International Journal of Statistics and Applied Mathematics 2020; 5(2): 191-195

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2020; 5(4): 191-195

www.mathsjournal.com
Received: 08-05-2020
Accepted: 12-06-2020

Hudson Nyang'wara Ongiri
Department of Mathematics and Actuarial Science, Kisii University, Kisii Kenya

Paul Wachiuri Warutumo
Department of Mathematics and Actuarial Science, Kisii University, Kisii Kenya

Corresponding Author:
Hudson Nyang'wara Ongiri
Department of Mathematics and Actuarial Science, Kisii University, Kisii Kenya

Use of triple exponential smoothing in the analysis of hydrological data

Hudson Nyang'wara Ongiri and Paul Wachiuri Warutumo

Abstract
Regular and correct hand hygiene is one of the most important measures to prevent infection with the COVID-19 virus. WASH practitioners should work to enable more frequent and regular hand hygiene by improving access to hand hygiene facilities to support good hand hygiene behavior. Performing hand hygiene at the right time, using the right technique with either an alcohol-based hand rub and soap and water is critical. It makes water to be an essential resource in the fight against the pandemic. This article ventured into the analysis of water demand by Kisii County householders. This article employed a triple exponential smoothing method. The exponential smoothing methods usually applied in the analysis of univariate time series data. This study employed the Cox-Stuart method to determine the trend of the data. Since p-value = 0.00001141 < the significance level (α) = 0.05, this study concluded that the water data has a trend. The parameters of the triple exponential smoothing were identified to be α=0.2358, β=0.0028 and γ = 0.0976. They were determined in such a way that the mean squared error (MSE) of the error is minimized. In-sample forecasting was employed. No significant difference was noted. The exponential smoothing model was employed in out of sample forecasting, and it was realized that the water demand was expected to decrease. This study recommends the use of other statistical models to establish if the same results could be realized.

Keywords: Trend, exponential smoothing, in-sample, out of sample, forecasting

1. Introduction

According to [4], Hand hygiene is essential to prevent the spread of the COVID-19 virus. Regular and correct hand hygiene is one of the most important measures to prevent infection with the COVID-19 virus. WASH practitioners should work to enable more frequent and regular hand hygiene by improving access to hand hygiene facilities to support good hand hygiene behavior. Performing hand hygiene at the right time, using the right technique with either an alcohol-based hand rub or soap and water, is critical. Existing WHO guidance on the safe management of drinking water and sanitation services applies to the COVID-19 outbreak. Water disinfection and sanitation treatment can reduce viruses. Many health co-benefits can be realized by safely managing water and sanitation services, and by applying good hygiene practices.

Hygiene makes water to be an essential resource in the fight against the pandemic. This article ventured into the analysis of water demand by Kisii County householders. This article employed a triple exponential smoothing method. According to [1], the presented method is found to have excellent forecast performance for time series with and without outliers, as well as for fat-tailed time series and under model misspecification.

The exponential smoothing methods are generally applied in the analysis of univariate time series data. It is generally considered for an alternative to the Box-Jenkins methodology (ARIMA) [2]. It is a rule of thumb technique for smoothing time series data using the exponential window function. Whereas in the simple moving average, the past observations are weighted equally, exponential functions are used to assign exponentially decreasing weights over time. It is a quickly learned and easily applied procedure for making some determination based on prior assumptions by the user, such as seasonality. Exponential smoothing is often used for the analysis of time-series data. Exponential smoothing is one of many window functions commonly applied to smooth data in signal processing, acting as low-pass filters to remove high-frequency noise. This method is preceded by Poisson's use of recursive exponential window functions in convolutions from the 19th century, as well as Kolmogorov and Zurbenko's use of recursive moving averages from their studies of turbulence in the 1940s [6].

2. Methodology

The article delved into the application of triple exponential smoothing in the analysis of household water data. We ventured into finding out if the data had a trend by using the Cox-Stuart method. The procedure is as follows:

2.1 Cox and Stuart (C-S) Trend Test

C-S test is customarily employed to detect the non-random pattern, which is the periodic pattern. C-S test compares the first half and the second half of the sample data. When the data has a downward trend, the observations in the first half are expected to be higher than the observation in the second half. When the data has an upward trend, the observations in the first half are expected to be smaller than the observations in the second half. When the data in question has no trend, then the researcher should expect smaller differences between the two halves of the sample data due to the randomness of the data [5].

To perform the C-S test, the sample differences are computed as shown below;

$$Y_1 = x_{1+m} - x_1, Y_2 = x_{2+m} - x_2, Y_3 = x_{3+m} - x_3, \ldots, \quad Y_m = x_n - x_{n-m} \quad (2.1)$$

Where $m = \frac{n}{2}$ if and only if n is even and $m = \frac{(n+1)}{2}$ if and only if n is odd.

The differences which are equal to zero are ignored in this case. Denoting the sample data with positive differences by Y_1, Y_2, Y_3 Y_m; then, the C-S test is a sign test applied to the sample data of non-zero differences $Y_1, Y_2, Y_3, \ldots, Y_m$ [5].

Let $sgn(b) = 1$, if $b > 0$ and $sgn(b) = -1$, if $a < 0$. Then the C-S statistic is given as:

$$C - S\ Statistic = \sum_{j=1}^{m} sgn(Y_j) \quad (2.2)$$

The hypotheses tested at the significance level (α) are

H_0: There is no periodic trend in the household water data

H_{A}: There is a periodic trend in the household water data

To use the triple exponential method, the time series data must have the trend and the seasonal component. The basic equations are indicated below.

2.2 Overall Smoothing Equation

$$S_t = \alpha \frac{x_t}{I_{t-L}} + (1 - \alpha)[S_{t-1} + b_{t-1}] \quad (2.3)$$

Where **x** is the observation, **S** is the smoothed observation, b is the trend factor, and t is the time

2.3 Trend Smoothing Equation

$$b_t = \gamma[S_t - S_{t-1}] + (1 - \gamma)b_{t-1} \quad (2.4)$$

Where **S** is the smoothed observation and b is the trend factor

2.4 Seasonal Smoothing

$$I_t = \beta \frac{x_t}{S_t} + (1 - \beta)I_{t-L} \quad (2.5)$$

Where I is the seasonal index

2.5 Forecasting Equation

$$F_{t+m} = [S_t + mb_t]I_{t+m-L} \quad (2.6)$$

Where I is the seasonal index and F is the forecast at m periods ahead

α, β and γ are constants that must be determined in such a way that the mean squared error (MSE) of the error is minimized. It is mostly achieved using a statistical software. To initialize the triple exponential smoothing method, we need at least season's data to estimate the initial seasonal indices $I_t - L$

A complete season's data comprises of L periods. We need to estimate the trend factor from one period to the next. It compels the researcher to use two complete seasons- that is- 2L periods [3].

2.6 Initial Values for the Trend Factor

The initial value of the trend is estimated by:

$$b = \frac{1}{L}\left[\frac{x_{L+1} - x_1}{L} + \frac{x_{L+2} - x_2}{L} + \cdots + \frac{x_{L+L} - x_L}{L}\right] \quad (2.7)$$

2.7 Initial Values for the Seasonal Indices

The initial values for the seasonal indices were computed using the steps indicated below.

In our study, we worked with hydrological data with a periodicity of 12 (that is 12 months per year).

STEP 1: Find the mean of the n years

$$A_p = \frac{\sum_{i=1}^{n} x_i}{12}, p = 1,2,3, \ldots, 12 \quad (2.8)$$

STEP 2: Divide the Observations by the appropriate yearly mean

1	2	
x_1/A_1	x_{13}/A_2	…
x_2/A_1	x_{14}/A_2	
.	.	
.	.	
.	.	
x_{12}/A_1	x_{124}/A_2	

The above process of finding the mean proceeds for the next n years.

STEP 3: The seasonal indices are formed by computing the average of each row. That is;

$$I_1 = \frac{x_1/A_1 + x_{13}/A_2 + x_{25}/A_3 + \cdots}{12}$$

$$I_2 = \frac{x_2/A_1 + x_{14}/A_2 + x_{26}/A_3 + \cdots}{12}$$

$$I_{12} = \frac{x_{12}/A_1 + x_{24}/A_2 + x_{36}/A_3 + \cdots}{12}$$

Sometimes zero coefficients for the trend or seasonality or both are encountered. It does not necessarily imply that there is no trend component or seasonality. It means that both of them were right on the money; hence no updating was required in order to achieve the lowest MSE [3].

3. Results

This study ventured into the analysis of hydrological data from 2012 to 2020 to determine the forecast of the future demand for water in Kisii County. This study conducted a trend test using the Cox-Stuart method.

3.1 Cox-Stuart Trend Test

Cox and Stuart's trend analysis method was subjected to household water data. The hypotheses were:
H_0: There is no periodic trend in the household hydrological data
$H_{A.}$: There was a periodic trend in the household hydrological data
The hypotheses were tested at a 5% significance level.
The test results from r software are as indicated in Table 4 below

Table 1: Cox-Stuart Trend Test

Description	Value
Cox-Stuart Statistic	4.8655
P-value	0.00001141
Significance level(α)	0.05

From the table above, it can be seen that the p-value = 0.00001141 < the significance level (α) = 0.05. Therefore this study can reject the null hypothesis in favor of the alternative hypothesis and conclude that the household water data has a periodic trend.

3.2 Estimated Parameters of Exponential Smoothing

The three parameters were estimated, and they are indicated in the table below

Table 2: Estimated Parameters

Parameter	Value
α	0.2358
β	0.0028
γ	0.0976

$\alpha = 0.2358, \beta = 0.0028\ and\ \gamma = 0.0976$ are constants that were determined in such a way that the mean squared error (MSE) of the error is minimized. It was achieved using a statistical software. Therefore the smoothing equations are:

$$S_t = 0.2358\frac{x_t}{I_{t-L}} + 0.7642[S_{t-1} + b_{t-1}],$$

$$b_t = 0.0976[S_t - S_{t-1}] + 0.9024b_{t-1} \text{ and}$$

$$I_t = 0.0028\frac{x_t}{S_t} + 0.9972I_{t-L}$$

3.3 Interpolation Fit of Exponential Smoothing

Figure 1 below indicates the interpolation fit for the chosen triple exponential smoothing

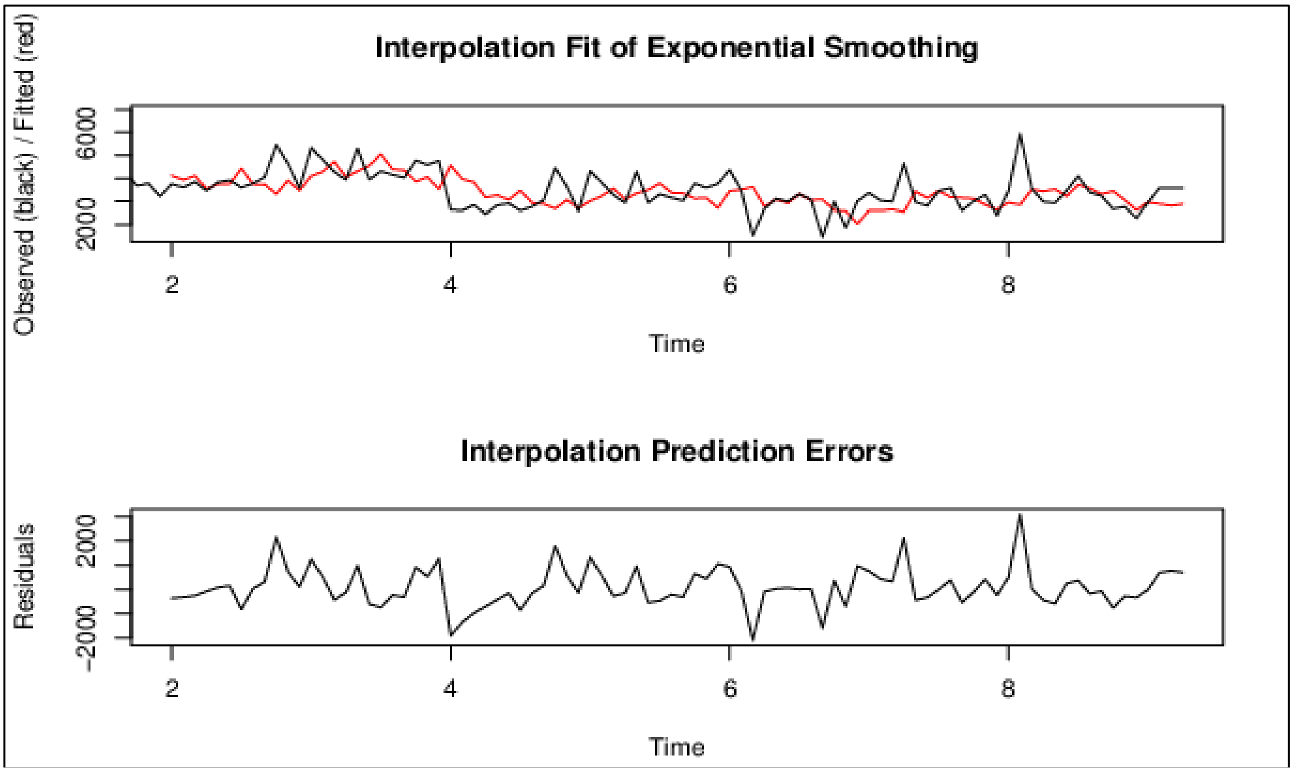

Fig 1: In-Sample forecasts and Prediction Errors

It can be deduced that there is no significant difference between the actual data plot and that of the in-sample forecasted plot. It suggests that the model gives a good representation of the data. Therefore this study deemed it fit to use in conducting out-sample forecasting.

3.4 Extrapolation Forecasts of Exponential Smoothing

Since the In-sample forecasting was identified to be within the limit bounds, this study went further and conducted out of sample forecasting. The out of sample forecasting are indicated in table 3 below

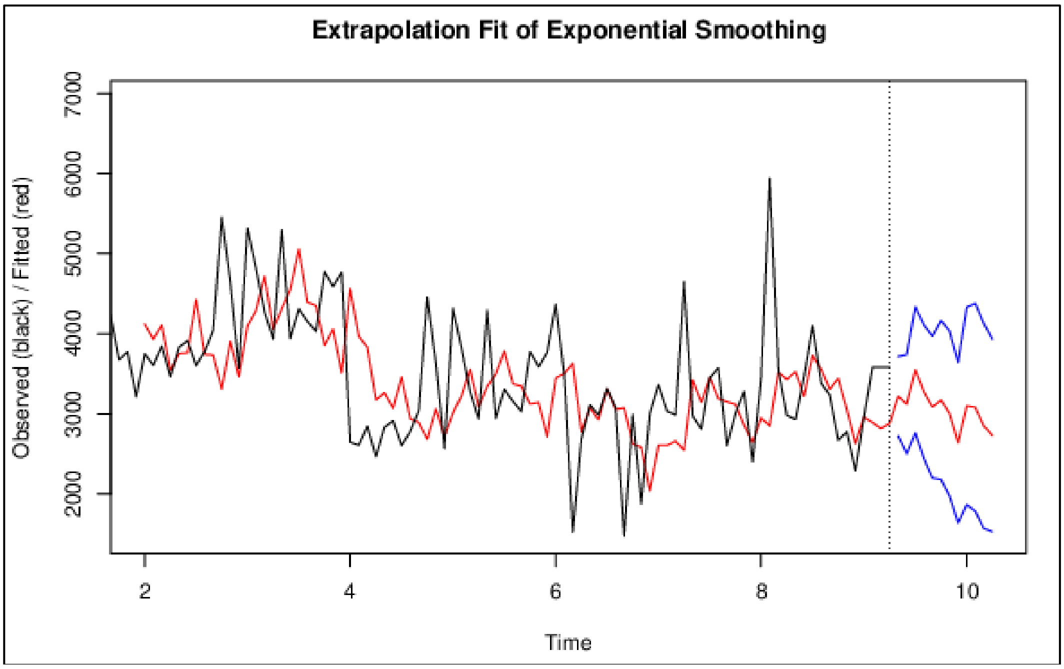

Fig 2: Plot of In-Sample and Out of Sample Forecasting

Figure 2 above indicates both the in-sample and out-of-sample forecasting

Table 3: Out of Sample Forecasting

t	Forecast	95% Lower Bound	95% Upper Bound
101	3217.37841842984	2720.61973301774	3714.13710384193
102	3119.33545688723	2499.53936475964	3739.13154901482
103	3550.14406298365	2763.54279757318	4336.74532839412
104	3274.98107709239	2440.63464247747	4109.32751170732
105	3083.96778036811	2197.64019282383	3970.29536791239
106	3170.74823948569	2177.623109805	4163.87336916638
107	3004.41923154056	1974.78515954325	4034.05330353786
108	2637.47398279769	1635.70232106092	3639.24564453446
109	3096.73574266512	1863.58403674882	4329.88744858141
110	3079.59523036195	1780.93969386018	4378.25076686371
111	2847.8308281784	1568.06448528927	4127.59717106754
112	2728.58642352323	1525.30491632836	3931.8679307181

Figure 3 below indicates the plots of the residual plots of the forecasted values compared to the observed data.

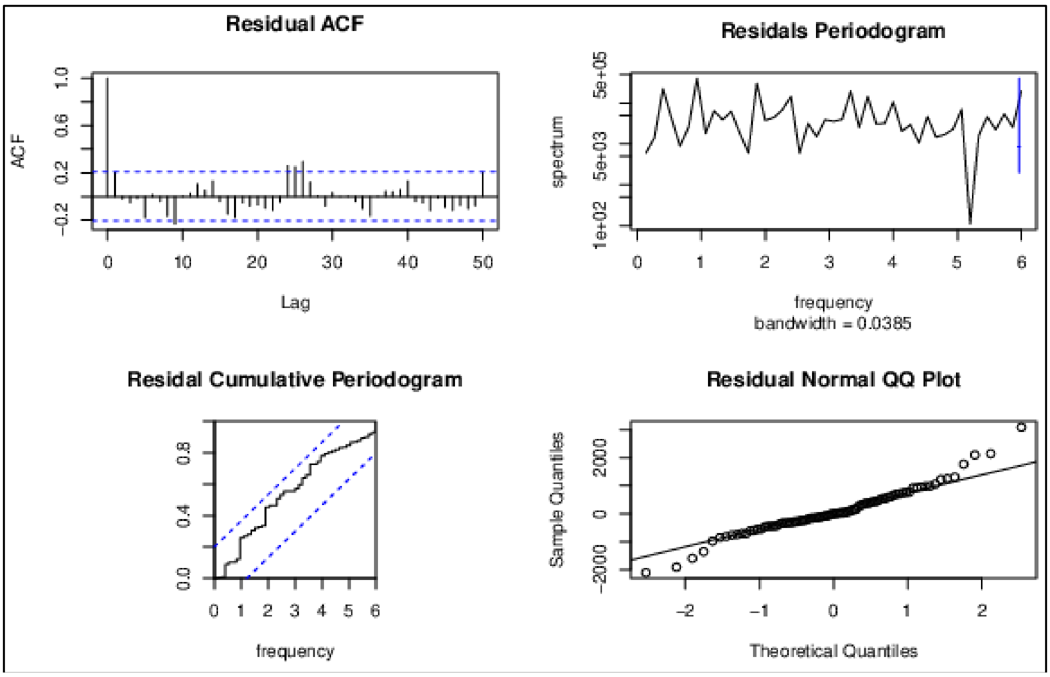

Fig 3: Residual Plots

It can be established that the residuals are normally distributed and fall with the limit bounds indicating a good fit of the model.

4. Conclusions

- It can be seen that the p-value = 0.00001141 < the significance level (α) = 0.05 from the Cox-Stuart test. Therefore this study can conclude that the household water data has a trend.
- Since the in-sample forecasts show no significant difference compared to the observed data, the estimated parameters can be applied in forecasting the future demand of household water.
- From figure 2 above, it can be deduced that the quantity of household water demanded is expected to decrease.

5. Recommendations

- This study employed the Cox-Stuart method to determine the trend of water demanded. Future researchers can employ other methods like the ordinary least square method to estimate the trend.
- The future researchers can employ other models like harmonic regression models, principal component regression models to determine if the same results could be realized.
- The methods employed in this article can be applied in other areas to analyze and forecast future values.

6. Acknowledgment

I wish to thank all individuals whose information was of value in drafting this article. God bless you all.

7. References

1. Gelper Sarah, Fried Roland, Croux Christophe. Robust Forecasting with Exponential and Holt-Winters Smoothing. Journal of Forecasting. 2010; 29(3):285-300. doi: 10,2139/ssrn.1089403
2. Jason Brownlee. A Gentle Introduction to Exponential Smoothing for Time Series Forecasting in Python. Retrieved from Machine Learning Mastery, 2020. https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/
3. Mary Natrella, Paul Tobias, Carroll Croarkin. e-Handbook of Statistical Methods, 2012. doi: doi.org/10.18434/M32189
4. Matt Arduino *et al*. Water, sanitation, hygiene, and waste management for the COVID-19 virus. World Health Organization, 2020.
5. Robert Orwenyo Onyoni, George Otieno Orwa, Zablon Maua Muga. Modeling household water demands using sinusoidal models. International Journal of Statistics and Applied Mathematics. 2019; 4(6), 128-132.
6. Wikipedia Contributors. Exponential Smoothing. Retrieved from Wikipedia, The Free Encyclopedia, 2020. https://en.wikipedia.org/w/index.php?title=Exponential_smoothing&oldid=966519029