARIMA(0,1,0)	+	+
3	ARIMA(1,1,0)	+	+
4	ARIMA(0,1,1)without constant	+	+
5	ARIMA(0,1,1) with constant	+	+
6	ARIMA(0,2,1)or(0,2,2) without constant	-	-
7	ARIMA(1,1,2) without constant	+	+
8	Adoptive smoothing model	+	+
9	Non linear model	+	+

Table-6 represents the positive and negative signs. Here the comparison is made-up of MAE values of maximum temperature data with minimum temperature data. Also the RMSE values of maximum temperature data with minimum temperature data. Here the positive (+) sign indicates increasing trend and negative (-) sign indicates decreasing trend.

V. CONCLUSIONS

In this paper we fitted nine models for temperature data in INDIA. The nine models are First order Auto Regressive model, Random Walk model, Differenced first order Auto Regression model, Simple Exponential Smoothing without constant, Simple Exponential Smoothing with growth, Linear Exponential Smoothing without constant, Damped-Trend linear Exponential smoothing without constant, Adaptive Smoothing Model, and the Non linear model. An error criteria called Mean Absolute Error and Root Mean Square Error are listed out for choosing the best model among different models using temperature data for minimum and maximum. The equations for maximum and minimum Temperature data are fitted for nine models and are as follows.

SNO	Model used for minimum and maximum temperature data	The equation for the model
1	ARIMA(1,0,0)	$\hat{Y}_t = \mu + \phi_1 Y_{t-1}$
2	ARIMA(0,1,0)	$\hat{Y}_t = \mu + Y_{t-1}$
3	ARIMA(1,1,0)	$Y_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2})$
4	ARIMA(0,1,1)without constant	$\hat{Y}_t = Y_{t-1} - \phi_1 e_{t-1}$
5	ARIMA(0,1,1) with constant	$\hat{Y}_t = \mu + Y_{t-1} - \theta_1 e_{t-1}$
6	ARIMA(0,2,1)or(0,2,2) without constant	$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
7	ARIMA(1,1,2) without constant	$\hat{Y}_t = Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}) - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
8	Adoptive smoothing model	$F_{t+1} = \alpha_t Y_t + (1-\alpha_t) F_t$, $\alpha_{t+1} = \lvert A_t / M_t \rvert$, $A_t = \beta e_t + (1-\beta) A_{t-1}$, $M_t = \beta \lvert e_t \rvert + (1-\beta) M_{t-1}$, $e_t = Y_t - F_t$
9	Non linear model	$y_t = y_{t-1} e^{(\beta_0 + \beta_1 t)} e_t$

The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are listed out for choosing best model , among different models using temperature data for minimum and maximum are as follows.

SNO	Model	Minimum		Maximum	
		MAE	*RMSE*	*MAE*	*RMSE*
1	ARIMA(1,0,0)	3.291	3.741	2.540	2.940
2	ARIMA(0,1,0)	4.538	5.014	3.079	4.041
3	ARIMA(1,1,0)	4.543	5.143	2.985	3.921
4	ARIMA(0,1,1)without constant	3.839	4.374	2.658	3.258
5	ARIMA(0,1,1) with constant	3.848	4.496	2.664	3.346
6	ARIMA(0,2,1)or(0,2,2) without constant	2.120	2.862	2.603	3.376
7	ARIMA(1,1,2) without constant	2.844	3.609	2.259	3.302
8	Adoptive smoothing model	2.862	3.353	2.280	3.076
9	Non linear model	4.069	4.136	2.814	3.280

The best model for minimum temperature data for seasons from the above table is ARIMA(0,2,1)or(0,2,2) without constant and the equation for this model is

$$\hat{Y}_t = 2Y_{t-1} - Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

The best model for maximum temperature data for seasons from the above table using RMSE is ARIMA (1,0,0) or first order Auto Regressive model and the equation for this model is $\hat{Y}_t = \mu + \phi_1 Y_{t-1}$

By using the best models, the forecasted values for four seasons for 4 years are as follows.

year	Quarter/ season	Minimum forecasted temperature	Maximum forecasted temperature
1	1	19.98	27.96
	2	17.56	28.72
	3	22.62	29.26
	4	17.71	29.98
2	1	18.65	28.26
	2	18.33	28.98
	3	22.73	29.51
	4	17.30	30.23
3	1	18.43	28.52
	2	18.17	29.22
	3	22.09	29.76
	4	17.37	30.49

4	1	18.73	28.78
	2	18.45	29.49
	3	22.31	30.03
	4	17.12	30.73

REFERENCES

[1] Grant Foster, Stefan Rahmstorf, "*Global Temperature Evolution 1979-2010*", Environ. Res. Lett.6, 044022,2011.

[2] Grant Foster, Niamh Cahill, Stefan Rahmstorf, , "*Global Temperature Evoluation: Recent Trends and Some Pitfalls*" , Environmental Research Letters, 12 , 054001,2017.

[3] Chonggunag Li, Lu Zhang, Tao Xiong , Yukun Bao, Zhongyi Hu, "*A Combination Method for Interval Forecasting of Agricultural Commodity Futures Prices"*, Knowledge-Based Systems ,77,92-102,2015.

[4] Markus Glaser, Martin Weber, Thomas Langer, " *True Overconfidence in Interval Estimates: Evidence Based on a New Measure of Miscalibration"*, journal of Behavioral Decision Making, 26:405-417,2012.

[5] Aidan McDermott, Francesca Dominici, Jonathan M. Samet, Scott L.Zeger, "*On the Use of Generalized Additive Models in Time Series Studies of Air Pollution and Health",* American Journal of Epidemiology, vol. 156, No. 3,2002.

[6] Heather A. Walton, H. Ross Anderson, Inga C. Mills, Richard W. Atkinson, "*Fine Paricle Componens and Health- a Systematic review and Meta-Analysis of Epidemological Time Series Studies of Daily Mortality and Hospital Admissions*", Journal of Exposure Science and Environmental Epidemiology,25,208-214,2015.

[7] Alesssandro Baccini, Alexndra Tyukavina, Belinda Arunarwati Margono, Ilona Zhuravleva, Matthew C Hansen Peter Potapov, Scott Goetz, Svetlana Turubanova, "*Mapping and Monitoring Deforestation and Forest Degradation in Sumatra (Indonesia) using Land sat time series data sets from 1990 to 2010*", Environ. Res. Lett.7, 034010, 2012.

[8] http//:www.data.gov.in.

AUTHORS PROFILE

Mr. A.Srinivasulu.pursed Bachelor of Science from Sri Venkateswara University, Tirupati in 2002 and Master of Science from Sri Venkateswara University, Tirupati in 2004. He is presently pursuing Ph.D. in the Department of Statistics, Sri Venkateswara University, Tirupati, Andhra Pradesh, India. He has published three papers in National and International Journals. He has presented two papers in International conference and two in National Conference. He has 14 years of teaching experience.

Dr.B.Sarojamma received Ph.D. degree in statistics in 2006 from Sri Venkateswara University, Tirupati, Andhra Pradesh, India. She taught the subjects Time series and Forecasting, Operations Research, Distribution Theory, Sampling Techniques and Biostatistics, etc. to the students of M.Sc. Statistics, Applied Statistics, Statistics with Computer applications, M.Sc.Home science, M.Sc. Aquaculture, Biotechnology and Bio-informatics. She has given several invited lectures in national/International workshops/Conferences. She has published more than 62 research papers in internationall reputed journals and presented more than 83 papers at National/International Conferences/Simposia. She is the Life Member in manyProfessional Bodies like Indian Society for Probability and Statistics (ISPS), Andhra Pradesh Society for Mathematics Sciences Hyderabad, International Indian Statistics Association (IISA) ,Cochin,Indian Mathematical Society (IMS), Aurangabad and Indian Statistical Association(AISA), Pune. At present she is working as an Assistant Professor in the Deparment of Statistics, S.V.University, Tirupati.

