Modeling and Forecasting of Rainfall Trends based on Historical Data in Bungoma County, Western Kenya using Holt Winters Method

Vanessa A. Otieno [a] and Samson W. Wanyonyi [b*]

[a] Department of Mathematics, Kibabii University, Bungoma, Kenya.
[b] Department of Mathematics and Actuarial Science, Kabarak University, Nakuru, Kenya.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Original Research Article

Abstract

Rainfall is termed as a meteorological phenomenon which is very useful for daily human activities. Most of the population depend on it for domestic purpose and agriculture, hence it is vital for farmers to know rainfall trends and patterns prevailing in their locality. In Kenya, rainfall variability has caused hunger to many people. In agriculture section, water problem is the most critical constraint in food productions. Extreme weather conditions which is occasioned by climatic change and weather variability makes a small scale farmer in Bungoma to face very high risks of reduced productivity. Scarcity of water is a severe environmental constraint to plant productivity hence drought causes loss in crop yield. Majority of the population in Bungoma do depend on agriculture either directly or indirectly therefore analysis of rainfall data for long periods will provide a lot of information about rainfall variability and this will help to better the agricultural activities of Bungoma farmers. The main aim of the study was to model rainfall patterns and hence predict the future rainfall trends in Bungoma region using Holt winters method in the context of the time series. This is because time series analysis plays an important role in modelling, predicting and forecasting meteorological data such as humidity, temperature, rainfall and other environmental variable. Therefore, data for Bungoma monthly and yearly rainfall patterns for the period 1988-2021 was obtained from the Kenya meteorological department. Collected data was analyzed using Holt winters method by R software.

_____

Corresponding author: Email: samsonwangila@kabarak.ac.ke, samsonwangila@gmail.com;

Rainfall data was found to be seasonal implying that most of the rainfall occurred in a specific period each year. Forecasted rainfall had increasing and decreasing prediction intervals and this implied that rainfall could either start decreasing or increasing. The data was found to be non-stationary due to presence of seasonality and rainfall trends. Findings of the study will make it possible to facilitate planning and management of water for both domestic and agricultural use in Bungoma region.

1. Introduction

The study of time series data is useful in modeling, predicting, and forecasting meteorological data such as humidity, temperature, rainfall, and other environmental variables [1]. Rainfall forecasting is critical for making critical choices and carrying out strategic planning. The ability to statistically predict and forecast rainfall informs the management of water-related problems such as severe rainfall situations such as floods and droughts, among other issues [2].

As a result, anticipating hydrological variables such as rainfall, flood stream, and run-off flow as probabilistic occurrences is an important topic in water resource planning. Typically, these hydrological variables are monitored longitudinally throughout time. This makes time series analysis of their discrete time occurrences suitable for monitoring and recreating their hydrological processes. Rainfall is one of the most complex and difficult components of the hydrological cycle to model and forecast due to a variety of dynamic and environmental elements, as well as random changes both geographically and temporally [1].

There is a great challenge in the agricultural sector due to recently noted extremities in rainfall variability and this heavily impacts East African Counties where irrigation schemes are not developed making agriculture to be entirely rain-fed [3]. In Eastern Africa the mean annual rainfall is more likely to occur and this is contrast to and northern Sahara and Mediterranean since they are likely to experience decrease in rainfall [4]. Even though the mean annual rainfall in East Africa is expected to increase, the increase might not be uniform across time and space [5].

These changes however are uncertain due to under estimation of the warming impacts of Indian Ocean in many general climate models making it possible to overestimate rainfall increase in the region [3]. Some records shows that there is an upward trend in Kenyan rainfall however there are a range of models and scenarios which suggest both an increase and decrease in rainfall [4, 6].

Many studies have proved that the total rainfall projection in Kenya increases from 0.2-0.4% each year until the 2090s 2090s [5, 7]. There is an increase in heavy rainfall and there is likely to be more frequent extreme events and this results to greater rainfall variability Macenzie F. Zamani. [8]. Correct forecasting of future rainfall will not only play a major role in the management of water resources but also assist in boosting agriculture since farmers will plant most crops during rainy seasons, Oyamakin [9].

2. Research Methodology

This section lays out how the research was conducted and the process that was used to achieve the purpose of the study. The study area and analysis tool used to analyze data is also discussed in detail.

2.1 Data Source

Monthly and yearly rainfall data was obtained from Kenya meteorological department for the period 1988-2021 which was used in the study. This long period of rainfall data was preferred so as to accurately predict the 2022 rainfall patterns and trends of Bungoma region.

2.2 Study Area

The geographical location of Bungoma is 0.4213^0N to 1.1477^0N along the latitude and 34.3627^0E to 35.0677^0E along the longitude. Bungoma is found in the western region of Kenya and is bordering Uganda. Its area is

approximately 2069 km^2. The population of Bungoma is 1,670,570 from the 2019 Kenya population census. Its climate is tropical and also a lot of rain is experienced in summers than winters. The average temperature is 20.3^{0}C | 68.5^{0}F and the rainfall is 1102 mm | 43.4 inch in a year.

2.3 Method of Analysis

Collected data was analyzed using R software. The analyzed data is presented in graphical forms where features such as rainfall trends and seasonality component are observed. In this study Holt-Winters forecasting method was used to predict the future rainfall trends in Bungoma region. According to Kamruzzaman et al. [10], Holt-winters method can be applied in forecasting non-stationary seasonal rainfall and temperature data. A study which aimed in forecasting rainfall in Nakhon Ratchasima province, Thailand was conducted by Spipan (2014) [11]. Holt-winters method was used in forecasting monthly rainfall at 15 stations which have historical monthly data from April 2005 up to March 2013. Mean squared error and mean absolute error were used in measuring the performance and the results showed that forecasting using this method is acceptable. Holt-Winters' multiplicative method was the preferred model because the data had seasonality component, and this implied that there are changes in the widths of the seasonal periods over time. The equation y_t is used to show the relationship between time series and state variables. Parameter p_t represents the time series level, q_t represents the period of growth, r_t is the seasonal factor and e_t the error term. Parameters $\hat{p}_t$, $\hat{q}_t$ and $\hat{r}_t$ represents the equation for the Holt winters multiplicative smoothing. According to Booranawong & Booranawong, (2017) [12], single exponential smoothing and double exponential smoothing were the techniques used to forecast and model monthly lime price of Thailand. Parameter m represents the number of seasons in a particular year. The first step is to show the relationship between the time series y_t in the first equation and the state variable p_t, q_t and r_t in the second, third and fourth equation at time t.

$$y_t = (p_{t-1} + q_{t-1}) \times r_{t-m} \times (1 + \varepsilon_t)$$

$$p_t = p_{t-1} + q_{t-1} + \beta_1 + (p_{t-1} + q_{t-1}) \times \varepsilon_t$$

$$q_t = q_{t-1} + \beta_2 \times (p_{t-1} + q_{t-1}) \times \varepsilon_t$$

$$r_t = r_{t-m} + \beta_3 \times r_{t-m} \times \varepsilon$$

Where $\beta_1, \beta_2, \beta_3$ are the smoothing constants.
The second step is to calculate the forecast error ε_t.

$$\varepsilon_t = y_t - (p_{t-1} + q_{t-1}) \times r_{t-m}$$

The third step is to compute the error term ε_t from m number of seasons.

$$\varepsilon_t = \frac{y_t - (p_{t-1} + q_{t-1}) \times r_{t-m}}{(p_{t-1} + q_{t-1}) \times r_{t-m}} = \frac{\varepsilon_t}{(p_{t-1} + q_{t-1}) \times r_{t-m}}$$

The forth step is to use $y_1, y_2 \ldots y_n$ to find the estimates of $\hat{p}_0, \hat{q}_0, \hat{r}_{j-m}$
The fifth step is to forecast periods h forward at time n.

$$\hat{y}(h) = (p_n + hq_n) \times \hat{r}_{n+h-jm} \text{ Where } j = \left(h/m\right) + 1$$

3. Results and Discussion

Figure1 displays average monthly rainfall data of Bungoma region from 1988-2021. Seasonality component can be observed from the graph. This is because there is an irregular distribution of rainfall patterns during the years. Most of the rainfall occurs in specific period in each year. In this plot there is no specific trend of the rainfall pattern in Bungoma County.

A time series is usually made up of components such as trend, seasonality and randomness therefore studying these components in a separate manner gives a good comprehensive insight about the time series. The time series data was therefore decomposed so that features such as trend, seasonality and randomness can be observed. Smoothing parameters which were observed are; an alpha (level smoothing factor) level of 0.07712996, this indicates that much weight is given to the latest values when being included in the current level estimate. Beta (trend) level of 0.3836055 was observed, this indicates that much weight is given to the older values. Gamma (seasonal smoothing) level of 0.2416752 was observed. Gamma specifies how smooth the seasonal component of the time series is. From this gamma value it's evident that older values weighted more heavily.
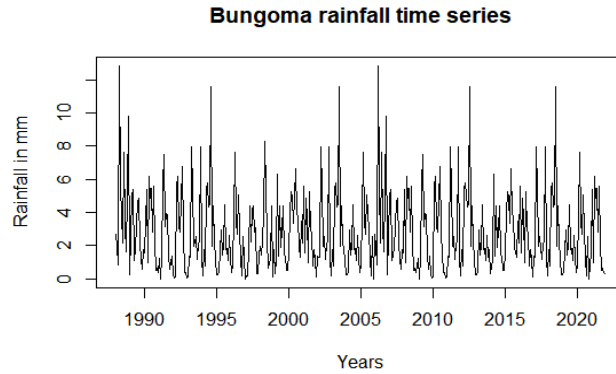

Fig. 1. Plot of average monthly rainfall of Bungoma region from 1988-2021

Table 1. Smoothing Parameters

| Smoothing Parameters	
α	0.07712996
β	0.3836055
γ	0.2416752

Fig. 2 displays the decomposed rainfall series data which is more detail based on the observed series, trend of the rainfall pattern, seasonality and randomness. This decomposition result will give more precise insight into rainfall behavior during 1988–2021 periods. The first plot shows the observed rainfall trends for the period 1988-2021 in Bungoma County. Under seasonal plot it's observed that there's seasonal fluctuation. This is due to the upward and downward trend. Under trend, it's observed that there is no specific trend for the rainfall patterns in the period 1988-2021. The last plot shows the randomness of the rainfall patterns.
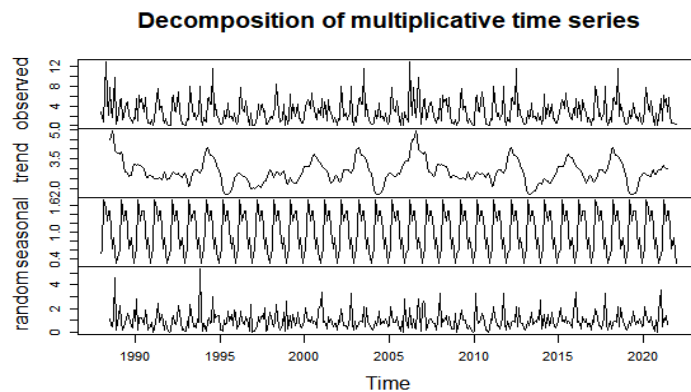
**Fig. 2. Plot of rainfall decomposition**

Fig. 3 showed the series followed by the forecast as the blue line and the upper and lower predictions limit as grey lines. Holt-winters was used to model and predict the behavior of Bungoma rainfall patterns this is because

it is one the most popular forecasting techniques for time series. The rainfall patterns from 1988 to 2021 was used to forecast the rainfall patters for the year 2022. The increasing and decreasing prediction intervals showed that rainfall could either start decreasing or increasing. The forecasted rainfall patterns is displayed by the blue line.

**Forecasts from HoltWinters**
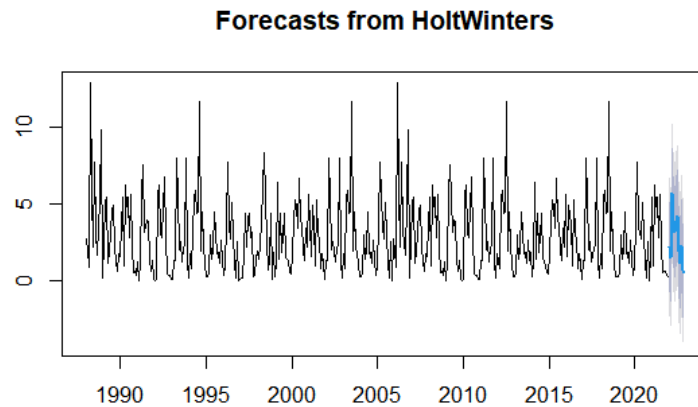
Fig. 3. Forecasted rainfall patterns

From Fig. 4 and Fig. 5, heavy rains are expected in the months of March, April, May, June and July while short rains are expected in the months of February, August, September October, November and December and no rainfall in January. This prediction can be useful for a farmer who wants to know which the best month to start planting and also for the government in case there's need to prepare for any policy of preventing flood on rainy season & drought on dry season.
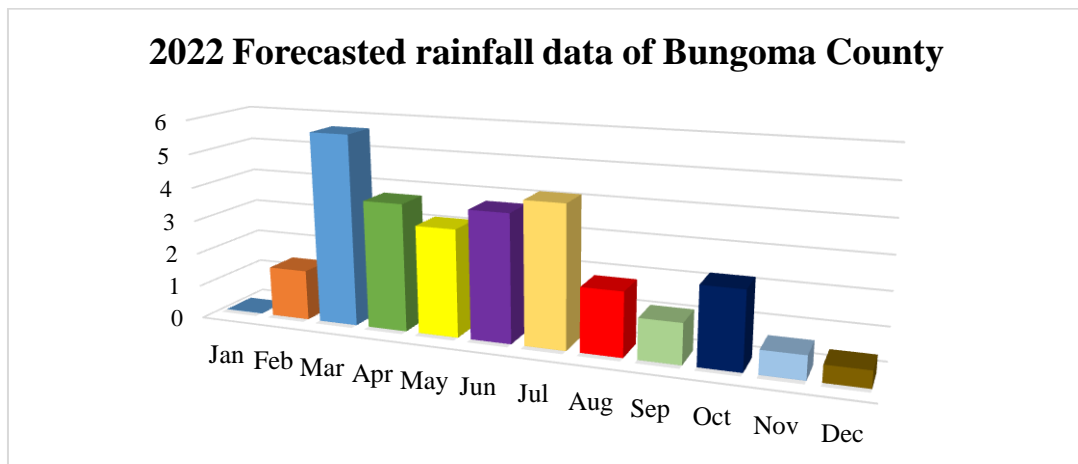
Fig. 4. 2022 Forecasted rainfall data of Bungoma in millimeters
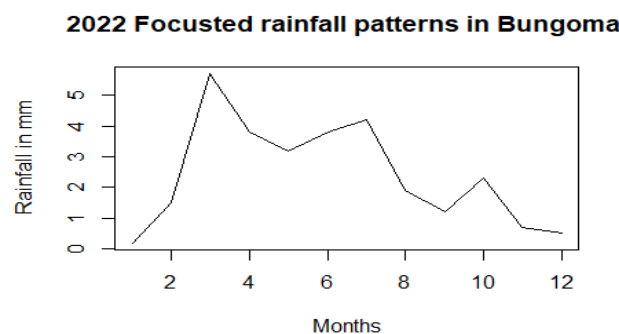
Fig. 5. Line graph of 2022 Forecasted rainfall patterns of Bungoma in millimeters

A ljung box test hypothesis was used to test whether errors are independent and identically distributed (white noise) hence the null hypothesis is:

H_0: Residuals are independently distributed.
H_1: Residuals are not independently distributed.

```
                Box-Ljung test

        data:  Bungomatimeseriesforecast2
        X-squared = 1.4681, df = 1, p-value = 0.2256
```

Fig. 6. Box-ljung test

The test statistics of the test is 1.4681 and the p-value of the test is 0.2256 which is larger than 0.05 therefore we do not reject the null hypothesis and conclude that the residuals are independently distributed.

4. Summary and Conclusion

This present study investigated the variability and trend of annual rainfall for the period 1988-2021. From the Holt winters model rainfall pattern in Bungoma region was found to change significantly over time. There were periods of low variability and others of extreme variability. Rainfall data was found to be non-stationary due to presence of rainfall trends and seasonality. Forecasted rainfall had increasing and decreasing prediction intervals and this implies that rainfall could either start decreasing or increasing. From the forecasted data, heavy rains are expected in the months of March, April, May, June and July while short rains are expected in the months of February, August, September October, November and December and no rainfall in January. The planning and management of climate sensitive activities such as water resources will play a major role in boosting agricultural sector because farmers will be able to plant lots of food crops during the rainy season. From this study, it can be concluded that, rainfall patterns for Bungoma region would change over time.

5 Recommendation

Based on the findings of the study, it is recommended that Bungoma farmers in conjunction with the Bungoma County government need to develop affordable and appropriate water harvesting technologies for increased water use efficiency in crop and livestock production.

Acknowledgement

I would like to thank my co-author for making this research successful. I am also grateful to all the farmers and agriculture extension workers for being ready to provide useful information during the research. I also thank the Almighty God for giving us good health during the research period, it is through His grace that we have been able to come up with this research.

Competing Interests

Authors have declared that no competing interests exist.

References

[1]   Htike, Khalifa. Time series data in modeling, predicting, and forecasting meteorological data eg humidity, temperature, rainfall; 2010.

[2]   Ansari. Predicting and forecasting rainfall informs the management of water-related problems such as floods and droughts; 2013.

[3] Eriksen S, Lind J. Adaptation as a political process: Adjusting to drought and conflict in Kenya's drylands. Env Manag. 2009;43:817-835.

[4] Chen H, Githeko GF, Zhou JI, Githure, Yan. New records of *Anopheles arabiensis* breeding on the Mount Kenya highlands indicate indigenous malaria transmission. Malaria J. 2006;5:5-17.

[5] Muthama NJ, Manene M. Simulation of Decadal Precipitation over Nairobi in Kenya. *SQU* Journal For Science. 2008;43-54.

[6] Helsel DR, Hirsch RM. Hydrological analysis and interpretation: statistical methods in water resources. US Geological Survey, Reston; 2002.

[7] Moazzam MFU, Lee BG, Rahman G, Waqas T. Spatial rainfall variability and an increasing threat of drought according to climate change in Uttaradit Province, Thailand. Atmos Clim Sci. 2020;10:357‑371.

[8] Macenzie F. Zamani: A survey of East African history. East African publishers, Nairobi, Kenya. 1974.

[9] Oyamakin SO, Ayoola FJ, Dare OT. Time series analysis of rainfall and temperature in South West Nigeria,Pacific. Journal of sciences and Technology. 2010;552-564.

[10] Kamruzzaman M, Beecham S, Metcalfe AV. 2011. Non‑ stationarity in rainfall and temperature in the Murray Darling Basin. Hydrological Processes 25(10), 1659-1675.

[11] Spipan. Time series analysis of rainfall using Holt winters in Nakhon Ratchasima province, Thailand. Journal of sciences and Technology; 2014.

[12] Booranawong T, Booranawong A. Simple and double exponential smoothing methods with designed input data for forecasting a seasonal time series: in an application for lime prices in Thailand. Suranaree Journal of Science & Technology. 2017;24(3).

_____