



# International Journal of Emerging Technologies in Computational and Applied Sciences(IJETCAS)

(Open Access, Double Blind Peer-reviewed, Refereed and Indexed Journal)

[www.iasir.net](http://www.iasir.net)

## Visualizing the Future on the Incidents of Dengue Fever in India by Using the Techniques of Time Series Analysis

T. Suneetha, G. Jayalakshmi, V. Kavitha, S.V.N.S. G. Mounika, M. Kavya, G. Lakshmi Ramya, S. Madhavi, P. Madhusa

Department of Statistics,  
Sri Padmavati Mahila Visvavidyalam,  
Tirupati – 517502, A.P., INDIA.

**Abstract:** Time series analysis had been increasingly used to forecast the number of dengue fever in many studies. This study was conducted to determine trend and forecasting the dengue cases in the years, the three successive years from 2018 to 2020. A time series analysis was conducted by secular trend values to the live data by fitting the trend line using least square techniques.

**Key Words:** time series, dengue fever, forecasting.

### I. Introduction

Planning for the future is an integral part of any business activity whether it is short term or long term, planning requires the ability to look ahead or look into the future. That is look into an area where it is impossible to see as clearly as one has to undertake the responsibility for looking ahead and forecast the future levels of activity on which planning can be used. Even though many factors have to be taken into consideration for visualizing the future one of the factors that an investigator always depends upon making reasonable guesses using the data of the past. Such data are provided by the time series. A time series as “The numerical information about a particular subject presented in historical form”.

It customary to classify the different influences that together give rise to apparently irregular fluctuations of a time series into four groups. These four groups more commonly known as components of time series are: secular trend, seasonal variations, cyclical variation and irregular variation. The secular (or long term) trend of a time series can be defined as the underlying smooth or regular movement of the time series over a long period of time. Variations that are periodic in nature and that reoccur regularly within a period of one year or less are known as seasonal variations. cyclical variations too, like seasonal variations a reperiodic in nature and involve oscillatory (up and down) movements. But, unlike seasonal variations, cyclical variations extend over a period of several years and don't secure with any predictable regularity. Any movement in time series that cannot be attributed to secular trend or seasonal or cyclical variation is termed as irregular or residual variation.

### II. Methodology

#### TREND ANALYSIS:

By definition, trend analysis is based on historical price movements. As a result, it may seem to fit best under the jurisdiction of technical analysis. Technical analysis looks at historical trends and changes in price to determine the future direction prices. By contrasts, fundamental analysis looks at changes in the performance of an asset, such as earnings or revenue. That said, fundamental analysts can also look for trends in earnings per share and revenue growth. If earnings have grown for the past four quarters, this represents a positive trend. However, if earnings have declined for the past four quarters, it represents a negative trend.

A trend analysis is an aspect of technical analysis that tries to predict the future movement of a stock based on past data. Trend analysis is based on the idea that what has happened in the past gives traders an idea of what will happen in the future. There are three main types of trend.

- Short
- Intermediate
- Long

A trend analysis is an aspect of technical analysis that tries to predict the future movement of a stock based on past data. Trend analysis is based on the idea that what has happened in the past gives traders an idea of what will happen in the future.

MS EXCEL:

A spreadsheet is essentially a matrix of rows and columns. Consider a sheet of paper on which horizontal and vertical lines are drawn to yield a rectangular grid. The grid namely a cell, is a result of the intersection of a row with a column. Such a structure is called a spreadsheet.

A spreadsheet package contains electronic equivalent of a pen, an eraser and large sheet of paper with vertical and horizontal line to give rows and columns. The cursor position uniquely shown in dark mode indicates where the pen is currently pointing. We can enter text or numbers at any position on the worksheet. We can enter a formula in a cell where we want to perform a calculation facility jumps into action each time we update the cell contents with new data.

MS-Excel is the most powerful spreadsheet package brought by Microsoft. The three main components of this package are

- ❖ Electronic spreadsheet
- ❖ Database management
- ❖ Generation of charts

Each workbook provides three worksheets with facilities to increase the number of sheets. Each sheet provides 16384 (XFD) columns and 10,48,576 rows to work with. Though the spreadsheet packages were originally designed for accountants, they have become popular with almost everyone working with figures. Sales executives, book-keepers, officers, students, research scholars, investors bankers etc, almost any one find some form of application for it.

### METHOD OF LEAST SQUARES:

#### STRAIGHT LINE TREND:

The straight line trend is represented by the equation:  $y_t = a + bx$ , where 'y<sub>t</sub>' are the trend values 'a' is the y intercept or the value of y when x=0, 'b' represents the slope of line or the amount by which y variable will change if x is increased or decreased by one unit, and x variable is the time period.

In order to determine the value of the constants 'a' and 'b', following two equations, called normal equations, are to be solved

$$\sum X_i = Na + b \sum X_i$$

$$\sum X_i y_i = a \sum X_i + b \sum X_i^2$$

Where N is the number of time periods.

While fitting the trend line by the method of least squares three things are always specified.

1. Which year was selected as origin?
2. What is unit of time represented by x.i.e.; whether x represents 1 month ½ month, ½ year or 1 year?
3. In what kind of units is y being measured?

For ease in calculation, generally the mid-point of the given time series is taken as origin so that negative values of the first half of the series balance out the positive values of the second half and we get  $\sum X_i = 0$

When  $\sum X_i = 0$  the two normal equations will be reduced to

$$\sum Y_i = Na \text{ and } \sum X_i Y_i = b \sum X_i^2, \text{ this means}$$

$$a = \sum Y_i / N = \bar{y} \text{ and } b = \sum X_i Y_i / \sum X_i^2$$

### TREND FUNCTION TO CREATE TREND LINES ON EXCEL CHARTS:

Use the TREND function to calculate the data of a trend line. This shows you the data in the worksheet so you can paste it into the chart and use other functions to control its appearance and behaviour. TREND is an array function. This is a function that uses a collection data as input and produces a collection of data as an output. You must enter an array function with a special key stroke, which I'll describe in just a bit. You can see in the formula bar in the figure that after you have entered that TREND function as a array it appears in the formulas bar differently from a normal formula. It appears with {} braces around it. This indicates it is an array and that one cell cannot be changed, all cells in the array must be changed simultaneously.

Excel has two functions it uses for calculating trends and filing the data points into a worksheet. TRENDS and GROWTH. Use TRENDS for linear or straight-line regressions. The TREND function is an array function. You must enter it into multiple cells all of one time, so follow these steps exactly:

1. Enter the title SL Trend. For straight line trend in cell
2. Select the cells where you want the Trend function results. The selection in which you enter TREND must match the width and height of the data used by TREND. Notice the selection in the figure is the same size and shape as the data cell above it. Array functions are entered into all cells of an array at one time so you must select them all prior to entering the function.
3. As you create the formula. With select it will appear as though y or are only creating it in cell =TREND(D5:D21,E5:E21)

The TREND function requires.=TREND(Known Y's, Known X's)

In this example the known X's are evenly spaced dates. In this case they are data values. The Y's are data values. All the cells from should still be selected.

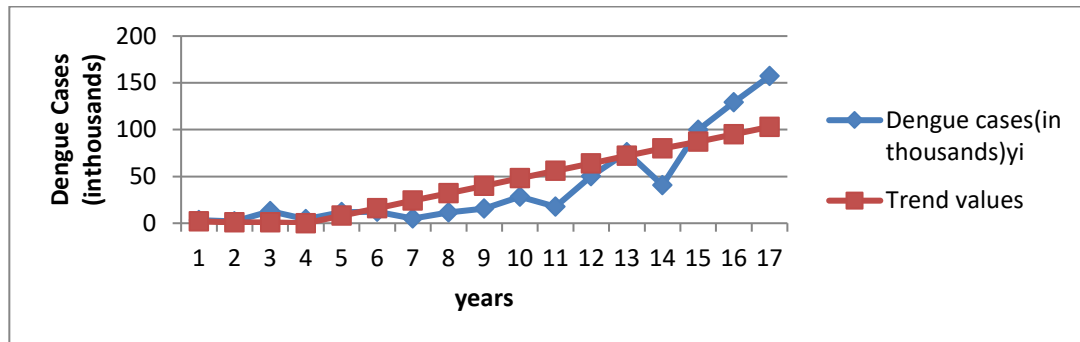
4. Enter the array formula by pressing

Shift+ctrl+Enter. (Hold down shift+ctrl, then press enter). This enters the formula into all selected cells. You'll see the trend data in row like in the figure above. But now you have to add the trend line data to the Excel chart.

1.select cells, right click and choose copy.

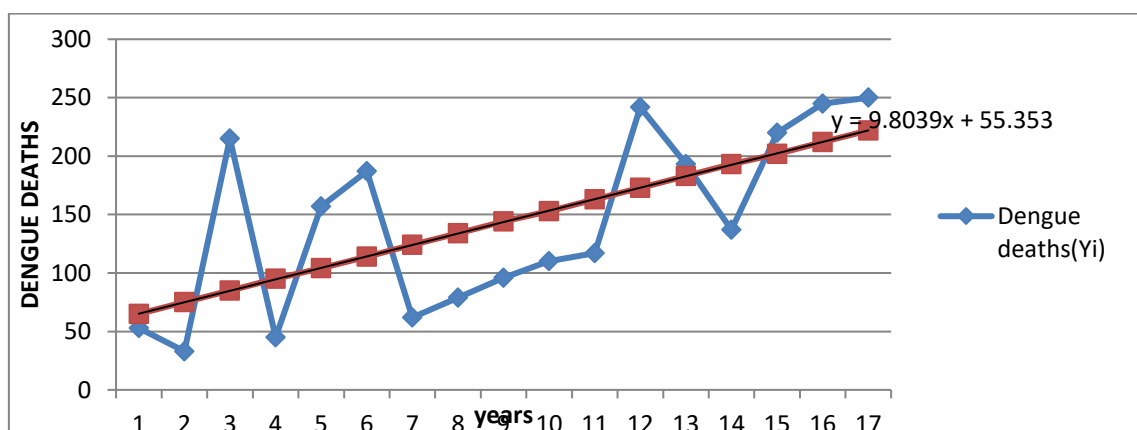
2.select the chart, right click and choose , paste the trend line will appear over the data line.

FIGURE 1: TREND LINE FOR DENGUE CASES IN INDIA



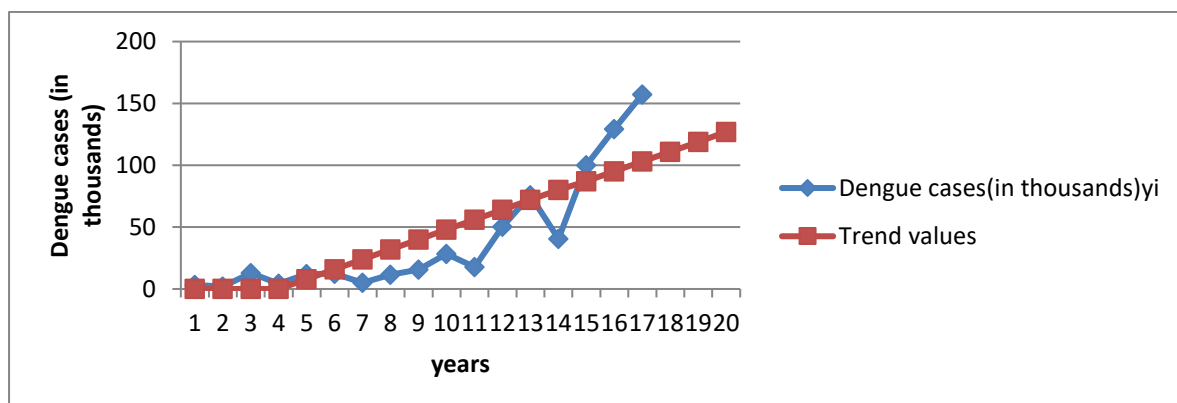
Obviously noticed from the above figure that using the Least Squares techniques the cases due to Dengue fever in India is moderately fluctuating around the fitted trend line and attaining the maximum dengue cases in 2017. Also, it clear that there exist an increasing trend values.

FIGURE 2: TREND LINE FOR DENGUE DEATHS IN INDIA



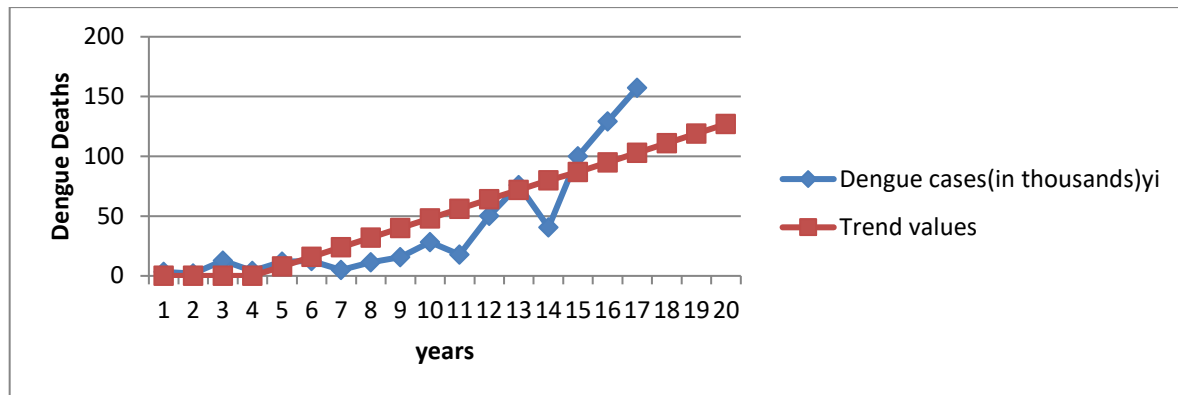
It is apparently observed from the above figure where the Dengue deaths in India are considerable increasing over a period from 2001 to 2017 and attained maximum in the year 2017. Also, the trend values show that there exist perfect increase in the said epidemic and reaching maximum in the same year 2017.

FIGURE 3 : FORECASTING TREND LINE FOR DENGUE CASES IN INDIA



It is apparently observed from the above figure where the Dengue cases in India are considered increasing over a period from 2001 to 2017 and attained maximum in the year 2017. Also, the trend values show there exist perfect increase in the said epidemic and reaching maximum in the same year 2017. Further, it is also noticed that there exist even an increasing trend line for the projected years of 2018 to 2020.

FIGURE 4: FORECASTING TREND LINE FOR DENGUE DEATHS IN INDIA



It is apparently observed from the above figure where the Dengue deaths in India are considered increasing over a period from 2001 to 2017 and attained maximum in the year 2017. Also, the trend values show there exist perfect increase in the said epidemic and reaching maximum in the same year 2017. Further, it is also noticed that there exist even an increasing trend line for the projected years of 2018 to 2020.

### III. Empirical Work

#### FORECASTING:

Meaning forecasting is a process of predicting or estimating the future based on past and present data. Forecasting provides information about the potential future events and their consequences for the organizations. It may not reduce the complications and uncertainty of the future. However, it increases the confidence of the management to make important decisions. Forecasting is the basis of premising. Forecasting uses many statistical techniques. Therefore, it is also called as statistical analysis.

#### SIMPLE EXPONENTIAL SMOOTHING:

If you have a time series that can be described using an additive model with constant level and no seasonality, you can use simple exponential smoothing to make short-term, forecast.

#### HOLT'S EXPONENTIAL SMOOTHING:

If you have a time series that can be described using an additive model with increasing or decreasing trend and no seasonality we can use holt's exponential smoothing to make short-term forecasts.

#### WINTER'S THREE PARAMETERS LINEAR AND SEASONAL EXPONENTIAL SMOOTHING:

If you have a time series that can be described using an additive model with increasing or decreasing trend and seasonality, you can use holt- winter's exponential smoothing to make short-term forecasts.

### IV. GRAPHICAL VIEWS

#### EXPONENTIAL SMOOTHING:

Here, we have alpha value that is smoothing constant and this method is called simple exponential smoothing method which consider other two factors as constant (i.e., seasonality and trend). Double's (Holt's) exponential smoothing and winter's exponential smoothing methods dealing two factors that is trend and seasonality (i.e., beta and gamma). Larger the alpha, closer to the actual data points and vice versa. This method is suitable for forecasting data with no trend or seasonal pattern (alpha = smoothing constant)

Exponential smoothing is another method of quantitative forecasting that some businesses have found useful due to its short-term precision and cost effectiveness. The use of exponential smoothing as a means of forecasting is slightly more complex than traditional moving average forecasts. Specifically, exponential smoothing combines a prior forecast with weighted error measurement. This error measurements comes from the difference between the prior forecast and the actual observation at the time of the prior forecast. It also requires managers to arbitrarily select a value for alpha. It must be stated that although any value of alpha between 0 and 1.0 inclusive is acceptable, certain values may yield better forecasts. By rewriting the basic exponential smoothing model shown in(3) to resemble(4) insight into good values alpha can be obtained.

The formula that the exponential smoothing approach is based on can be expressed as

$$F_{t+1} = F_t + \alpha \cdot e_t$$

Given that,  $t$  is time,  $F_t$  is the forecast at time  $t$  and  $\alpha$  is a smoothing constant (a value between 0 and 1 inclusive), and  $e_t = D_t - F_t$ , where  $D_t$  denotes the actual observation at time  $t$ , equation can be written as follows :

$$F_{t+1} = F_t + \alpha(D_t - F_t)$$

From it can be seen that the new forecast,  $F_{t+1}$  is equal to the previous forecast,  $F_t$  plus an adjustment (i.e.,  $\alpha(D_t - F_t)$ ). Specifically, the forecast for period  $t+1$  is obtained by adjusting the forecast in period  $t$  by a fraction of forecast error (i.e.  $D_t - F_t$ ). If the time series contains substantial variability, a small value of the smoothing constant is preferred so not to over react and adjust forecast too quickly. For a time series with small variability a larger value of alpha is preferred so the forecast can quickly adjust when forecasting errors do occur and therefore allow the model to react faster to changing conditions. Overall, a user should choose an alpha value that minimizes the error of the model.

With some mathematical manipulation, equation can be rearranged to yield to following equations.

$$F_{t+1} = F_t + \alpha(D_t - F_t)$$

$$F_{t+1} = (1 - \alpha)F_t + \alpha D_t$$

Equation describes the forecast at period  $t+1$  as a proportion of the forecast at period  $t$  plus a proportion of the actual observation at period  $t$ . If alpha is 1.0 the forecast at period  $t+1$  is the actual observation,  $D_t$  from period  $t$ . Conversely, if alpha is 0.0 the forecast at period  $t+1$  is the forecast,  $F_t$  from period  $t$ .

#### EXPONENTIAL SMOOTHING USING EXCEL FUNCTIONS:

The data from table1 will be used to demonstrate exponential smoothing. For this example, alpha is arbitrarily set at four levels 0.3, 0.8, 0 and 1. In addition, a common method of boot strapping is used to initialize the first periods forecast. This boot strap method is to assume the first years forecast is the same as the first year's dengue death cases. Using the same data set as presented previously the exponential smoothing forecast is created in EXCEL.

**TABLE 1 : FORECASTING DENGUE CASES IN INDIA**

| year | Dengue cases | Forecast (Trial 1) | Forecast (Trial 2) | Forecast (Trial 3) | Forecast (Trial 4) |
|------|--------------|--------------------|--------------------|--------------------|--------------------|
|      |              | 0.3                | 0.8                | 0                  | 1                  |
| 2001 | 3.306        | 39.8477059         | 39.8477059         | 39.8477059         | 39.84770588        |
| 2002 | 1.926        | 28.8851941         | 10.6143412         | 39.8477059         | 3.306              |
| 2003 | 12.754       | 20.7974359         | 3.66366824         | 39.84771           | 1.926              |
| 2004 | 4.153        | 18.3844051         | 10.9359336         | 39.84771           | 12.754             |
| 2005 | 11.985       | 14.1149836         | 5.50958673         | 39.84771           | 4.153              |
| 2006 | 12.317       | 13.4759885         | 10.6899173         | 39.84771           | 11.985             |
| 2007 | 5.023        | 13.128292          | 11.9915835         | 39.84771           | 12.317             |
| 2008 | 11.476       | 10.6967044         | 6.41671669         | 39.84771           | 5.023              |
| 2009 | 15.535       | 10.9304931         | 10.4641433         | 39.84771           | 11.476             |
| 2010 | 28.292       | 12.3118451         | 14.5208287         | 39.84771           | 15.535             |
| 2011 | 17.744       | 17.1058916         | 14.5208287         | 39.84771           | 28.292             |
| 2012 | 50.222       | 17.2973241         | 25.5377657         | 39.84771           | 17.744             |
| 2013 | 75.808       | 27.1747269         | 45.2851531         | 39.84771           | 50.222             |
| 2014 | 40.571       | 41.7647088         | 69.7034306         | 39.84771           | 75.808             |
| 2015 | 99.913       | 41.4065962         | 46.3974861         | 39.84771           | 40.571             |
| 2016 | 129.166      | 58.9585173         | 89.2098972         | 39.84771           | 99.913             |
| 2017 | 157.22       | 80.0207621         | 121.174779         | 39.84771           | 129.166            |
| 2018 |              | 103.180533         | 150.010956         | 39.84771           | 157.22             |

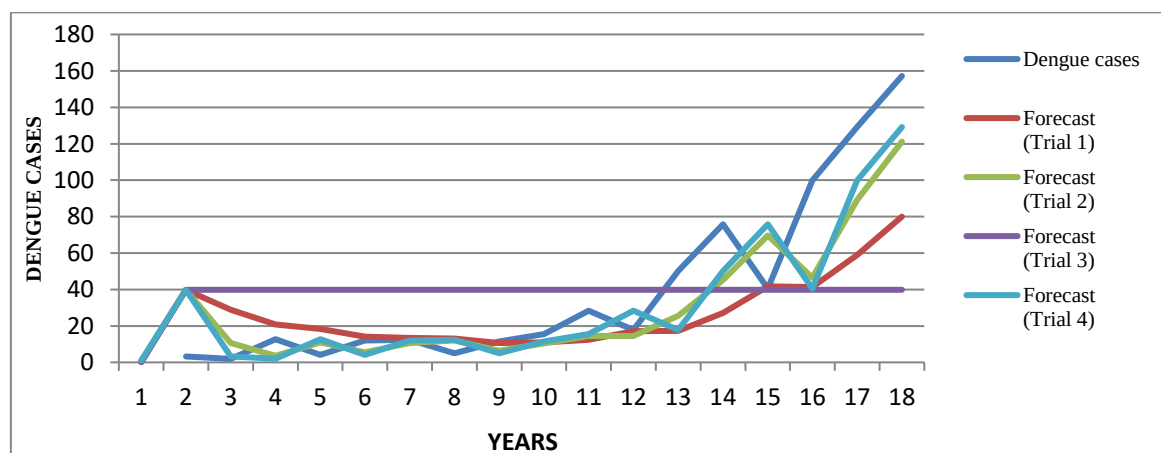


**FIGURE 5: FORECASTING DENGUE CASES IN INDIA**

**TABLE 2: FORECASTING DENGUE DEATH IN INDIA**

| Year | Dengue deaths | forecast (trial 1) | forecast (trial 2) | forecast (trial 3) | Forecast (trial 4) |
|------|---------------|--------------------|--------------------|--------------------|--------------------|
|      |               | 0.3                | 0.8                | 0                  | 1                  |
| 2001 | 53            | 143.5882353        | 143.5882353        | 143.5882           | 143.58824          |
| 2002 | 33            | 116.4117647        | 71.11764706        | 143.5882           | 53                 |
| 2003 | 215           | 91.38823529        | 40.62352941        | 143.5882           | 33                 |
| 2004 | 45            | 128.4717647        | 180.1247059        | 143.5882           | 215                |
| 2005 | 157           | 103.4302353        | 72.02494118        | 143.5882           | 45                 |
| 2006 | 187           | 119.5011647        | 140.0049882        | 143.5882           | 157                |
| 2007 | 62            | 139.7508153        | 177.6009976        | 143.5882           | 187                |
| 2008 | 79            | 116.4255707        | 85.12019953        | 143.5882           | 62                 |
| 2009 | 96            | 105.1978995        | 80.22403991        | 143.5882           | 79                 |
| 2010 | 110           | 102.4385296        | 92.84480798        | 143.5882           | 96                 |
| 2011 | 117           | 104.7069708        | 106.5689616        | 143.5882           | 110                |
| 2012 | 242           | 108.3948795        | 114.9137923        | 143.5882           | 117                |
| 2013 | 193           | 148.4764157        | 216.5827585        | 143.5882           | 242                |
| 2014 | 137           | 161.833491         | 197.7165517        | 143.5882           | 193                |
| 2015 | 220           | 154.3834437        | 149.1433103        | 143.5882           | 137                |
| 2016 | 245           | 174.0684106        | 205.8286621        | 143.5882           | 220                |
| 2017 | 250           | 195.3478874        | 237.1657324        | 143.5882           | 221                |
| 2018 |               | 211.7435212        | 247.4331465        | 143.5882           | 250                |

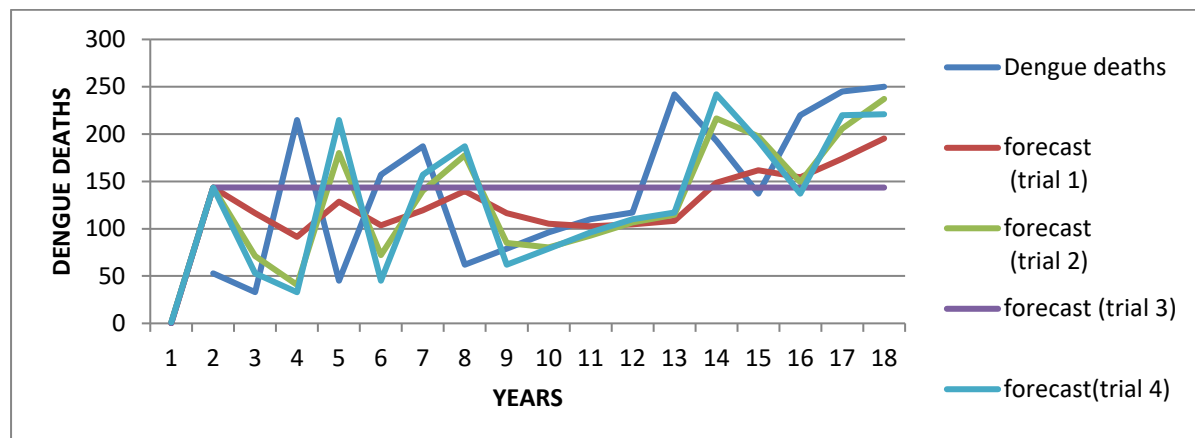


**FIGURE 6: FORECASTING DENGUE DEATHS IN INDIA**

The above figures are constructed to compare forecast values with different alpha points using the exponential smoothing of Ms-Excel. Through this it may be interpreted that for different alpha points the incidents of Dengue fever are showing

Trial 1: forecast relatively smoothing as compared to the actual Dengue fever which as higher variation this is smoothing taking place.

Trial 2: The green line shows the variation more than blue line. So  $\alpha$  has increased ability to smoothing out the dengue cases it was down.

Trial 3: the next period forecast  $\alpha$  was 0 it is very sensitive method the forecast do not change.

Trial 4: first period actual value is same as second period forecast value; second period actual value is same as the third period forecast value as the third period forecast value. This is called naïve forecasting which is highly sensitive forecasting. If  $\alpha = 1$  it is highly sensitive forecasting method like naïve forecasting when previous periods actual becomes the current periods forecasting.

## V. Conclusions

It is predicted during the year 2020, the incidents of Dengue fever will predominantly affect the population where more no. of death cases are expected, hence steps may be taken to prevent in reducing the mosquito habitat and limiting exposure to bites. It is suggested a novel vaccine for Dengue fever as precautionary measure may be applied in the vulnerable area of society.

## VI. References

1. Armstrong, J S, ed. 2001, Principles of Forecasting: A handbook for researchers and Practitioners. Kulwer Academic Publishers.
2. Bhatt S, Gething PW, Brady OJ et al. The global distribution and burden of Dengue, Nature 2013; 496: 504-507.
3. Box, George E P, and Gwilym M Jenkins. 1970 Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
4. Brockwell, Peter J, and Richard A Davis. 2016. Introduction to Time Series and Forecasting. 3<sup>rd</sup>Ed. New York, USA.
5. Chakravarti A, Aror R, Luxemburger C fifty years of Dengue in India. Trans R SOC trop Med Hyg 2012, 106: 273-282.

6. Chaturvedi UC, Nagar R. Dengue and Dengue haemorrhagic fever: India Perspective. J Biosci 2008; 33: 429-441.
7. Durbin J. and Koopman, S.J. (2000), "Time Series Analysis by State Space Methods", Oxford University press, Oxford.
8. Hamilton, E.J. (1960), "Time Series Analysis" Methuen, London.
9. Harvey, A. (1990), "The Econometric Analysis of Time Series", the MIT Press, 2<sup>nd</sup> Ed. Cambridge Mass., 1990, Chap 5.
10. KaleharPrajakta S (2014), "Time Series Forecasting Using Holt – Winters Exponential Smoothing".
11. Lam SK, Burke D, GublerDet al. call for a World Dengue Day. Lancet 2012; 379: 411-412.
12. Rissanen J. (1983), "A Predictive Least Squares Principle", J. Math. Control Info, 3, 211-22.
13. Taylor, J W. 2003, "Exponential smoothing with a Damped Multiplicative Trend" international journal of forecasting 19: 715-25.
14. World health Organization. Dengue Vaccine: WHO position. WKLY Epidemiol Rec 2016; 30: 349-364.

## VII. Acknowledgments

We sincerely acknowledged the help extended by Mr. G. Megha Shyam, Lecturer in Computer Science, Dr. A.E.R. Degree college, Tirupati.