Comparative study of different exponential smoothing models in simulation of meteorological drought: A study on Purulia district, West Bengal, India

Shrinwantu Raha*
Department of Geography, Coochbehar Panchanan Barma University, Coochbehar-736101 (West Bengal), India

Shasanka Kumar Gayen
Department of Geography, Coochbehar Panchanan Barma University, Coochbehar-736101 (West Bengal), India

*Corresponding author. Email: rahashrinwantu@gmail.com

INTRODUCTION

Drought is a slow onset, costliest, recurrent and complex phenomenon (Drumond et al., 2017). It is very difficult to describe its status most of the time because of its’ inherent character (Rossi 2003). The general feature of drought is peculiar, which significantly deviates from other natural hazards and disasters (Zarch et al., 2015) as it may stretch over a month or even years (Singh et al., 2019). Over 4 types of drought (e.g. meteorological drought, hydrological drought, ecological drought and socioeconomic drought), meteorological drought is more common and complex in nature (Stagge et al., 2015). Its complexity further increases when it assimilates with other meteorological and socioeconomic phenomena (Dogan et al., 2012). Drought forecasting and simulation is an essential element in drought risk management (Hagenlocher et al., 2019). Although considerable works have been done in spatio-temporal assessment of drought, still simulation and prediction of drought are relatively new and challenging aspects in weather forecasting (Yves et al., 2020). This research tried to utilize the exponential smoothing models to simulate the meteorological drought, completely new in this field. Thus this research bears the fo two objectives a) To simulate meteorological drought at 1 month lead time using two exponential smoothing models, and b) To make a comparative assessment of two exponential smoothing models to determine which model is a better fit for this context.

MATERIALS AND METHODS

Study area

Purulia is one of the westernmost districts of West Bengal and it is an extended tract of Rarh terrain (Dey et al., 2020). This region is affected by the shortage of rainfall and due to low water holding capacity of the

How to Cite


Keywords: Additive Model, Multiplicative model, Standardized Precipitation Index (SPI)
soil, 50% of its rainfall passes away as runoff (Hazra et al., 2017). Average annual rainfall varies between 1100 to 1500 mm within Purulia. 75% to 90% relative humidity is observed in monsoon season. According to SAFE (2011), Purulia was the worst hit, with agricultural production falling to 27% and around 280,000 hectares of agricultural land was lying vacant facing drought. Most of the families of Purulia (almost 70-75%) were dependent upon paddy production and the drought condition of 2010 had put an end to their only means of livelihood. Therefore effective early warning system of drought is desperately needed to combat the situation in Purulia. Simulation of drought helps to understand the true nature of drought and helps to perceive the severity of drought in the near future. Fig. 1 depicts the location map of the study area.

**Source of data**

Rainfall data was downloaded from Climate Forecast System Reanalysis (CFSR) Soil and Water Assessment daily database and those daily databases were converted into monthly values. CFSR is an open access high resolution coupled atmosphere-ocean-land-surface-sea ice system which provides the best estimate of these coupled domains over 1979-2013 period completely. Fig. 2 denotes the stations’ name, location, mean and standard deviation of rainfall. SPI was simulated by Holt-Winter additive and Holt-Winter multiplicative models and ultimately superiority of every model were judged using statistically robust Root Mean Squared Error (RMSE), and R² value. Fig. 3 denotes methodological framework of this research.

**Standardized precipitation index (SPI)**

Mckee et al. (1993) proposed that SPI is found suitable to monitor drought events worldwide (Durdu 2010). This research strictly follows the scheme of Mckee et al., (1993) for the calculation of SPI which is as follows:

The first step is to determine the cumulative difference between total rainfall data points and long term rainfall means (Sönmez et al., 2005). The rainfall data points at j-th time frame are denoted as xij, mean of the rainfall (at j-th time frame) is denoted as x̄j and the standard deviation of rainfall (i-th data point and j-th time step) is denoted as δij so the initial equation is as follows:

\[ SPI = \frac{x_{ij} - x̄j}{δ_{ij}} \]  

(1)
Thom (1958) found that the gamma distribution function can be utilized to fit the various climatological time series in a structured manner. So, in the next step the rainfall series is fitted with gamma probability distribution function (Thom 1958) which is denoted as follows:

\[ P(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \]  

(2)

Here, \( P \) denotes the gamma probability distribution function.

Here, the appearance of parameters is denoted by \( \alpha \) and the range of the parameter is denoted by \( \beta \), \( x \) denotes the amount of rainfall, and \( \Gamma \) is the gamma function. The value of \( \alpha \) and \( \beta \) is greater than 0. Can be written as follows:

\[ i(\alpha) = \lim_{n \to \infty} \int_{0}^{\infty} y^{n-1} e^{-y} dy = \int_{0}^{\infty} y^{\alpha-1} e^{-y} \]

(4)

Here, \( \alpha \) and \( \beta \) are estimated using maximum likelihood solutions (Thom 1958) which is as follows:

\[ \hat{\alpha} = \frac{1}{4A} (1 + \sqrt{1 + \frac{4A}{3}}) \]

(5)

\[ \hat{\beta} = \frac{x}{\hat{\alpha}} \]

(6)

Where,

\[ A = \ln(x) - \frac{\ln(x)}{n} \]

(7)

After that \( \alpha \) and \( \beta \) are adopted to detect the increasing probability distribution function \( p'(x) \) at a selected time step

\[ p'(x) = \int_{0}^{\infty} P(x) dx = \frac{1}{\beta^\alpha} \int_{0}^{\infty} \alpha x^{-\alpha} e^{-\frac{x}{\beta}} dx \]

(8)

Rainfall is not always continuous over time. In this research, zero values are considered as the no rainfall. True feasibility of non exceedence was proposed by Edwards and Mckee (1997) which can be obtained as follows:

\[ H(x) = q + (1 - q)p'(x) \]

(9)

Where, \( q \) is obtained using the method of Thom (1958)

\[ q = \frac{m}{n} \]

(10)

\( q \) is the possibility of zero event, \( m \) is the frequency of zero event in the rainfall dataset.

**Holt-winter exponential smoothing**

Holt-Winter exponential smoothing is a statistical method which is used here to simulate drought at 1 month lead time. It provides an effective way of forecasting of future value of time series data (Razali et al., 2018). \( \alpha, \beta, \gamma \) are the 3 smoothing constants are required for updating the level, trend and seasonal index respectively at time, \( t \). The value of these constants lies between 0 to 1 (Elmunim et al., 2015). This value is selected depending on the weight (high smooth constant mean ensures more weight). Holt-Winter method has two seasonal models i.e. additive and multiplicative.

Here, monthly SPI (\( R_t \)) is assessed from monthly rainfall dataset. Additive and multiplicative model is applied for SPI separately in this research using following equations (Hyndman and Athanasopoulos, 2018):

Simulation and prediction of SPI using Additive Model:

\[ \begin{align*}
L_t &= a_t (R_t - S_{t-1}) + (1 - a_t)(L_{t-1} + b_{t-1}) \\
b_t &= \beta_t (L_{t-1} - L_{t-2}) + (1 - \beta_t) b_{t-1} \\
S_t &= \lambda_t (R_t - L_{t}) + (1 - \lambda_t) S_{t-1} \\
F_{t+m} &= L_t + b_t m + S_{t+m}
\end{align*} \]

(11)
Where, $L_t$ is the level, $b_t$ is the trend, $S_t$ is the seasonal factor, $R_t$ is the actual monthly SPI, and $t$ is the time period of $L_t$, $b_t$, $S_t$ and $R_t$ component. $S$ is the number of seasons of a year, here monthly data has been used so $S = 12$. $S = 11^{th}$ month (i.e. November) to $2^{nd}$ month (i.e. February) is considered as post monsoon season, $S = 3^{rd}$ month (i.e. March) to $5^{th}$ month (i.e. May) considered as pre-monsoon season and rest of the period i.e. $6^{th}$ month (i.e. June) to $10^{th}$ month (i.e. October) is considered as the monsoon season. $F_t$ is the forecast value of $R_t$ ahead of six period, $F_{t+m}$ is the monthly forecasted time period; $\alpha$, $\beta$, $\lambda$ level, trend and seasonal smoothing coefficient estimators respectively, $m$ is the forecast period and $S_t$ is the seasonal duration. The initial value of seasonal component $S_t$ determined using eq. 12 (Granger and Newbold 1986),

$$S_1 = R_1 - I_s$$
$$S_2 = R_2 - I_s$$

(12)

$$L_s = \frac{1}{S}(R_1 + R_2 + R_3 + \cdots \cdots + R_S)$$

(13)

Simulation and prediction of SPI using Multiplicative model:

$$\begin{cases}
L_t = a_t + (1 - \alpha_t)(L_{t-1} + b_{t-1}) \\
b_t = \beta_t L_t - L_{t-1} + (1 - \beta_t)b_{t-1} \\
S_t = \lambda_t (S_{t-1} + (1 - \lambda_t)S_{t-s}) \\
F_t = \left(I_{t-m} - b_{t-1}S_{t-s} \right)S_{t-s} \\
F_{t+m} = \left(I_{t} + b_{m}S_{t-s+m} \\
\end{cases}$$

(14)

The initial seasonal component value ($S_1$) and the level of seasonal duration $s$ is determined by the following equations (15) and (16)

$$S_1 = \frac{R_1}{L_s}, S_2 = \frac{R_2}{L_s}, \cdots \cdots S_t = \frac{R_t}{L_t}$$

(15)

$$L_s = \frac{1}{S}(R_1 + R_2 + \cdots \cdots + R_S)$$

(16)

**Accuracy Assessment**

Error estimation judges the statistical significance of the models. The RMSE and $R^2$ are utilized here which can be expressed as follows:

**Root Mean Squared Error (RMSE)**

Root Mean squared error (RMSE) is calculated using the following equation (Hyndman and Athanasopoulos 2018):

$$\text{RMSE} = \sqrt{\frac{1}{n}(F_t - F_t')^2}$$

(17)

Where, $n$ is the number of observations, $F_t'$ is the actual value, $F_t'$ is the forecasted value at period $t$. We use constant multiplicative factors to adjust the variability, randomness and validation of the data set $z$. Lower the Mean Square Error higher the accuracy and vice versa (Wichitarapongsakun et al., 2016). For the convenient and sophisticated use this research adjusts RMSE within 0 to 1 for pictorial representation.

**Assessment of correlation coefficient ($R^2$) value:**

Appraisal of precision of different simulation models can be estimated by correlation-coefficient ($R^2$) value (Suryanarayana and Mistry 2016). Pearson’s product moment correlation coefficient is the robust statistical technique which can be utilized to estimate efficiency of statistical models (Kamari et al., 2013). $r$ is expressed as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

(18)

Where, $r$ is the correlation coefficient; $x_i$ is the observed value and $y_i$ is the modelled value at time $i$, $\bar{x}$ and $\bar{y}$ is the average value of observed and modelled value, $n$ is the number of observations.

**RESULTS**

The month to month variation of the actual and forecasted SPI of $\delta$ meteorological stations for the period of 1979-2013 are shown in Fig. 5. The observed and simulated SPI showed inter and intra-seasonal variation of peaks (at rainy season) and drops (at dry season) at all the stations. The seasonally simulated SPI for all the meteorological stations varied between -3 to 2.5. From Fig. 5, it was obvious that in all stations, droughts were occurring in the pre-monsoon and post-monsoon period. Station 233859, 229866 and 233866 exhibited the same nature of seasonal variation (Fig. 5), which was increased slightly with the significant difference between actual and forecasted SPI during post-monsoon phase, especially for the months of November-December. Additive models achieved a higher correlation coefficient ($R^2$) value (i.e. 0.82) in relation to the respective observed model (Fig. 4). Relatively low RMSE (0.235) was produced by Additive model in simulation of drought. The additive model produced better result and this result could be very essential to know the present and future conditions of drought of Purulia. The spatial assessment of pre-monsoon, monsoon, and post monsoon SPI (Fig. 6) indicates that in most cases, the eastern portions of the district face highest drought (lowest SPI) using all models. Using all models, average pre-monsoonal SPI (Fig. 6a, 6b, 6c) varied within -2.344 to -0.635. In the monsoon season SPI (Fig. 6d, 6e, 6f) varied within -3.218 to -0.603 (using all models) whereas in the post monsoon season average
SPI (Fig. 6g, 6h, 6i) varied within -3.218 to -0.603 (using all models). According to the observed model, drought was high at the western portions at pre monsoon season and it was low at the eastern portions. At post-monsoon season, northern and north-eastern portions were noticed with high drought and southern portions with low drought. At monsoon and post monsoon season similar feature was noticed. In the case of the multiplicative model, the middle and western portions of the study region were observed with relatively high drought in pre-monsoon, monsoon and post monsoon seasons.

**DISCUSSION**

The drought of Purulia matches with West Bengal as well as all India pattern (Chatterjee et al., 2016). The possible causes of such a picture are deeply rooted in

![Fig. 4. Showing correlation between observed and simulated models (at 1 month lead time) c) Correlation between observed and multiplicative model d) Correlation between observed and additive model.](image)

![Fig. 5. Showing station wise assessments of observed and simulated SPI (at 1 month lead time).](image)
geo-environmental condition as well as global climate change. The seasonal pattern of drought with ENSO (i.e. the warm ENSO event is associated with weak Indian Monsoon) experiences maximum positive relationship in the last 40 years (Kripalani et al., 2003). The southeast ward shift of walker circulation anomalies and higher surface temperatures in winter and spring creates an enormous impact on this changing relationship (Kumar, 1999). Hence the ENSO event cannot truly express the variability of drought in the Indian subcontinent as well as for Purulia. Broadly the drought in West Bengal as well as in India depends on a number of factors, i.e., 'weakening of easterly jet, 'warming of equatorial Indian Ocean than northern latitudes of the tropics', 'weakening of gradient of Sea Surface Temperature (SST) from the equatorial region to South Asian Coast during summer' (Kundu and Mondal, 2019). The study of the simulation of the drought of Purulia is devoted to the micro level variability, which fluctuates seasonally. Therefore no prominent role of ENSO and other phenomena have been assessed to illustrate the simulation of drought event in Purulia. Rather this research strongly encourages the local level factors viz. occurrence of tropical disturbances (Dhar et al., 1981), several topographic variables such as local geology, relief, slope, soil and aspect (Basist et al., 1994), increasing concentration of aerosol at lower atmosphere, changes in agricultural land use by means of initiation of irrigated agriculture (Douglas et al., 2006) and deforestation (Nair et al., 2003). Originally belonging to a granitic terrain, Purulia has crystalline basement rocks that are thinly covered by a weathered mantle (Baidya, 2015). The scattered shallow fracture zones of Purulia cannot store a sufficient quantity of groundwater (Dolui et al., 2021). According to Mukherjee and Singh (2018), the groundwater of Purulia is restricted due to the weathered fracture zone. The fracture zone is having secondary porosity indicating a thick profile of in situ soft porous material developed as a disintegration product on the uppermost part of the

Fig. 6. Showing spatial assessment of SPI using observed and simulated models (at 1 month lead time) in pre-monsoon, monsoon and post-monsoon phase.
hard, consolidated rock due to weathering. Weathering generally imparts secondary porosity to the hard rock, which is fractured at the peninsular Purulia. According to Nag and Kundu (2016) relatively low groundwater level forms the depository of ground water as shallow aquifers in the area occupied by the hard rocks in Purulia.

**Conclusion**

A comparison between Additive and Multiplicative Holt-Winter statistical method was carried out during 1979-2014 for Purulia and accuracy of each model was tested. In general, the trend of forecasted rainfall and SPI for both analysed models, was similar to the actual and forecasted models considering all the stations. Additive model was the most accurate model with relatively higher accuracy than the multiplicative model in simulating SPI over Purulia. Spatial assessment using both the models indicated the fact that in every post monsoon, monsoon and pre-monsoon phase, eastern and north eastern portions were more sensitive to drought whereas the western portions were at moderate to normal condition.

The comparative study of drought in seasonal aspect made the research interesting and unique. The identification and monitoring of the drought at seasonal aspect made the study a fruitful one for implementation. The study might also be very useful to improve the resiliency of the water management infrastructure, including a more accurate drought simulation tool for sustainable planning and assessment of drought of Purulia District.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**REFERENCES**


