Available online at http://scik.org
J. Math. Comput. Sci. 11 (2021), No. 4, 3887-3915
https://doi.org/10.28919/jmcs/5400
ISSN: 1927-5307

MODELLING AND FORECASTING OF WEB TRAFFIC USING HOLT'S LINEAR, BATS AND TBATS MODELS

AMR BADR[1], TATIANA MAKAROVSKIKH[2], PRADEEP MISHRA[3,*], MOSTAFA ABOTALEB[2], ABDULLAH MOHAMMAD GHAZI AL KHATIB[4], KADIR KARAKAYA[5], SONYA REDJALA[6], ANURAG DUBEY[7], EMMANUEL ATTAL[8]

[1] University of New England, Faculty of Science, School of Science and Technology, Australia
[2]Department of System Programming, South Ural State University, Chelyabinsk, Russia
[3]College of Agriculture, Powarkheda, Jawaharlal Nehru Krishi Vishwa Vidyalaya, (M.P.), India
[4]Department of Banking and Insurance, Faculty of Economics, Damascus University, Syria
[5]SelçukUniversity, Faculty of Science, Department of Statistics, Konya, Turkey
[6]Department de Génie Mécanique, Université Mouloud Mammeri de Tizi-Ouzou UMMTO, Algeria
[7]Laboratoire de Mécanique Gabriel Lamé (LaMé), INSA Centre Val de Loire, Blois, France
[8]Institut FEMTO-ST, CNRS, Université de Bourgogne Franche-Comté, Besançon, France

Abstract: In the recent era, internet consumption has increased. Due to this heavy and regular use of internet web traffic is increased. Sometimes due to high web traffic server may also affect. In this study, cumulative data on the number of visitors to Wikipedia, Facebook, Energy, Android, and Apple, are analyzed in detail. Some descriptive statistics of visitor’s pages in Wikipedia are given such as mean, minimum, maximum, standard deviation, skewness

*Corresponding author
E-mail address: pradeepjnkvv@gmail.com
Received January 8, 2021

and kurtosis. The present study used different time series models like Holt's Linear Trend, BATS and TBATS for different web pages. From the results, it was found that the Android page as well as apple page in Wikipedia holt's Linear Trend model performance is better compared to other models. This kind of projection is helpful for web traffic to solve the server breakdown problem for larger users of the server. These Wikipedia pages have been chosen to Forecast the number of visitors to these pages through time series models BATS, TBATS, and Holt's Linear Trend Model in order to face future problems to mitigate over loading that may occur with the increase in the number of visitors to these pages and also to experiment the suitability of these models of Time series to Forecast the number of visitors and that to achieve the highest level of accuracy. MAPE for accuracy was used to compare model performance.

Keywords: Holt's linear trend; forecasting; TBATS; BATS; time series; web traffic; Wikipedia; webpages.

2010 AMS Subject Classification: 37M10.

1. INTRODUCTION

Nowadays the web traffic consumption has been increased significantly and it became necessary to be forecasted in order to design the web server to balance and distribute the request loads [1]-[7]. Forecasting the load can allow building load balancing techniques to schedule client requests to meet the demands and to be dynamically scaled up and down for the server based on the forecasted visits[21]. Forecast of the upcoming traffic on multiple pages on a website is very challenging specifically if we are trying to forecast the traffic for Wikipedia pages [2] as it's becoming the primal source of knowledge for billions of users [3], who are constantly contributing and expanding the content. Wikipedia consented users to access the traffic by querying each article over a given window [7], but we used in this research the available dataset created by the Kaggle competition [6]. [22-23] compared times series models and used for forecasting purposes. In this research, we employed the time series forecasting techniques to estimate future request loads based on historical visits[4], traditional time series such as Autoregressive integrated moving average "ARIMA"[5], Seasonal ARIMA "SARIMA"[4] have been used extensively for forecasting and based on a recent research Abotaleb [9] found that

Holt's Linear Trend is giving accurate results In predicting the daily Infection cases for COVID-19 research in (Italy, China and USA) reaching to a conclusion that Holt's model is giving more accurate results than ARIMA models. So, we used Holt's Linear Model, and Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend and Seasonal components "TBATS" in which time series seasonality is not forced to be periodic, however, it's allowed to be dynamic, in which results proved to be more accurate.

2. MATERIAL AND METHODS

This study was performed on data obtained from a database by the Kaggle competition from01-07-2015 to 31-12-2016 [10]. We used cumulative data on the number of visitors to Wikipedia and four pages, namely Facebook, Energy, Android, and Apple, were selected. In this paper, we used time series models because these model a combination of Fourier with an exponential smoothing state-space model and a Box-Cox transformation, Both the BATS and TBATS models dealing with seasonality.

2.1. BATS and TBATS Models

TBATS is an improvement modification of BATS that allows multiple seasonal incorrect cycles. TBATS has the following equation [11].

Equation (1) is a Box-Cox transformation

$$Y_t{}^{(\omega)} = \begin{cases} \dfrac{y_t{}^{(\omega)} - 1}{\omega} \omega \neq 0 \\ log y_t \omega = 0 \end{cases} \tag{1}$$

Equation (2) represents the seasonal M pattern

$$Y_t{}^{(\omega)} = l_{t-1} + \emptyset b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t \tag{2}$$

Equations (3), (4) and (5) are global trends and local trends

$$l_t = l_{t-1} + \emptyset b_{t-1} + \alpha d_t \tag{3}$$

$$b_t = \emptyset b_{t-1} + \beta d_t \tag{4}$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t \tag{5}$$

equation (6) is the error modeled by ARMA

$$d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{6}$$

where $m_1, \dots, m_T$ denote that seasonal period, l_t and b_t denote that the level and trend of components of the time series at time t, $s_t^{(i)}$ denote that seasonal component at time t, d_t represents to ARMA(p, q) component and ε_t is white noise process.

The smoothing parameters are given by α, β, γ_i for i=1…T and $\emptyset$ is the dampening parameter, which gives more control over trendextrapolation when the trend component is damped [12]. For seasonal data the following equations representing Trigonometric exponential smoothing models

$$s_t^{(i)} = \sum_{j=1}^{k_i} a_{j,t}^{(i)} \cos\left(\lambda_j^{(i)} t\right) \tag{7}$$

$$a_{j,t}^{(i)} = a_{j,t-1}^{(i)} + k_1^{(i)} d_t \tag{8}$$

$$\beta_{j,t}^{(i)} = \beta_{j,t-1}^{(i)} + k_2^{(i)} d_t \tag{9}$$

where $k_1^{(i)}$ and $k_2^{(i)}$ are the smoothing parameters.

$\lambda_j^{(i)} = 2\pi j/m_i$. This is an extended, modified single source of error version of single seasonal multiple sources of error representation suggested by [13]and is equivalent to index seasonal approaches when $k_i = m_i/2$ for evenvalues of m_i and when $k_i = (m_i - 1)/2$ for odd values of m_i. But most seasonal terms will require much smaller values of k_i, thus reducing the number of parameters to be estimated.

In the single seasonal multiple sources of error setting [14] an alternative, but equivalent formulation of representation (2) is preferred [15] which can be obtained hyper-parameterizing the single seasonal multiple sources of error version of (2) using

$$a_{j,t}^{(i)} = s_{j,t}^{(i)} \cos\left(\lambda_j^{(i)} t\right) - s_{jt}^{*(i)} \sin\left(\lambda_j^{(i)} t\right) \tag{10}$$

$$\beta_{j,t}^{(i)} = s_{j,t}^{(i)} \sin\left(\lambda_j^{(i)} t\right) - s_{jt}^{*(i)} \cos\left(\lambda_j^{(i)} t\right) \tag{11}$$

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t'}^{(i)} \tag{12}$$

Where

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos\lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin\lambda_j^{(i)} + \left[k_1^{(i)} \cos(\lambda_j^{(i)} t) + k_2^{(i)} \sin(\lambda_j^{(i)} t)\right] d_t \tag{13}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1} \sin\lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos\lambda_j^{(i)} + \left[k_2^{(i)} \cos(\lambda_j t) - k_1^{(i)} \sin(\lambda_j t)\right] d_t \tag{14}$$

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t'}^{(i)} \tag{15}$$

Equations (16) and (17) are seasonal patterns modeled by the Fourier model.

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos\lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin\lambda_j^{(i)} + \gamma_1^{(i)} d_t \tag{16}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1} \sin\lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos\lambda_j^{(i)} + \gamma_2^{(i)} d_{t'} \tag{17}$$

The notation TBATS$(p, q, \{m_1, k_1\}, \{m_2, k_2\}, \ldots, \{m_T, k_T\})$ is used for these trigonometric models.

2.2. Holt's Linear Trend Method

The exponentially weighted moving average is also the averages of smoothing random variability with the following properties: (1) that is very important that older data have a declining weight; (2) it is very simple to calculate; and (3) the most important for data set is that minimal data is needed. Holt, C.E. 1957 had given equation.

Forecast Equation

$$\hat{y}_{t+h\backslash t} = m_t + h z_t \tag{18}$$

Level Equation

$$m_t = \alpha\, y_t + (1-\alpha)(m_{t-1} + z_{t-1}) \tag{19}$$

Trend Equation

$$b_t = \beta^*(m_t - m_{t-1}) + (1 - \beta^*) z_{t-1} \tag{20}$$

where $m_t$represent an estimate of the level of series at time t, b_t denotes an estimate of the trend (slope) of the series at time t,α is the smoothing parameter for level, $0 \le \alpha \le 1$ and β^* is the smoothing parameter for the trend, $0 \le \beta^* \le 1$ that's with simple exponential smoothing.

2.2.1. Hyperparameters & Overfit

BATS differs from TBATS only in the way it models seasonal effects. In BATS we have a more traditional approach where each seasonality is modeled by that Equation $s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t$ that's mean that BATS can only model integer period lengths. Approach taken in BATS requires m_i seed states for season i, if this season is long the model may become intractable.

TBATS Choose the Final Model Under the hood TBATS will consider various alternatives and fit quite a few models. It will consider models. 1) with Box-Cox transformation and without it.2) with the trend and without it. 3)with Trend Damping and without it. 4) with ARMA(p,q) process used to model residuals non-seasonal model and without it various amounts of harmonics used to model seasonal effects. The final model will be chosen using the Akaike information criterion (AIC). In particular auto, ARMA is used to decide if residuals need modeling and what p and q values are suitable.

Holt's Linear Trend Method, as with simple exponential smoothing, the level equation here shows that m_t is a weighted average of observation and the one-step-ahead training forecast for time $\hat{y}_t$ and the one-step-ahead training forecast for time t. here given by $m_{t-1} + z_{t-1}$. The trend equation shows that z_t is a weighted average of the estimated trend at time t based on $m_t - m_{t-1}$ and z_{t-1} the previous estimate of the trend. The smoothing parameters α and β^*, and the initial values m_0 and z_0 are estimated by minimising the SSE for the one-step training errors.

2.2.2. Overfit problem

If the model is overfitting on training data when the model performs well on the training data but does not perform well on the evaluation data so when we comparing training data with testing data the accuracy of testing data is better than training data

3. RESULT AND DISCUSSION

From table 1, we find that: from 01-07-2015 to 31-12-2016, the visitors to Facebook page in Wikipedia have increased during the period from (65) to (28328). Average daily visitors to the Facebook page in Wikipedia are (13447.63). Kurtosis value is (1.6) indicates the data follows a platykurtic distribution which shows a tail that's thinner than a normal distribution which means the number of outliers will not be large. Followed by a positive value of skewness (0.12) which is between -0.5 and 0.5, the distribution is approximately symmetric.

The Visitors of Energy page in Wikipedia have increased from (16) to (5014) during the same period, with average daily Visitors of Energy page in Wikipedia about (2284). Kurtosis value is (1.6) indicates the data follows a platykurtic distribution which shows a tail that's thinner than a normal distribution which means the number of outliers will not be large. With the positive value of skewness (0.16) which is between -0.5 and 0.5, the distribution is approximately symmetric.

The visitors of Android page in Wikipedia have increased from (8) to (22740) during the same period, with the average daily visitors of Android page in Wikipedia about (11021). Kurtosis value is (1.6) indicates the data follows a platykurtic distribution which shows a tail that's thinner than a normal distribution which means the number of outliers will not be large. With the positive value of skewness (0.01) which is between -0.5 and 0.5, the distribution is approximately symmetric. The Visitors of Apple page in Wikipedia have increased from (4) to (4906) during the same period, with average daily Visitors of Apple page in Wikipedia about (2416). Kurtosis value is (1.7) indicates the data follows a platykurtic distribution which shows a tail that's thinner than a normal distribution which means the number of outliers will not be large. With the positive value of skewness (0.04) which is between -0.5 and 0.5, the distribution is approximately symmetric.

The parameter for the level smoothing is denoted by Alpha, and the parameter for the trend smoothing is denoted by Beta, α and β are constrained to 0-1 with higher values giving faster learning and lower values providing slower learning. From Table 2 it becomes clear that the best values of the level and the trends are 0.999 for level for all the series (Facebook page in Wikipedia), (Energy page in Wikipedia), (Android page in Wikipedia), (Apple page in Wikipedia),meaning fast learning in the day-to-day visits for pages in Wikipedia, and 0.14 for trend for the series (Facebook page in Wikipedia), and 0.56 for trend for the series (Energy page in Wikipedia), and 0.10 for trend for the series (Android page in Wikipedia), and 0.96 for trend for the series (Apple page in Wikipedia), which means slow learning for the trend.

From Table 3, Facebook page in Wikipedia: BATS is the best-suited model (0.845, {0,0}, 1, -), In this model, Box-Cox transformation =0.845, the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). Energy page in Wikipedia: BATS is the best-suited model (0.718, {5,0}, 1, -), In this model, Box-Cox transformation =0.718, the order of ARMA error is (5, 0), the damping parameter = 1 (essentially doing nothing).

Android page in Wikipedia: BATS is the best-suited model (0.854, {0,0}, 1, -), In this model, Box-Cox transformation =0.854, the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). Apple page in Wikipedia: BATS is the best-suited model (0.972, {1,0}, 1, -). In this model, Box-Cox transformation =0.972, the order of ARMA error is (1, 0), the damping parameter = 1 (essentially doing nothing). By comparing table 2 with table 3, we find: The forecasting accuracy by the Holt's Linear Trend model is very high because the values of the (sigma, AIC) were lower than TBATS Model for all series.

Table 4 is given the model fitting of TBATS models in all web pages considered in the study. In Facebook page in Wikipedia: BATS is the best-suited model (1, {0,0}, 1, {<6,2>}).In this model, the damping parameter = 1 (essentially doing nothing).

In Table 5, the best-fitted models based on, lowest values of ME, RMSE, MAE, MPE, MAPE and ACF1, TBATS model is better than Holt's linear trend model for all the series (Facebook page in Wikipedia, Energy page in Wikipedia, Android page in Wikipedia, Apple page in Wikipedia). In other words, the forecasting accuracy by the TBATS model is very high and outperform the forecasting accuracy of Holt's Linear Trend Model, because the values of the accuracy criteria (ME, RMSE, MAE, MPE, MAPE and ACF1) were lower than the values of the accuracy criteria of Holt's Linear Trend Model for all series (Facebook page in Wikipedia, Energy page in Wikipedia, Android page in Wikipedia, Apple page in Wikipedia).

In Table 6, MAPEs are given for Holt’s linear model for all series. From Table 6, it is seen that the 7-day error averages can be ranked as Facebook page in Wikipedia, Android page in Wikipedia, Energy page in Wikipedia, and Apple page in Wikipedia in ascending order. This clearly states that the best forecasting for the Holt’s linear model is Facebook page of Wikipedia.

In Table 7, MAPEs are given for the BATS model for Facebook page in Wikipedia, Energy page in Wikipedia, Android page in Wikipedia, Apple page in Wikipedia. From Table 7, it is easily observed that the 7-day error averages can be ranked as Facebook page in Wikipedia, Android page in Wikipedia, Energy page in Wikipedia, and Apple page in Wikipedia in ascending order. These results show that the best forecasting for BATS linear model is Wikipedia's Facebook page. In Table 8, MAPEs are also presented for the TBATS model for Facebook page in Wikipedia, Energy page in Wikipedia, Android page in Wikipedia, Apple page in Wikipedia. In Table 8, similar results were seen in Tables 6 and 7, and it was seen that all three methods gave consistent results.

4. CONCLUSION

In the projection model, every model is showing a different error rate. In the present investigation; MAPE measure, from training data the best models with least MAPE error are Holt’s linear trend and Bats Models. The best models for Android page in Wikipedia and Apple page in Wikipedia on training data is Holt’s linear trend but the best model for Facebook page in Wikipedia and Energy page in Wikipedia on training data is the BATs model. However, the percentage of error (MAPE) on testing data for last 7 days in the data from (25-12-2016 to 31-12-2016), we concluded that the best model can be forecasting Facebook page in Wikipedia Holt’s Linear trend and Bats models because Facebook page in Wikipedia has least MAPE = 0.019 %, the best model can forecast. Energy page in Wikipedia is Bats model with MAPE = 0.1290%, for Android page in Wikipedia the best model with MAPE =0.0460%, and Apple page in Wikipedia TBATS model with MAPE = 0.34 %. This kind of study helps to understand the web traffic or different web pages trend in the future to invest towards it.

ACKNOWLEDGMENT

The work was supported by the Ministry of Science and Higher Education of the Russian Federation (government order FENU-2020-0022).

HOLT'S LINEAR TREND AND TBATS MODELS

- Holt's Linear Trend Model were fitted using R software.
- BATS and TBATS Model were fitted using R software.
- Data was used in fitting the Models1. (From 01-07-2015 To 31-12-2016) and testing data for last 7 days in the data from (25-12-2016 to 31-12-2016)

Table 1: Descriptive statistics of Visitors pages in Wikipedia

Page	**Mean**	**Minimum**	**Maximum**	**Standard Deviation**	**Skewness**	**Kurtosis**
Facebook page in Wikipedia	13447.63	65	28328	8855.564	0.1255453	1.605639
Energy page in Wikipedia	2284.327	16	5014	1551.386	0.1699649	1.633895
Android page in Wikipedia	11021.37	8	22740	7016.673	0.01107273	1.635909
Apple page in Wikipedia	2416.776	4	4906	1438.541	0.04719163	1.723823

Table 2: Holt's Linear Trend Models fitted for Visitors pages in Wikipedia

Page	Box-Cox transformation	Smoothing parameters		Initial states		Sigma	AIC
	Lambda	**Alpha**	**Beta**	**L**	**B**		
Facebook page in Wikipedia	0.826	0.9999	0.1466	22.1848	15.5383	5.6804	5311.728
Energy page in Wikipedia	0.6822	0.9999	0.566	6.8721	1.65	0.9233	3338.653
Android page in Wikipedia	0.78	0.9999	0.1062	0.4811	8.357	7.9947	5682.872
Apple page in Wikipedia	0.9675	0.9999	0.0297	10.0102	8.236	3.9439	4915.496

Table 3: BATS Model fitted for Visitors pages in Wikipedia

Page	Parameters					prediction error	
	Lambda	**Alpha**	**Beta**	Damping Parameter	AR coefficients	Sigma	AIC
Facebook page in Wikipedia BATS(0.845, {0,0}, 1, -)	0.845331	1.080399	0.1306299	1	-	6.735756	7031.399
Energy page in Wikipedia BATS(0.718, {5,0}, 1, -)	0.717972	1.981873	0.006224636	1	-0.527266 0.662549 0.219562 -0.092203 -0.092795	1.121119	5815.88
Android page inWikipedia BATS(0.854, {0,0}, 1, -)	0.853521	1.409863	-0.01404994	1	-	14.09601	7716.226
Apple page in Wikipedia BATS(0.972, {1,0}, 1, -)	0.971881	0.3868964	-0.004016503	0.999781	0.705365	4.047286	5181.524

Table 4: TBATS Model fitted for Visitors pages in Wikipedia

Page	Parameters			Gamma-1 Values	Gamma-2 Values	Sigma	AIC
	Alpha	Beta	Damping Parameter				
Facebook page in Wikipedia TBATS(1, {0,0}, 1, {<6,2>})	1.075378	0.1329358	1	4.842195e-05	-0.0005057618	27.94339	7055.907
Energy page in Wikipedia TBATS(1, {0,0}, 1, {<6,2>})	1.130439	0.3640595	1	-0.0005904406	-0.0006112289	10.30364	5972.422
Android page in Wikipedia TBATS(1, {0,0}, 1, {<6,2>})	1.478911	0.01387797	1	-0.002673922	0.0006009636	52.69904	7744.883
Apple page in Wikipedia TBATS(1, {0,0}, 1, {<6,2>})	1.039657	0.01173392	1	-0.003746758	0.0009470691	5.010726	5189.507

Table 5: Fitted Holt's Linear Trend, BATS and TBATS Model fitted for Vistorspages in

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
	Holt's Linear Trend Model						
Facebookpage in Wikipedia	-0.382917	28.20683	15.0246	-0.02479222	0.3233684	0.291907	0.06510676
Energy page in Wikipedia	-0.05710684	10.44684	4.623831	-0.01219886	0.539105	0.5077221	-0.0184911
Android page in Wikipedia	-0.3773315	58.70418	20.22983	-0.1105207	0.7482277	0.4897302	0.3402883
Apple page in Wikipedia	-0.1331319	5.00807	3.324313	0.1261135	0.6548812	0.3745121	0.02402595
	BATS Model						
Facebookpage in Wikipedia	-0.4404807	28.12292	14.98003	-0.05470518	0.3211768	0.2910412	0.001672377
Energy page in Wikipedia	0.1745692	9.689009	4.375281	0.1208824	0.5060333	0.48043	-0.04764241
Android page in Wikipedia	0.9798254	52.52552	18.28792	-0.1709344	0.874007	0.4427197	0.07114531
Apple page in Wikipedia	-0.02153136	4.976799	3.269713	0.6647574	1.135035	0.368361	-0.06509243
	TBATS Model						
Facebookpage in Wikipedia	0.04114569	27.94339	14.91394	-0.009450621	0.382524	0.2897572	-5.82108e-05
Energy page in Wikipedia	0.0346115	10.30364	4.601217	0.09734302	0.5599713	0.505239	0.02425513
Android page in Wikipedia	1.855597	52.69904	19.40097	0.4709047	1.305884	0.4696648	0.008057886
Apple page in Wikipedia	0.08187205	5.010726	3.299648	0.8127762	1.243233	0.3717333	0.002301955

Table 6: MAPE Holt's Linear Trend Model fitted for Vistors pages in Wikipedia

Date	Day	Facebook page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	28011	28011.96	0.003 %
2016-12-26	Monday	28059	28061.94	0.01 %
2016-12-27	Tuesday	28111	28111.94	0.003 %
2016-12-28	Wednesday	28161	28161.95	0.003 %
2016-12-29	Thursday	28221	28211.97	0.032 %
2016-12-30	Friday	28269	28262.01	0.025 %
2016-12-31	Saturday	28328	28312.07	0.056 %
MAPE				0.019 %

Date	Day	Energy page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	4969	4959.849	0.184 %
2016-12-26	Monday	4976	4967.703	0.167 %
2016-12-27	Tuesday	4985	4975.560	0.189 %
2016-12-28	Wednesday	4994	4983.422	0.212 %
2016-12-29	Thursday	5002	4991.287	0.214 %
2016-12-30	Friday	5005	4999.157	0.117 %
2016-12-31	Saturday	5014	5007.030	0.139 %
MAPE				0.175 %

Date	Day	Android page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	22439	22444.35	0.024 %
2016-12-26	Monday	22479	22491.73	0.057 %
2016-12-27	Tuesday	22519	22539.13	0.089 %
2016-12-28	Wednesday	22585	22586.55	0.007 %
2016-12-29	Thursday	22628	22633.99	0.026 %
2016-12-30	Friday	22666	22681.45	0.068 %
2016-12-31	Saturday	22740	22728.94	0.049 %
MAPE				0.046 %

Date	Day	Apple page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	4826	4823.601	0.05 %
2016-12-26	Monday	4834	4832.202	0.037 %
2016-12-27	Tuesday	4858	4840.803	0.354 %
2016-12-28	Wednesday	4868	4849.405	0.382 %
2016-12-29	Thursday	4883	4858.007	0.512 %
2016-12-30	Friday	4895	4866.610	0.58 %
2016-12-31	Saturday	4906	4875.214	0.628 %
MAPE				0.363 %

Table 7: MAPE BATS Model fitted for Vistors pages in Wikipedia

Date	Day	Facebook page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	28011	28011.88	0.003 %
2016-12-26	Monday	28059	28061.98	0.011 %
2016-12-27	Tuesday	28111	28112.09	0.004 %
2016-12-28	Wednesday	28161	28162.21	0.004 %
2016-12-29	Thursday	28221	28212.35	0.031 %
2016-12-30	Friday	28269	28262.50	0.023 %
2016-12-31	Saturday	28328	28312.67	0.054 %
MAPE				0.019 %
Date	Day	Energy page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	4969	4961.055	0.16 %
2016-12-26	Monday	4976	4971.897	0.082 %
2016-12-27	Tuesday	4985	4982.497	0.05 %
2016-12-28	Wednesday	4994	4994.245	0.005 %
2016-12-29	Thursday	5002	5005.343	0.067 %
2016-12-30	Friday	5005	5017.327	0.246 %
2016-12-31	Saturday	5014	5028.527	0.29 %
MAPE				0.129 %
Date	Day	Android page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	22439	22442.02	0.013 %
2016-12-26	Monday	22479	22483.89	0.022 %
2016-12-27	Tuesday	22519	22525.77	0.03 %
2016-12-28	Wednesday	22585	22567.66	0.077 %
2016-12-29	Thursday	22628	22609.56	0.081 %
2016-12-30	Friday	22666	22651.47	0.064 %
2016-12-31	Saturday	22740	22693.40	0.205 %
MAPE				0.07 %
Date	Day	Apple page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	4826	4823.733	0.047 %
2016-12-26	Monday	4834	4832.427	0.033 %
2016-12-27	Tuesday	4858	4841.091	0.348 %
2016-12-28	Wednesday	4868	4849.735	0.375 %
2016-12-29	Thursday	4883	4858.364	0.505 %
2016-12-30	Friday	4895	4866.981	0.572 %
2016-12-31	Saturday	4906	4875.591	0.62 %
MAPE				0.357 %

Table 8: MAPE TBATS Model fitted for Vistors pages in Wikipedia

Date	Day	Facebook page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	28011	28009.45	0.006 %
2016-12-26	Monday	28059	28063.59	0.016 %
2016-12-27	Tuesday	28111	28117.03	0.021 %
2016-12-28	Wednesday	28161	28167.62	0.023 %
2016-12-29	Thursday	28221	28217.72	0.012 %
2016-12-30	Friday	28269	28263.83	0.018 %
2016-12-31	Saturday	28328	28311.13	0.06 %
MAPE				0.022 %
Date	Day	Energy page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	4969	4961.410	0.153 %
2016-12-26	Monday	4976	4968.871	0.143 %
2016-12-27	Tuesday	4985	4975.989	0.181 %
2016-12-28	Wednesday	4994	4983.894	0.202 %
2016-12-29	Thursday	5002	4990.741	0.225 %
2016-12-30	Friday	5005	4998.485	0.13 %
2016-12-31	Saturday	5014	5007.631	0.127 %
MAPE				0.166 %
Date	Day	Android page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	22439	22442.64	0.016 %
2016-12-26	Monday	22479	22484.13	0.023 %
2016-12-27	Tuesday	22519	22518.31	0.003 %
2016-12-28	Wednesday	22585	22558.48	0.117 %
2016-12-29	Thursday	22628	22607.34	0.091 %
2016-12-30	Friday	22666	22647.34	0.082 %
2016-12-31	Saturday	22740	22685.96	0.238 %
MAPE				0.081 %
Date	Day	Apple page in Wikipedia		
		Actual	Forecasted	Error
2016-12-25	Sunday	4826	4823.971	0.042 %
2016-12-26	Monday	4834	4833.684	0.007 %
2016-12-27	Tuesday	4858	4842.021	0.329 %
2016-12-28	Wednesday	4868	4850.823	0.353 %
2016-12-29	Thursday	4883	4859.679	0.478 %
2016-12-30	Friday	4895	4867.321	0.565 %
2016-12-31		4906	4876.286	0.606 %
MAPE				0.34 %

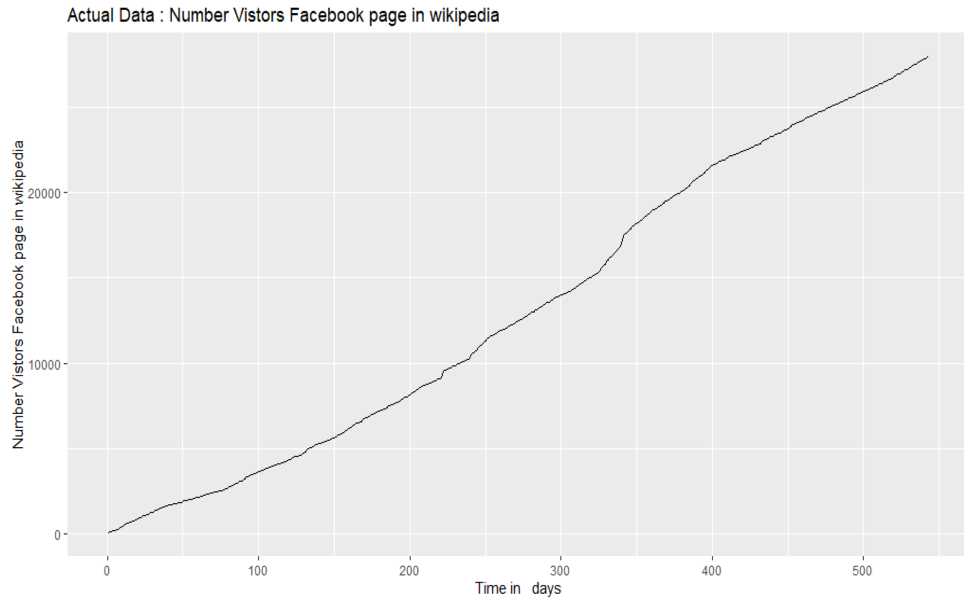

Figure 1 Actual Data of Number vistorsfacebook page in Wikipedia

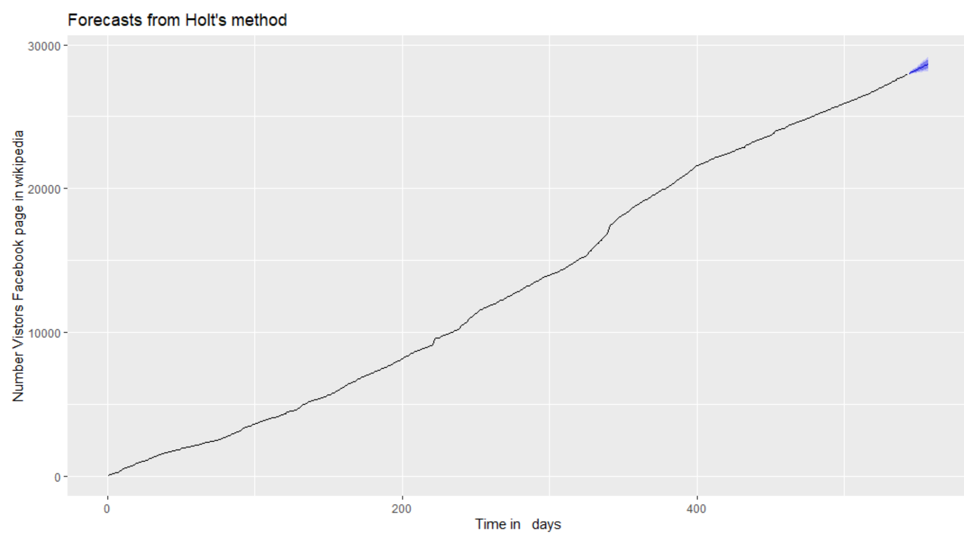

Figure 2 Forecasts From Holt's Method For Facebook page in Wikipedia

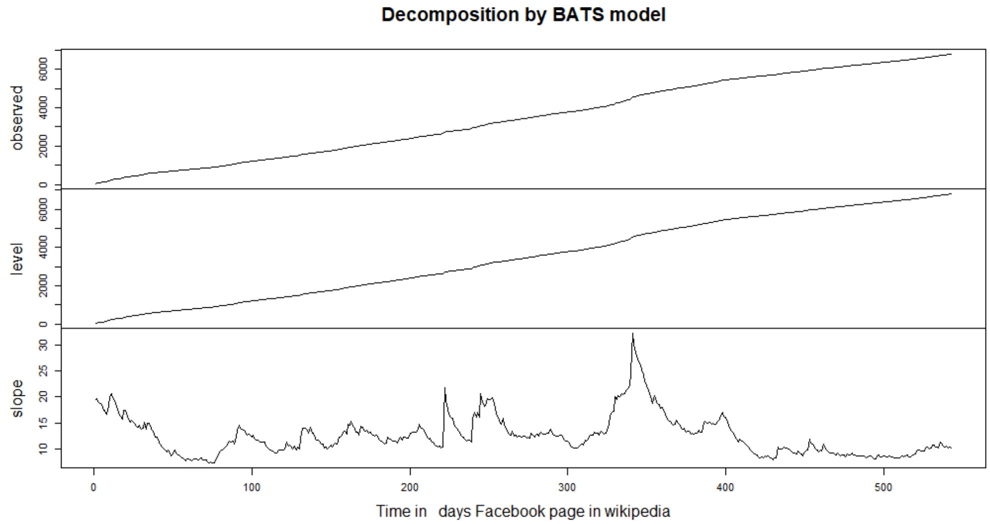

Figure 3 Decomposition by BATS Model Facebook page in Wikipedia

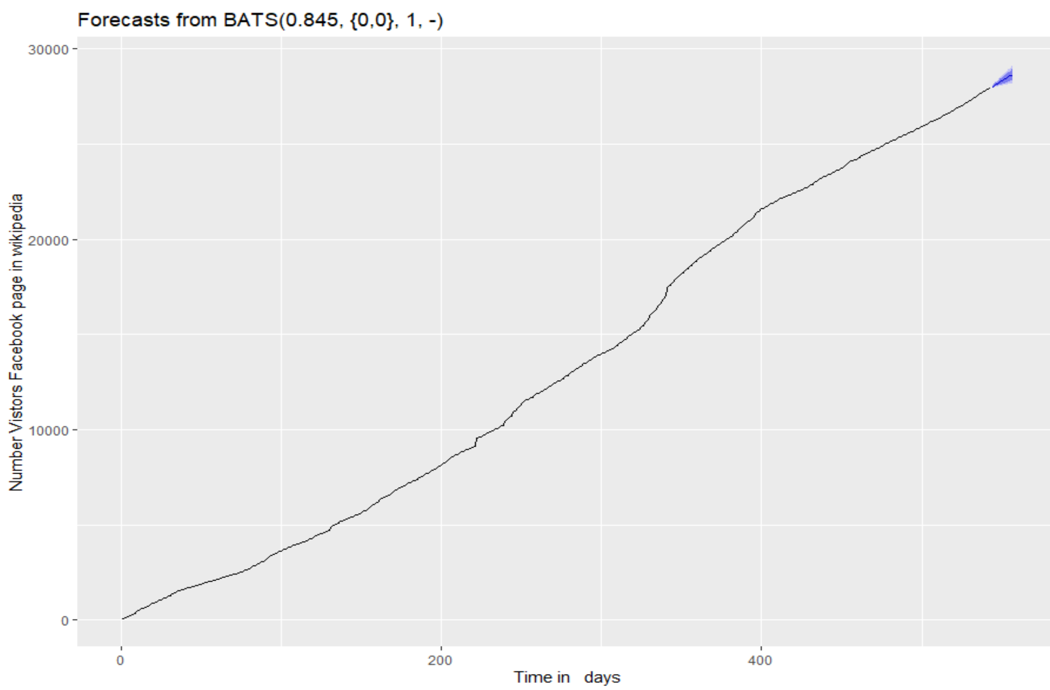

Figure 4 Forecasts From BATS Model in Facebook page in Wikipedia

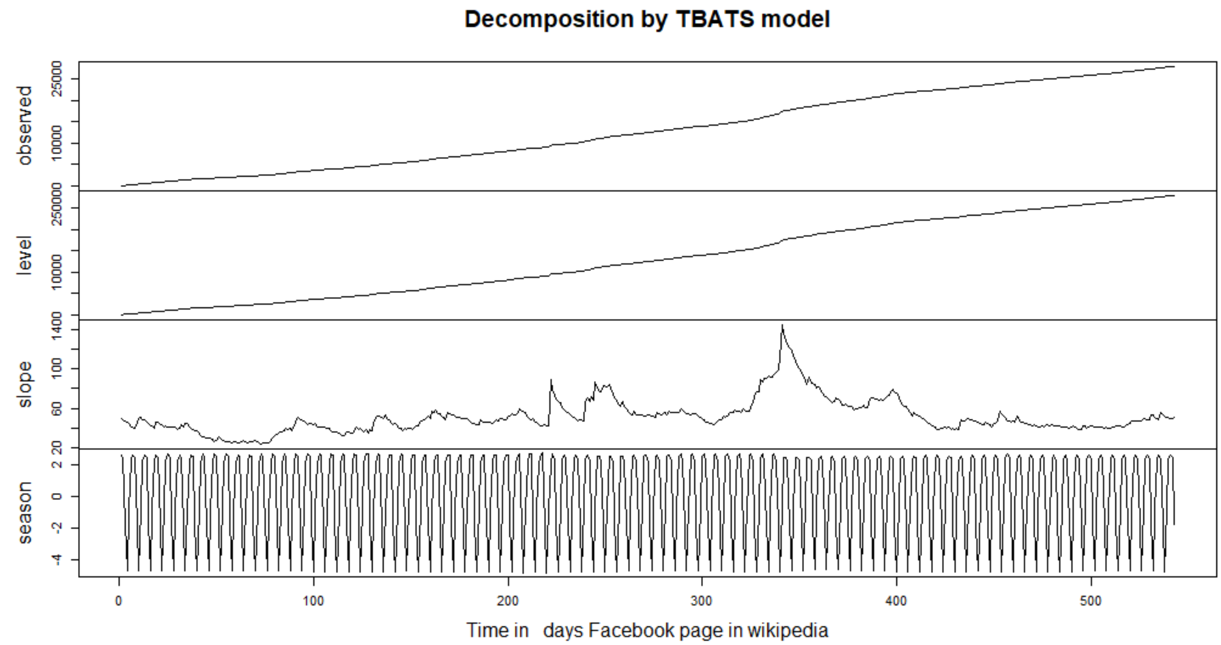

Figure 5 Decomposition by TBATS Model for Facebook page in Wikipedia

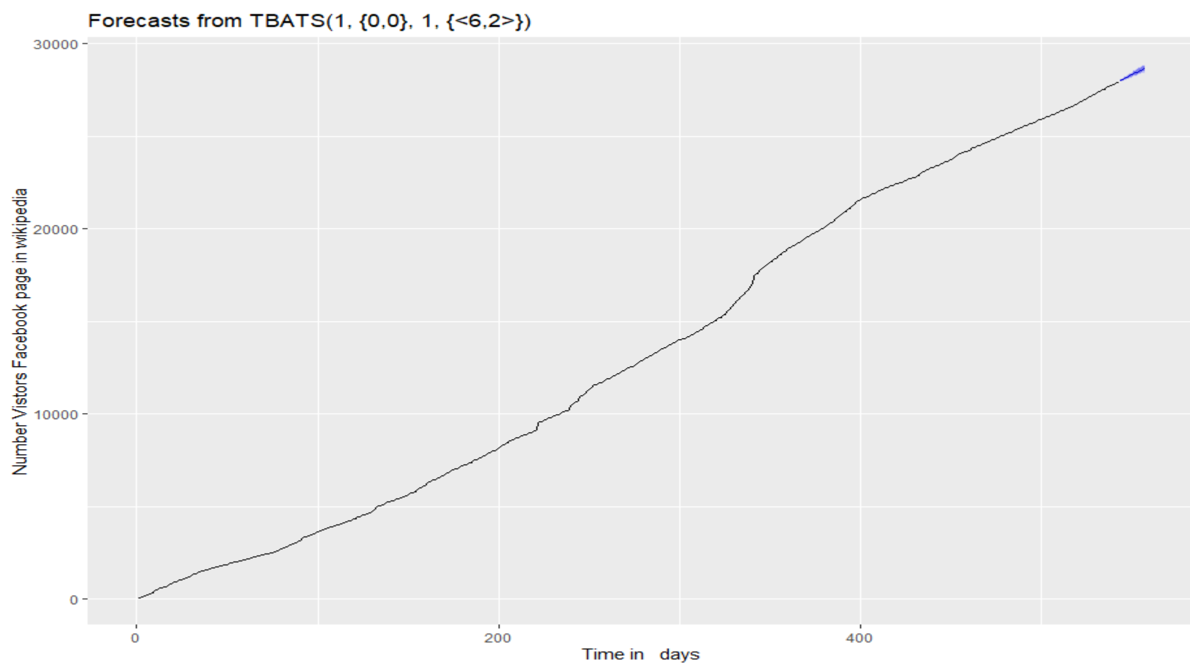

Figure 6 Forecasts from TBATS Model for Facebook page in Wikipedia

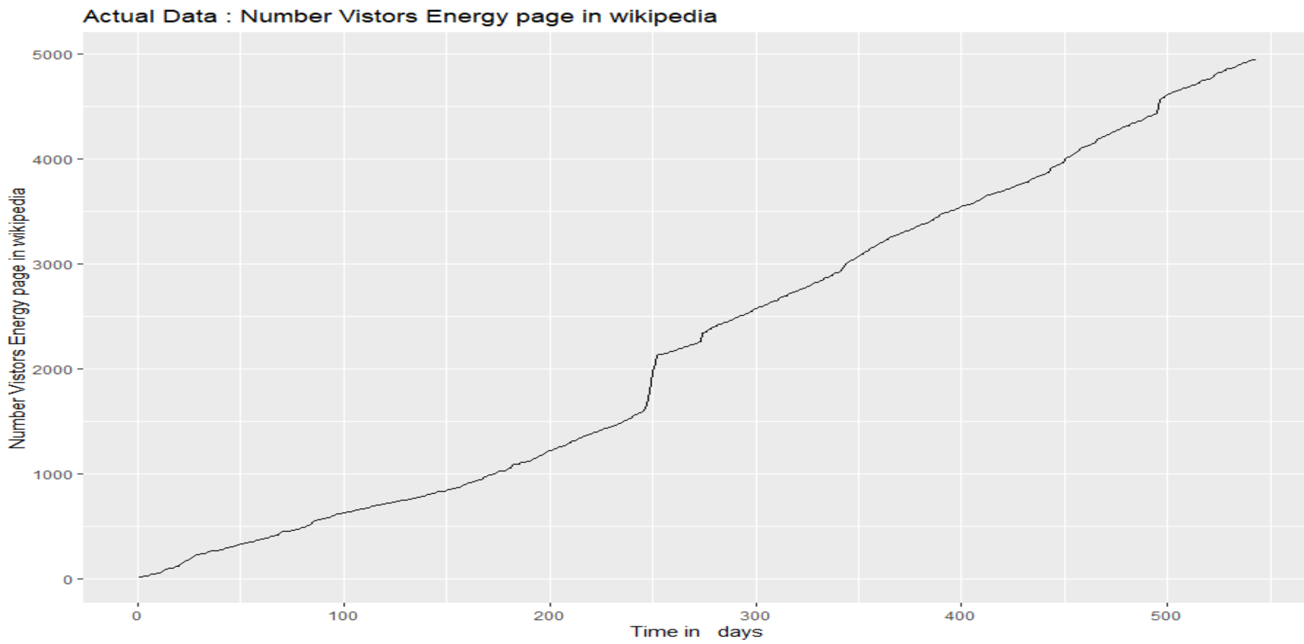

Figure 7 Actual data For Visitors Energy page in Wikipedia

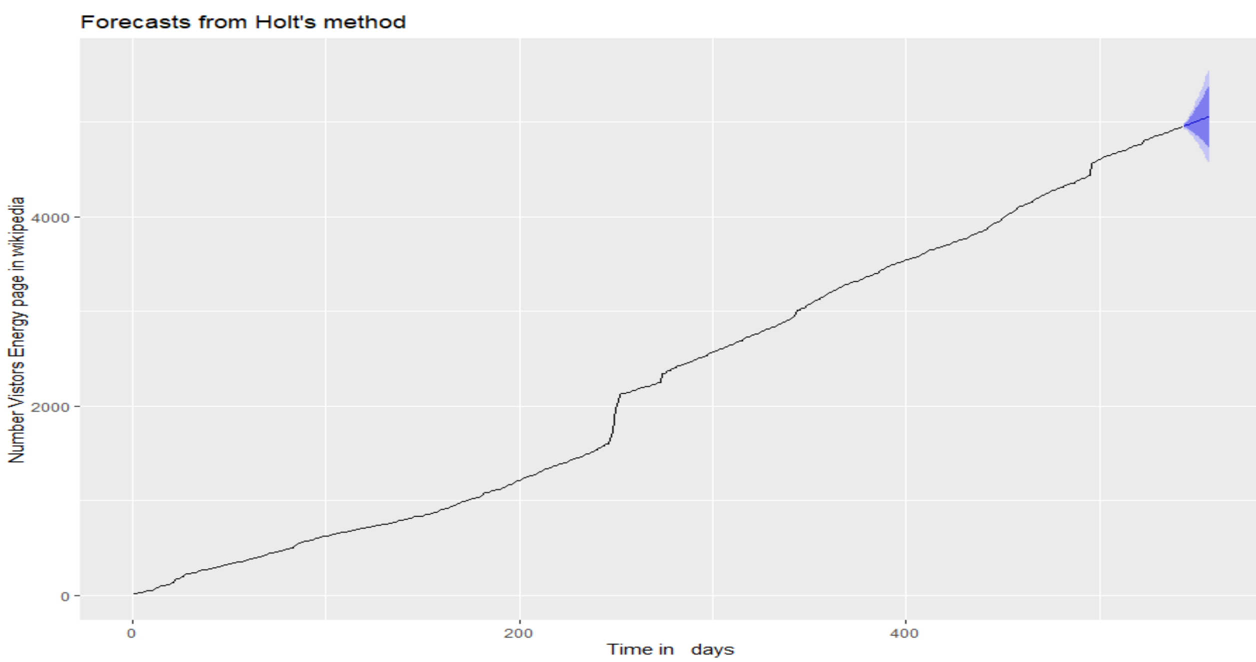

Figure 8 Forecasts from Holt's Method for Energy page in Wikipedia

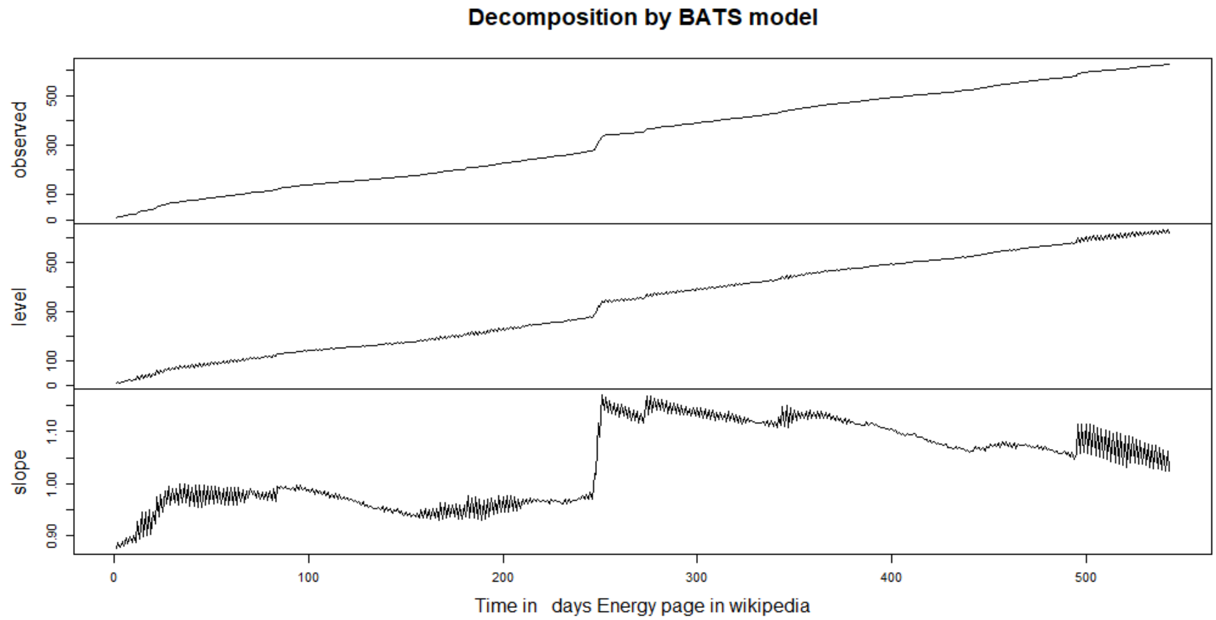

Figure 9 Decomposition by BATS Model for Energy page in Wikipedia

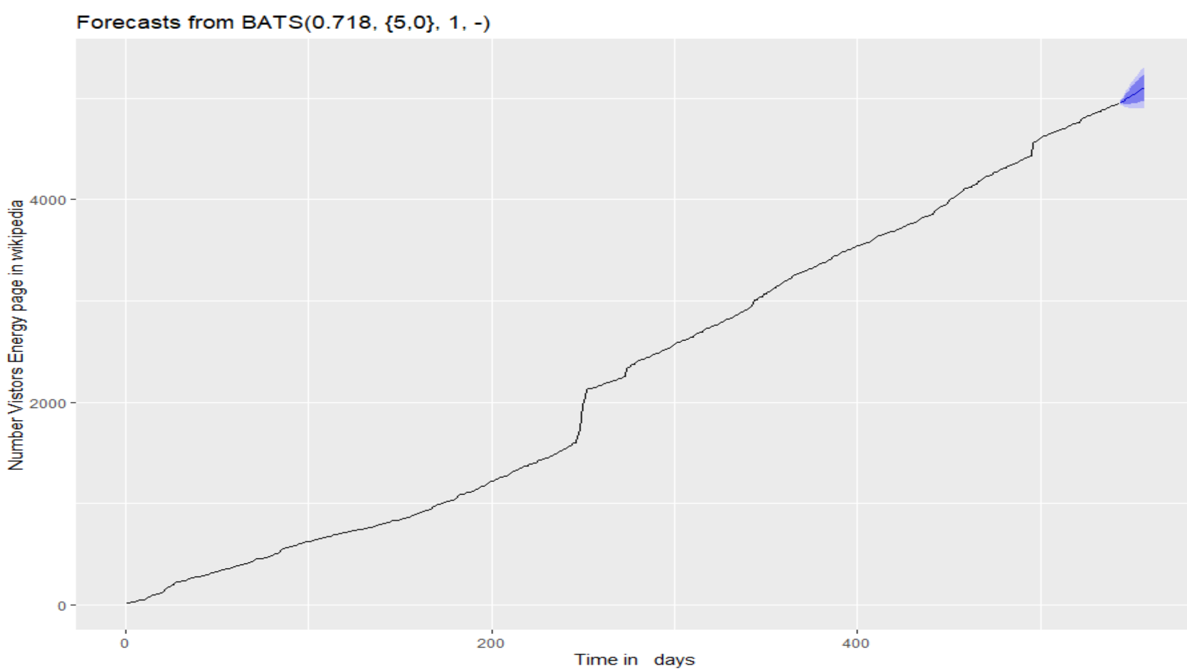

Figure 10 Forecasts from BATS Model for Energy page in Wikipedia

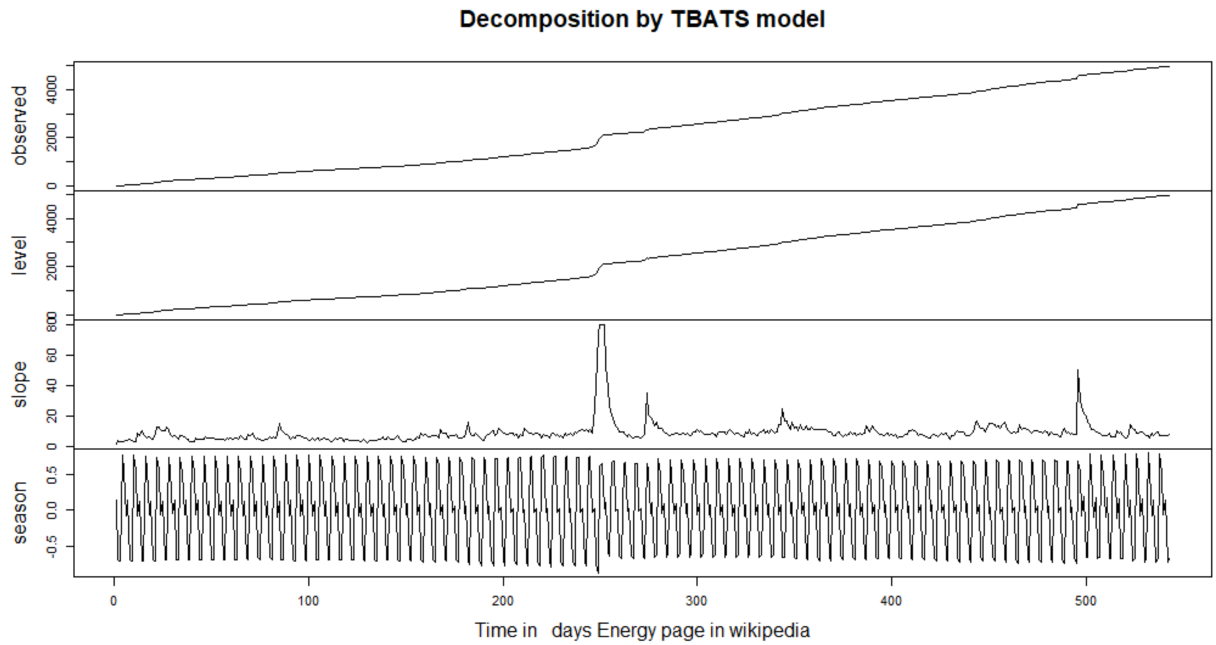

Figure 11 Decomposition by TBATS Model for Energy page in Wikipedia

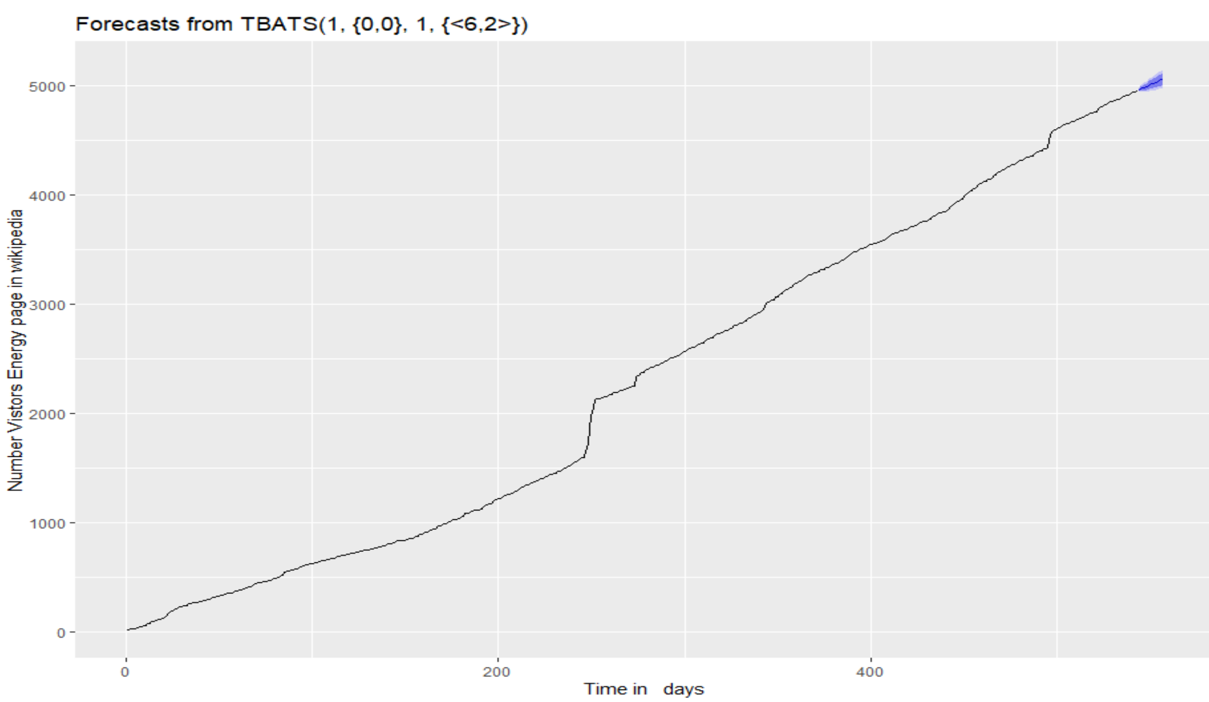

Figure 12 Forecasts from TBATS Model for Energy page in Wikipedia

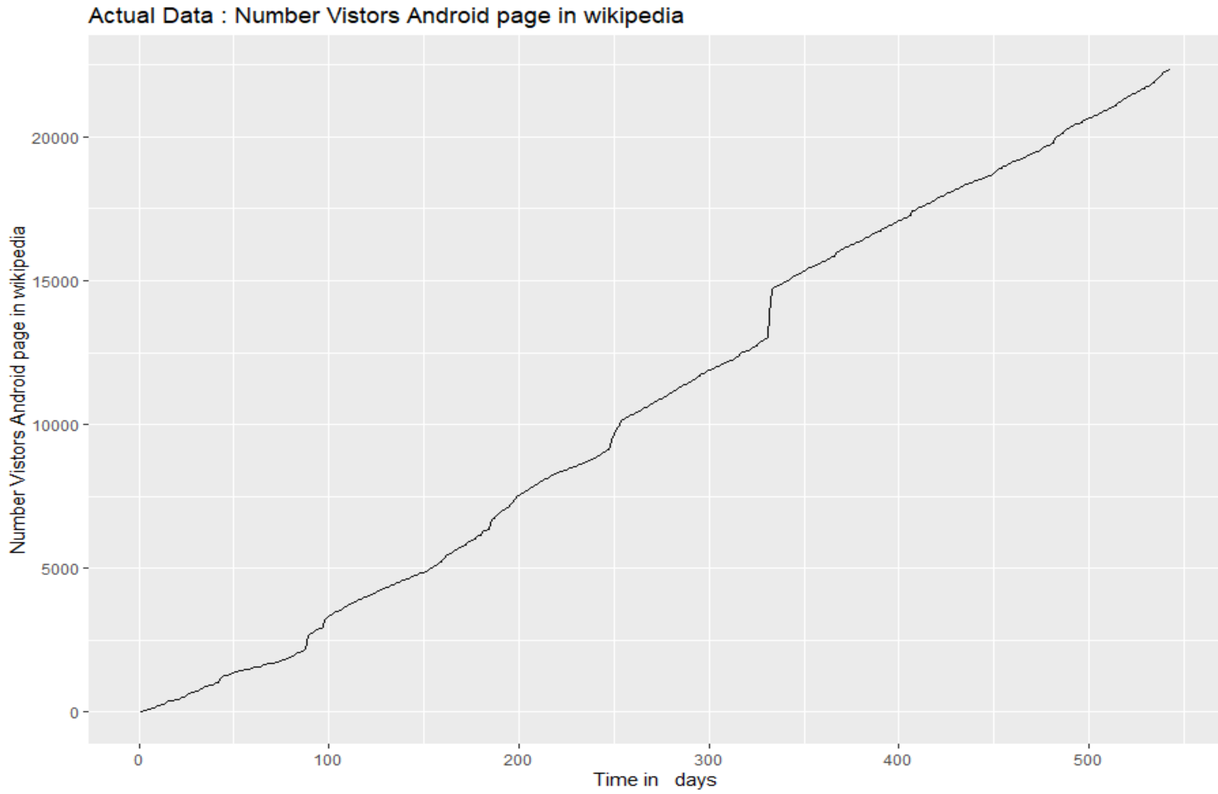

Figure 13 Actual data For Visitors for Androidpage in Wikipedia

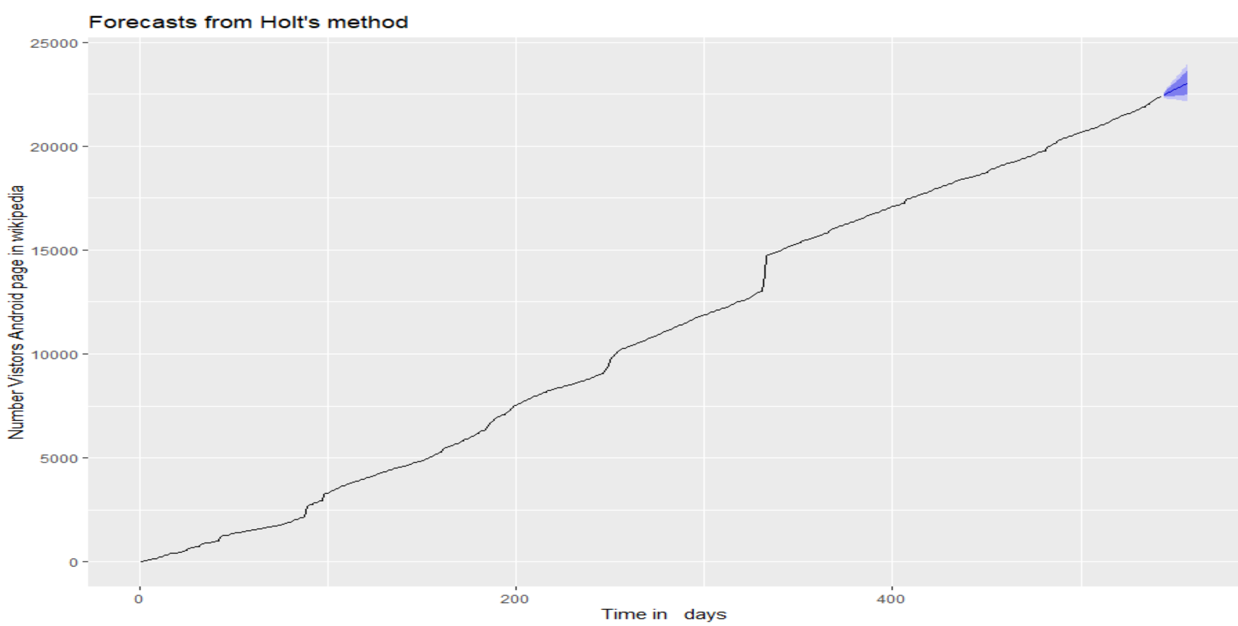

Figure 14 Forecasts from Holt's Method for Android page in Wikipedia

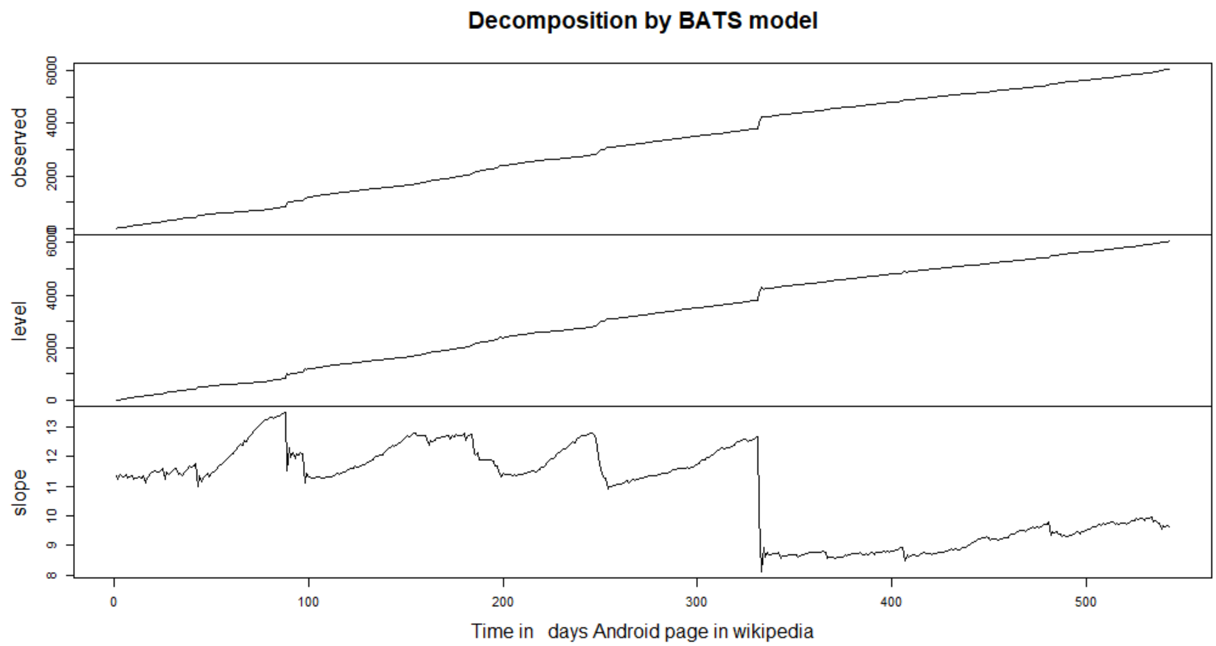

Figure 15 Decomposition by BATS Model for Android page in Wikipedia

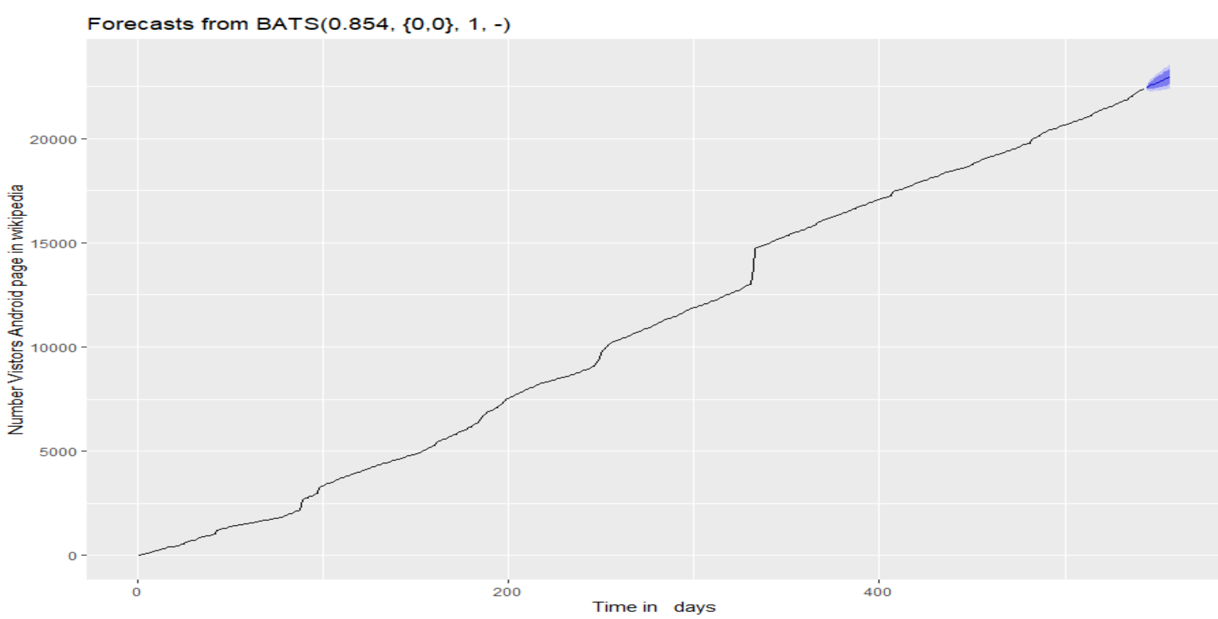

Figure 16 Forecasts from BATS Model for Android page in Wikipedia

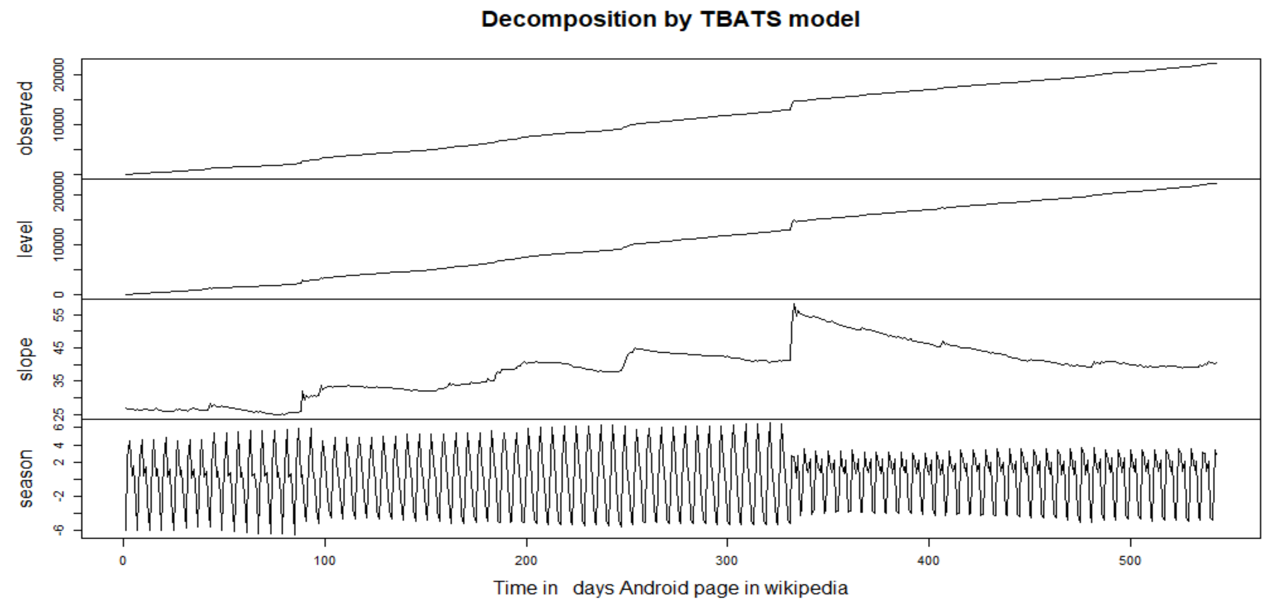

Figure 17 Decomposition by TBATS Model for Android page in Wikipedia

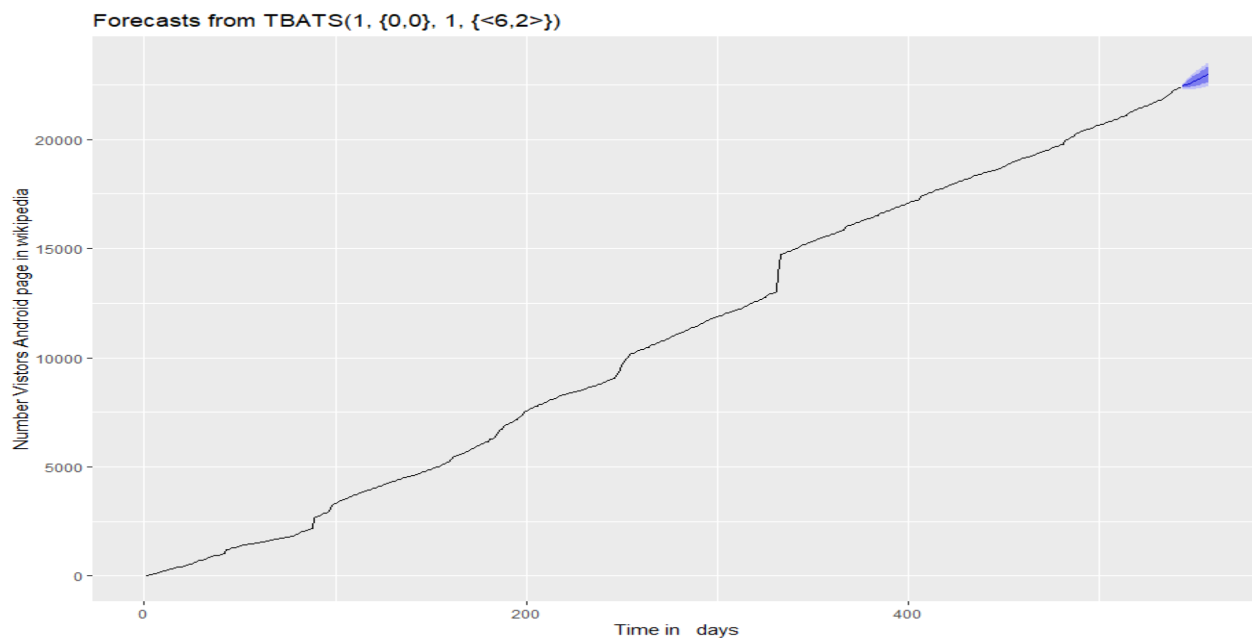

Figure 18 Forecasts from TBATS Model for Android page in Wikipedia

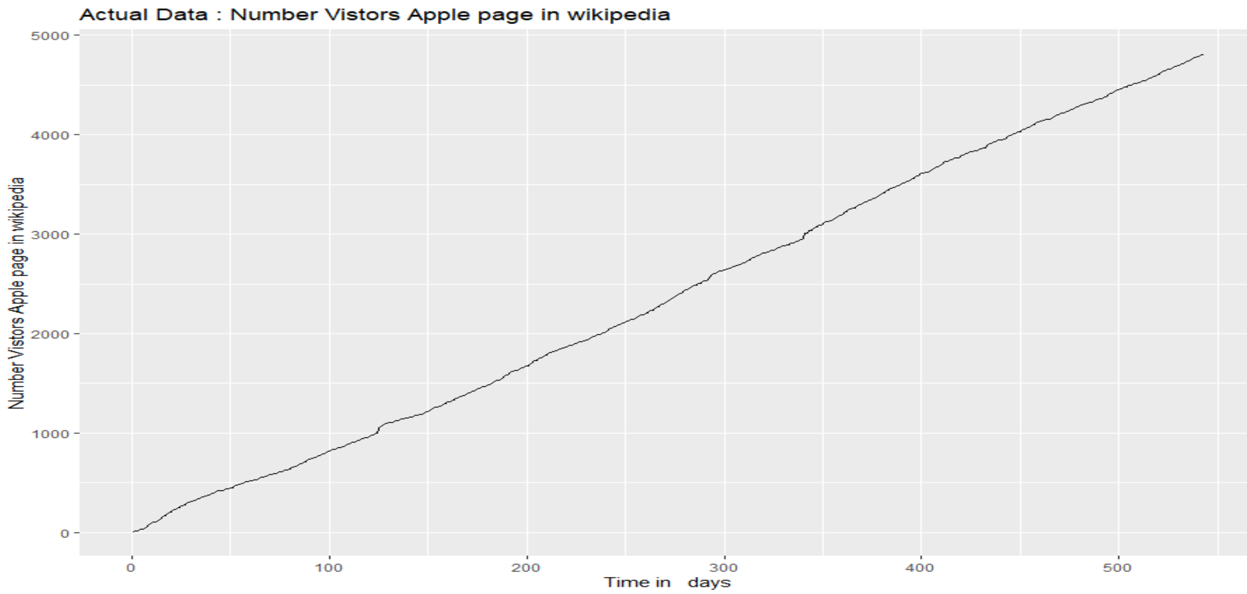

Figure 19 Actual dara for Visitors Apple Page in Wikipedia

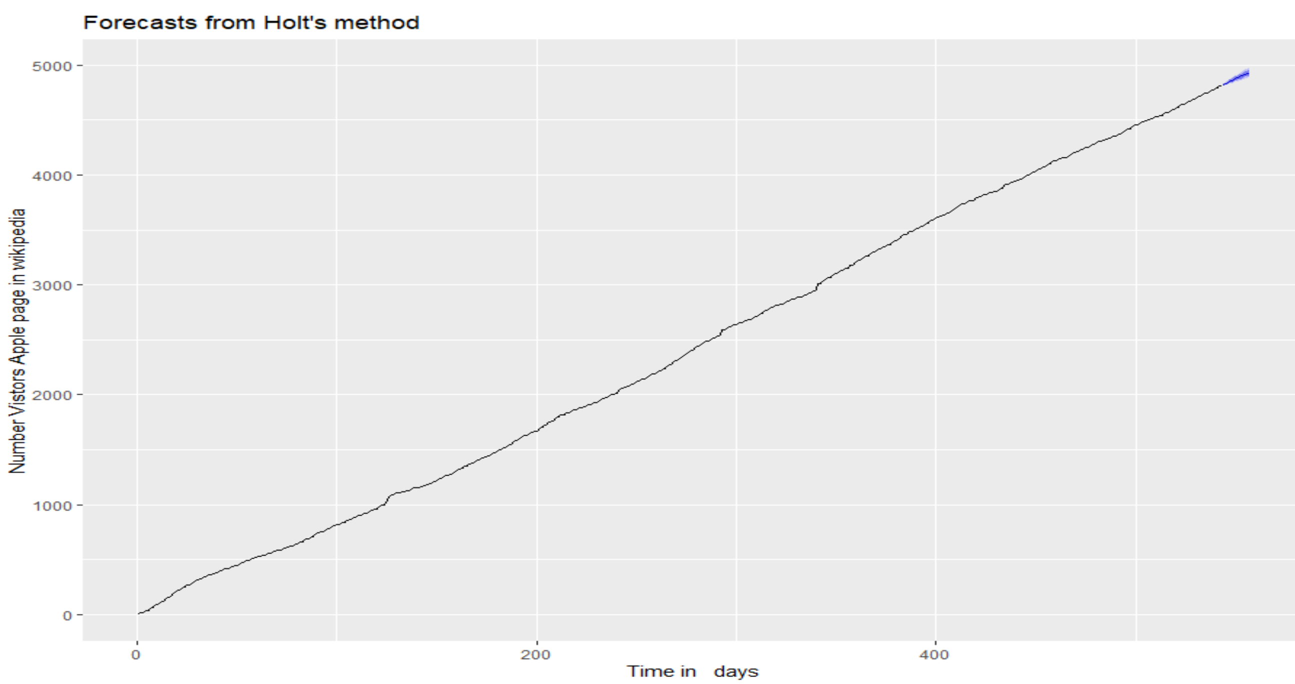

Figure 20 Forecasts from Holt's Method for Apple Page in Wikipedia

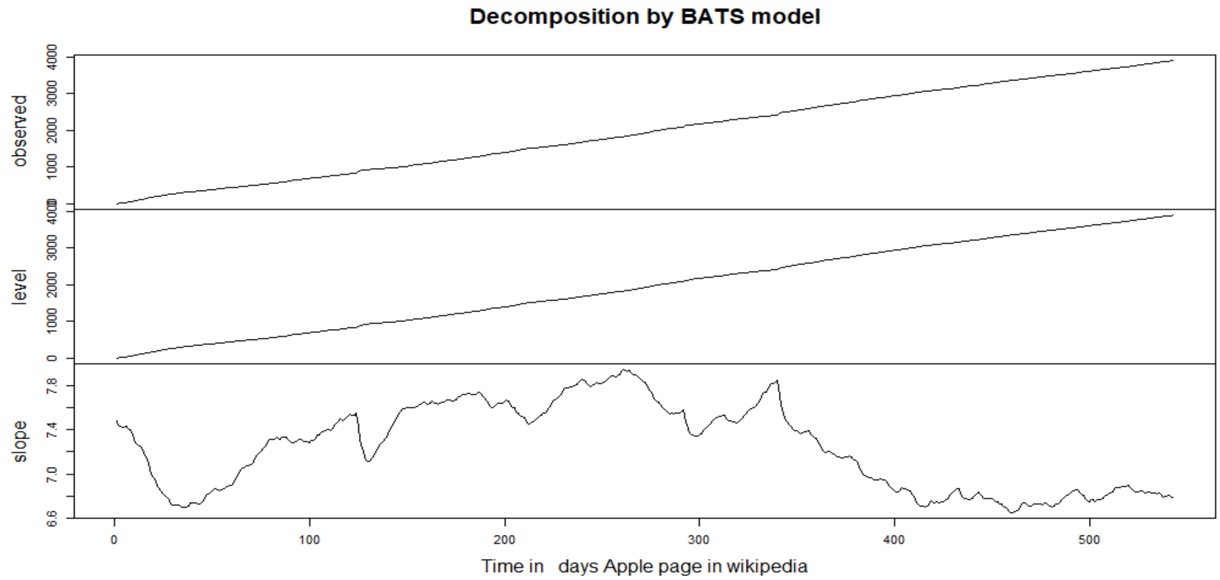

Figure 21 Decomposition by BATS Model For Apple Page in Wikipedia

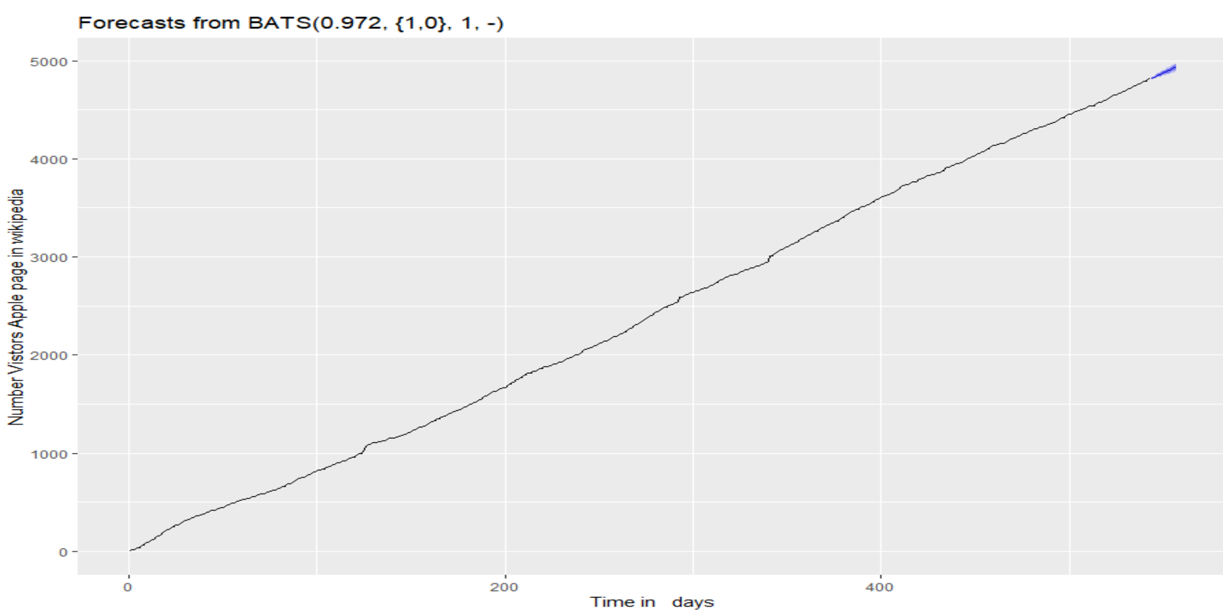

Figure 22 Forecasts from BATS Model for Apple Page in Wikipedia

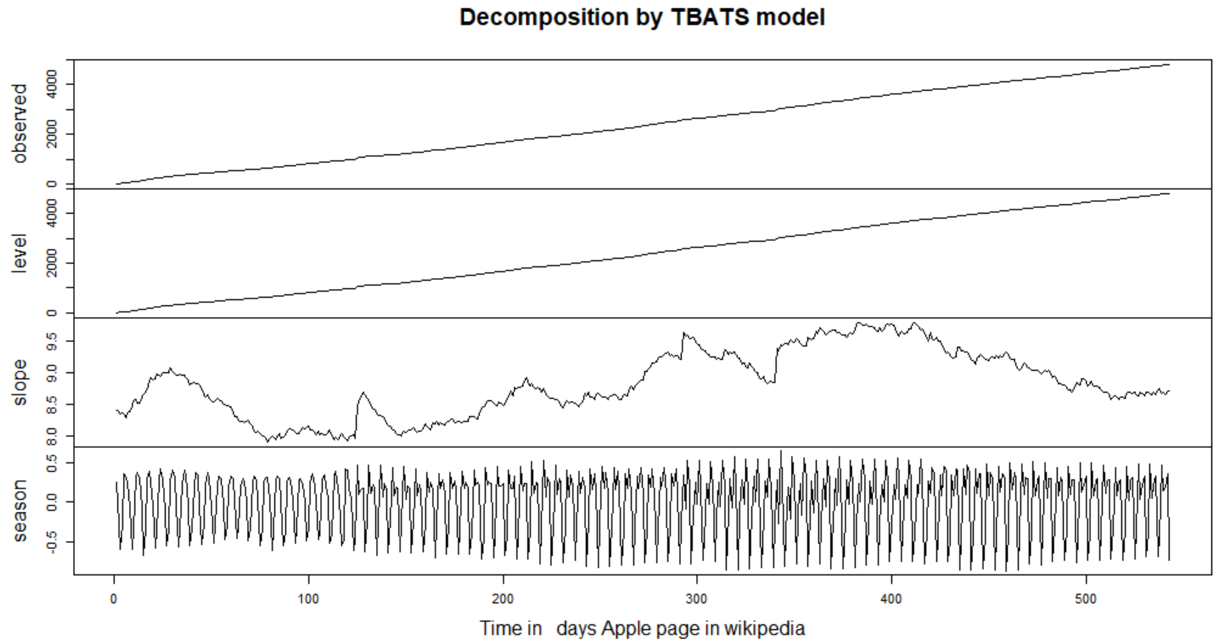

Figure 23 Decomposition by TBATS Model for Apple Page in Wikipedia

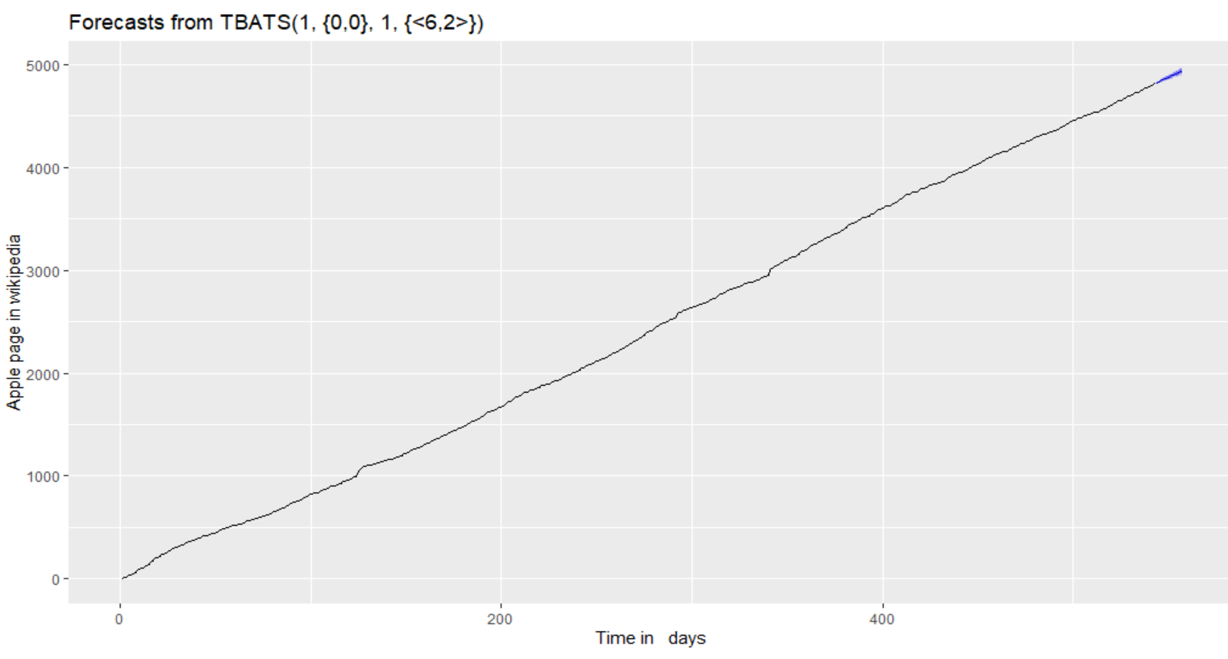

Figure 24 Forecasts From TBATS Model For Apple Page in Wikipedia

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] S. Yao, C. Hu, M. Sun, Prediction of web traffic based on wavelet and neural network, in: 2006 6th World Congress on Intelligent Control and Automation, IEEE, Dalian, China, 2006: pp. 4026–4028.

[2] N. Petluri, E. Al-Masri, Web traffic prediction of wikipedia pages, in: 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA,2018: pp.5427-5429.

[3] K.P. Paun, C.H. Makwan, A survey on web traffic forecasting on time series data, J. Appl. Sci. Comput. 6 (2019), 3588-3594.

[4] T. Shelatkar, S. Tondale, S. Yadav, S. Ahir, Web traffic time series forecasting using ARIMA and LSTM RNN, ITM Web Conf. 32 (2020), 03017.

[5] https://www.kaggle.com/c/web-traffic-time-series-forecasting

[6] https://www.mediawiki.org/wiki/API:Changing_wiki_content

[7] M.F. Iqbal, M. Zahid, D. Habib, L.K. John, Efficient prediction of network traffic for real-time applications, J. Computer Netw. Commun. 2019 (2019), 4067135.

[8] M.S.A. Abotaleb, Predicting COVID-19 cases using some statistical models: an application to the cases reported in China Italy and USA, Acad. J. Appl. Math. Sci. 6(4) (2020), 32-40.

[9] R.J. Hyndman, G. Athanasopoulos, Forecasting: principles and practice. OTexts, New York, (2014).

[10] A.M. De Livera, R.J. Hyndman, R.D. Snyder, Forecasting time series with complex seasonal patterns using exponential smoothing, J. Amer. Stat. Assoc. 106 (2011), 1513–1527.

[11] J.W. Taylor, Exponential smoothing with a damped multiplicative trend, Int. J. Forecast. 19 (2003), 715–725.

[12] E.J. Hannan, R.D. Terrell, N.E. Tuckwell, The seasonal adjustment of economic time series. Int. Econ. Rev. 11(1) (1970), 24-52.

[13] A.C. Harvey, Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge, (1990).

[14] J.Durbin, S.J. Koopman, Time series analysis by state space methods. Oxford University Press, Oxford, (2012).

[15] T. Proietti, Comparing seasonal components for structural time series models. Int. J. Forecast. 16(2) (2000), 247-260.

[16] S.J.Koopman, K.M.Lee,Measuring Asymmetric Stochastic Cycle Components. Available at SSRN: https://dx.doi.org/10.2139/ssrn.785024. (2005).

[17] C.C. Holt, Forecasting seasonals and trends by exponentially weighted averages. In Proceedings of the O.N.R. Memorandum, Carnegie Institute of Technology, Pittsburgh, (1957).

[18] https://www.kaggle.com/c/web-traffic-time-series-forecasting

[19] https://github.com/abotalebmostafa11/wta

[20] V. Cardellini, M. Colajanni, P.S. Yu, Dynamic load balancing on web-server systems, IEEE Internet Comput. 3 (1999), 28–39.

[21] P. Mishra, C. Fatih, H.K. Niranjan, S. Tiwari, M. Devi, A. Dubey, Modelling and forecasting of milk production in Chhattisgarh and India, Indian J. Animal Res. 54(7) (2020), 912-917.

[22] P. Mishra, A. Matuka, M.S.A. Abotaleb, W.P.M.C.N. Weerasinghe, K. Karakaya, S.S. Das, Modeling and forecasting of milk production in the SAARC countries and China, Model. Earth Syst. Environ. (2021). https://doi.org/10.1007/s40808-021-01138-z.