Forecasting Tourist Arrival To Bali-Indonesia From 3 Continents Using Thief-MLP Hybrid Method

Kadek Jemmy Waciko, Ismail, B.

Abstract: In this study, four different method namely Thief-MLP Hybrid, Thief-ELM Hybrid, TBATS and Theta methods are adopted and compared for forecasting tourist arrival to Bali-Indonesia from 3 different continents such as the Australia continent, Europe continent, and Asia continent. To evaluate performance of different methods, the criteria like the Mean Absolute Percentage Error (MAPE) and The Mean Absolute Scaled Error (MASE) are used. This study demonstrates that Thief-MLP Hybrid method outperforms Thief-ELM Hybrid, TBATS and Theta methods. Thus, we can produce short-term forecasts and give a contribution to exploring the best performances of the Thief-MLP Hybrid method.

Keywords : Thief-MLP Hybrid; Thief-ELM Hybrid; TBATS; Theta; MAPE; MASE.

—————————— ◆ ——————————

I. INTRODUCTION

Who does not know the island of Bali? This tourist destination is the pride of Indonesia because it is quite popular with its beauty that is quite charming which makes it visited a lot, both domestic and foreign tourists. Bali is well known internationally as a tourist destination because of its unique arts and culture and natural beauty. By foreign tourists, Bali is referred to as "an island with a thousand temples", "Morning of the world", and "the last paradise"[34]. Bali is the name of one of the provinces in Indonesia (see Figure 1), and also one of the largest islands that are part of the province of Bali. Besides comprising the island of Bali, the province of Bali also consists of several small islands around it such as Nusa Penida, Nusa Lembongan, Nusa Ceningan, and Serangan Island. Bali has a tropical climate like other regions in Indonesia. The capital of Bali is the city of Denpasar on the south side of the island and the total population of Bali is around 4 million people, with the majority 92.3% being Hindu. The rest are Buddhist, Muslim, Christian, and Catholic. Aside from the tourism sector, the Balinese also depend on agriculture and the fisheries sector. The Subak system in the agricultural sector in Bali is the most famous thing in the world, while other people choose to become artists.

- *Kadek Jemmy Waciko, Department of Statistics, Mangalore University, Mangalagangothri-574 199, Karnataka, India, Email: kadekjemmywaciko@gmail.com.*
- *Dr Ismail, B, Professor of Statistics, YENEPOYA (Deemed to be University), Derlakatte, Mangalore, Karnataka, India, Email: prof.ismailb@gmail.com*

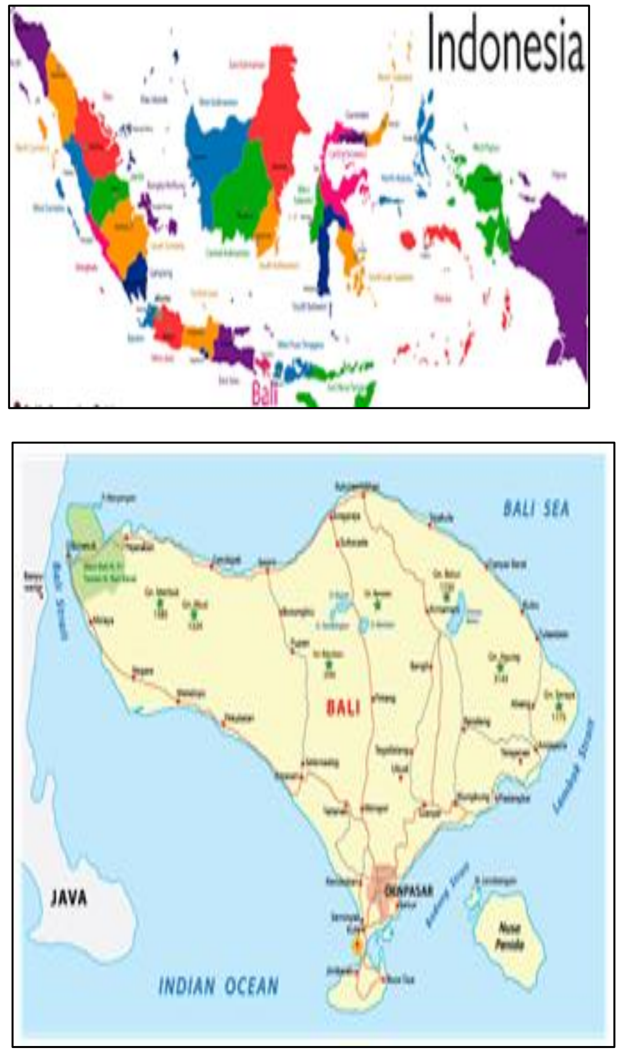

***Figure 1**. Indonesia & Bali Maps*

The latest data obtained and summarized from the Central Bureau of Statistics of the Province of Bali, that for the arrival of foreign tourists (tourists) to Bali in August 2019 recorded as many as 616,706 visits (through the airport as many as 615,027 visits, and through seaports 1,679 visits). in August 2019 the number of foreign tourists to the province of Bali rose as high as 2.02% where the most foreign tourists from Australia (19.22 %), China (17.68 %), France (5.83%), Japan (5.43 %) , and United Kingdom (5.07 %) (we present in Figure

2) while the Room Occupancy Rate of starred hotels in August 2019 was recorded at 67.10%, up 5.39 points compared to the July Room Occupancy Rate which reached 61.71% and the average length of stay of guests in August was 2.83 days, up 0.17 points compared to the average length of stay of guests in the previous month which reached 2.66 days.

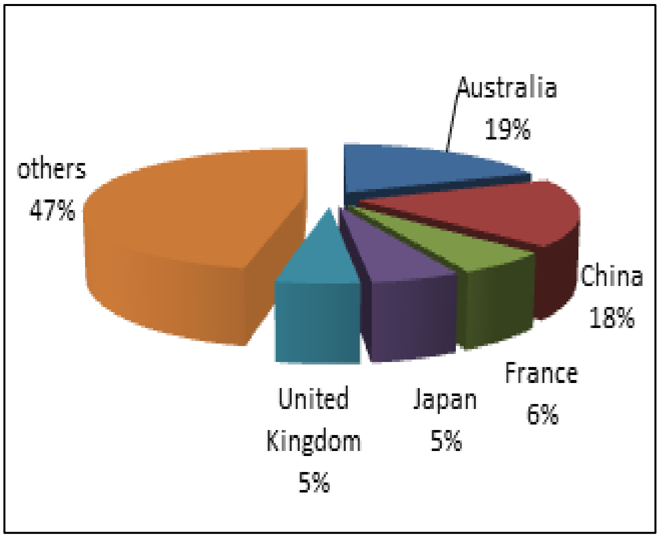

Figure 2. *Percentage of Tourist Arrival to Bali in August 2019*

Forecasting is an important activity in the fields of tourism, economics, marketing, commerce, and various other branches of science. Forecasting methods are procedures for calculating estimates from present and past values. The Theta forecasting method was developed [2] in the M3-Competition. They have decomposed the data into two theta lines with pre-fixed parameters so that they reflect the specific level of improvement in the behavior of short and long term data. Empirical studies of this method have also been examined in several studies, as an overview, we can see in [24] and [27]. Furthermore, [19] and [33] have provided theoretical insights into the Theta Method. While [12] has generalized the dynamically optimized Theta method so that the Theta method has progressed and has received more attention from the forecasting community, given its simplicity and superior forecasting performance. There are several models developed with the aim of accommodating more complex seasonal patterns, as an overview we can see in [15], [16], [26], [31], [13], and [32], unfortunately, there is no model that can handle all the complexities problem as mentioned above. Finally, [11] succeeded in developing the TBATS method that are able to tackle all seasonal complexities or the multiseasonality problem. Conventional models such as the ARIMA and SARIMA models [4] and [5] are very popular to be used to predict tourism demand. On the other hand, Artificial Neural Networks (ANN) have also grown rapidly in estimating tourism demand [30]. This can be seen in [7], [9], and [21]. Furthermore, [18] have succeeded in creating a new learning algorithm for the Single Hidden Layer Feedforward Neural Network (SLFN) architecture called Extreme Learning Machine (ELM). Meanwhile, [10] demonstrated that Multi-Layer Perceptron-Neural Networks (MLP-NN) model is better than the Elman network, similar things can also be seen in [34] where empirical study concluded that MLP -NN outperforms the SARIMA and Extreme Learning Machine (ELM) models. In the development of MLP and ELM models, [1] proposed Thief method framework that is based on using temporal hierarchies for time series forecasting. The selection of the forecasting model is an important criterion that will influence the forecasting accuracy [25]. Many time series models have been applied to forecasts for tourism demand but no single forecasting method has been found to outperform all others in all situations [28] and for forecasters, it's difficult to choose the right technique for their unique situations [35]. This paper focuses on forecasting methods based on the use of time-series analysis. We consider using univariate methods where forecasts depend only on the present and past values of the single series being forecasted [8]. So far there is no research results have been found yet especially develop Thief-MLP Hybrid and Thief-ELM Hybrid methods in tourism demand. In this study 4 different method namely Thief-MLP Hybrid, Thief-ELM Hybrid, TBATS and Theta methods are adopted and investigated. The aim of this study is to compare and select an appropriate forecasting method for forecasting tourist arrivals to Bali-Indonesia from 3 different continents such as the Australia continent, the Europe continent, and the Asia continent.

II. METHODOLOGY

A. Theta Method

In this paper, we using Theta method adopted from [19] where this method founded by [2].

$$Z_{t,\theta} = p_\theta + q_\theta(t-1) + \theta X_t \quad (t = 1,\ldots,n) \tag{1}$$

Where $Z_{t,\theta}$ represents a theta line and equivalent to a linear function of X_t with a linear trend added, for a fixed θ, p_θ and q_θ are constants.

From the observed univariate time series. $\{Z_1, Z_2, \ldots, Z_n\}$, construct $\{Z_{1,\theta}, \ldots, Z_{n,\theta}\}$ as a new series

$$Z''_{t,\theta} = \theta X''_t, \quad t = 3,\ldots,n \tag{2}$$

Where $Z''_{t,\theta}$ represents a second-order difference equation and has the solution [20]; X''_t represents the second differences of X_t and $Z''_{t,\theta}$ represents the second difference of $Z_{t,\theta}$.

B. TBATS Method

Reference [11] developing the TBATS method (Trigonometric, Box-Cox Transformation, ARMA errors,Trends, and Seasonal Components) with arguments TBATS $(\omega, \{p,q\}, \varphi, \{<m_1,k_1>, <m_2,k_2>, \ldots, <m_T,k_T>\})$

The model can be written as

$$Z_t^{(\omega)} = l_{t-1} + \varphi b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + \alpha d_t$$

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

$$s_t^{(i)} = \sum_{j=1}^{kt} s_{j,t}^{(i)}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos\lambda_j^{(i)} + s_{j.t-1}^{*} \sin\lambda_j^{(i)} + \gamma_1^{(i)} d_t$$
$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin\lambda_j^{(i)} + s_{j.t-1}^{*(i)} \cos\lambda_j^{(i)} + \gamma_1^{(i)} d_t$$
$$\lambda_j^{(i)} = \frac{2\pi j}{m_i} \quad (3)$$

Where,
ω is a Box-Cox transformation [3]; p, q are ARMA parameters [6]; φ is a damping parameters [29]; $m_1,..., m_T$ are seasonal periods; $k_1,....., k_T$ are the number of Fourier series pairs [14]; $i = 1,..., T$; d_t is an ARMA; (p,q) process [4]; α, β, γ_1 and γ_2 are smoothing parameters; l_o is the initial level and b_0 is slope value.

C. Thief-ELM & Thief-MLP

Reference [1] introduces Temporal Hierarchies Forecasting (Thief). Suppose in a time series $\{Z_t : t = 1,..., T\}$ observed at the highest available sampling frequency per year, m and in the k-aggregates that can be constructed where k is a factor of m. The various aggregated series can be written as

$$Z_j^{(k)} = \sum_{t=t^*+(j-1)k}^{t^*+jk-1} Z_t \quad (4)$$

For $j = 1,..., [T/k]$ and $M_k = m/k$ is the seasonal period of the aggregated series. The non-overlapping aggregation requires that the total number of observations has to be a multiple of m. To ensure this, we start the aggregation from. We denote the factors of m, in descending order, to be $\{k_p,..., k_3, k_2, k_1\}$ where $k_p = m,\ k_1 = 1,$ and p is the total number of aggregation levels,

The Extreme Learning Machine (ELM) algorithm in [18] as an original way of building a Single Hidden Layer Feedforward Neural Network (SLFN). Consider a set of n distinct samples (x_i, y_i), $1 \le i \le n$, with $x_i \in \Re^p$ and $y_i \in \Re$. A SLFN with m hidden neuron in the hidden layer can be expressed by the following sum

$$\sum_{i=1}^{m} \beta_i f(w_i x_j + b_i),\ \ 1 \le \mathrm{j} \le \mathrm{n}. \quad (5)$$

With β_i the output weights, f an activation function, w_i the input weights and b_i the biases. Denoting by $\hat{y}_i$ the outputs estimated by the SLFN, in the hypothetical case where the SLFN perfectly approximates the actual outputs y_i, the relation is

$$\sum_{i=1}^{m} \beta_i f(w_i x_j + b_i) = y_j,\ \ 1 \le \mathrm{j} \le \mathrm{m}. \quad (6)$$

Which is written in matrix form as $H\beta = y$, with

$$H = \begin{pmatrix} f(w_1 x_1 + b_1) & \cdots & f(w_m x_1 + b_m) \\ \vdots & \ddots & \vdots \\ f(w_1 x_n + b_1) & \cdots & f(w_m x_n + b_m) \end{pmatrix} \quad (7)$$

$\beta = (\beta_1,..., \beta_m)^T$ and $y = (y_1,...., y_n)^T$. The ELM approach is thus to initialize randomly the w_i and b_i and compute the output weights $\beta = H^{\div} y$ by a Moore-Penrose pseudo-inverse of H, $H^{\div}$.

Multi Layer Perceptron (MLP) Neural Network model with a single hidden layer is shown as written:

$$z_t = \beta_0 + \sum_{j=1}^{q} \beta_j f(\gamma_{0j} + \sum_{i=1}^{p} \gamma_{ij} z_{t-i}) + \varepsilon_t \quad (8)$$

Where, $z_{t-i} (i = 1,2, \dots, p)$ are the p inputs and $\hat{z}_t$ is the output; $\beta_j (j = 0,1,2, \dots, q)$ and $\gamma_{ij} (i = 0,1,2, \dots, p; j = 0,1,2, \dots q)$ are the conection weights and ε_t is random error term; The integers p, q are the number of input and hidden nodes each; β_0 and γ_{oj} are the error terms and f is activation function [17]. The main idea of proposed method is the process in (4) can be combine with ELM and MLP methods to become two hybrid methods namely Thief-ELM hybrid and Thief-MLP hybrid methods .

D. The Measures of Forecasting Accuracy

1. Mean Absolute Percentage Error (MAPE)

The Mean Absolute Percentage Error (MAPE) is a statistical measure that measures how accurate a forecasting system can be calculated as the average absolute percent error of the actual value minus forecast value divided by the actual value. MAPE is the most common measure used to forecast error and works best if there are no extreme to data (no zeros). MAPE is defined by the following equation:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Z_t - F_t}{Z_t} \right| \quad (9)$$

Where,
Z_t = The actual value
F_t = The forecast value

2. Mean Absolute Scaled Error (MASE)

The Mean Absolute Scaled Error (MASE) is a scale-free error metric that gives each error as a ratio compared to a baseline's average error. The advantages of MASE include that it never gives undefined or infinite values and so is a good choice for intermittent-demand series (which arise when there are periods of zero demand in a forecast). It can be used on a single series, or as a tool to compare multiple series. MASE is defined by the following equation:

$$MASE = \frac{|e_j|}{\frac{1}{n-1} \sum_{i=2}^{n} |Z_i - Z_{i-1}|} \quad (10)$$

Where:
i = 1,2,…, n is the set of forecasting sample periods.

$e_j = Z_t - F_t$ = Forecast error.

III. RESULT

The data in this study were collected from the Bali Government Tourism Office. Based on the data we obtained

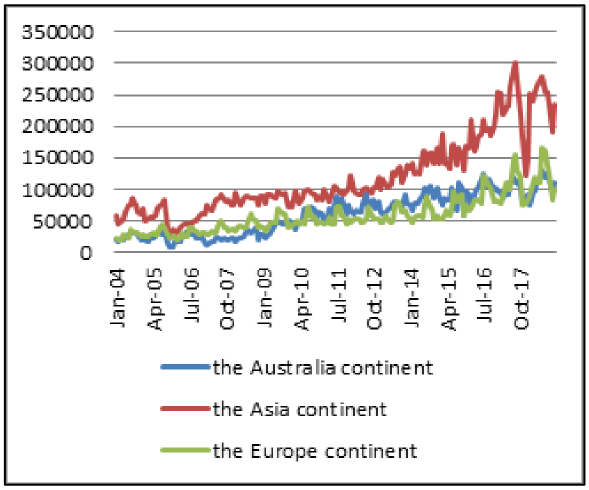

Figure 3. *Plot of tourist arrival to Bali from 3 continents*

we only used 180 monthly observations (January 2004-December 2018) from this data, divided into two parts, namely training data (January 2004-December 2013) and testing data (January 2014-December 2018), these data are tourist arrival to Bali from 3 continents chosen by purpose random sampling (Continents with the most dominant visitor countries coming to Bali) the three continents are the Asia continent, the Australia continent, and the Europe continent. For data analysis, we use R Software. The plot of tourist arrival from the three continents is shown in Figure 3. To assess the performance of forecasting from different methods and to investigate the best estimation method, we consider criteria such as MAPE and MASE. Furthermore, for each dataset, we have presented the results of the analysis of the calculations in the table 1.

Table 1. *. Forecast result of tourist arrival to Bali from 3 Continents*

Continent & Method	*The Measures of Forecasting Accuracy*	
the Europe Continent		
Method	*MAPE*	*MASE*
TBATS	*24.320*	*4.542*
Theta	*66.170*	*3.337*
Thief-MLP hybrid	*16.263*	*3.051*
Thief-ELM hybrid	*18.223*	*3.387*
the Asia Continent		
Method	*MAPE*	*MASE*
TBATS	*27.097*	*4.229*
Theta	*26.437*	*4.055*
Thief-MLP hybrid	*20.088*	*3.077*
Thief-ELM hybrid	*20.250*	*3.120*
the Australia Continent		
Method	*MAPE*	*MASE*
TBATS	*17.453*	*2.002*
Theta	*43.742*	*2.767*
Thief-MLP hybrid	*11.258*	*1.223*
Thief-ELM hybrid	*15.102*	*1.717*

We can see that Thief-MLP hybrid method is the best forecasting method for forecasting tourist arrivals to Bali-Indonesia with all MAPE and MASE measurement results are minimum, for the Europe continent (MAPE = 16.263 and MASE = 3.051); the Asia continent (MAPE = 20.088 and MASE = 3.077); and the Australia continent (MAPE = 11.258 and MASE = 1.223).

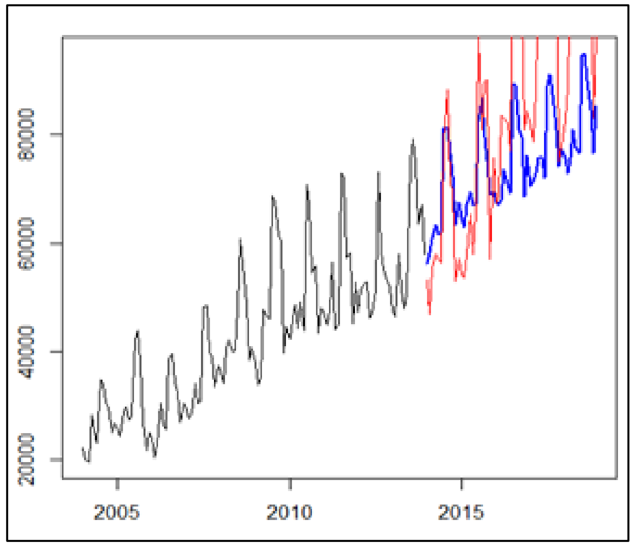

Figure 4. *Forecasting tourist arrival to Bali from the Europe Continent using Thief-MLP hybrid method*

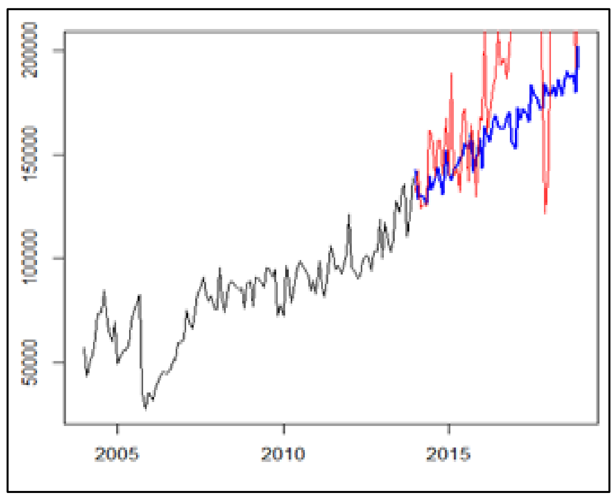

Figure 5. *Forecasting tourist arrival to Bali from the Asia Continent using Thief-MLP hybrid method*

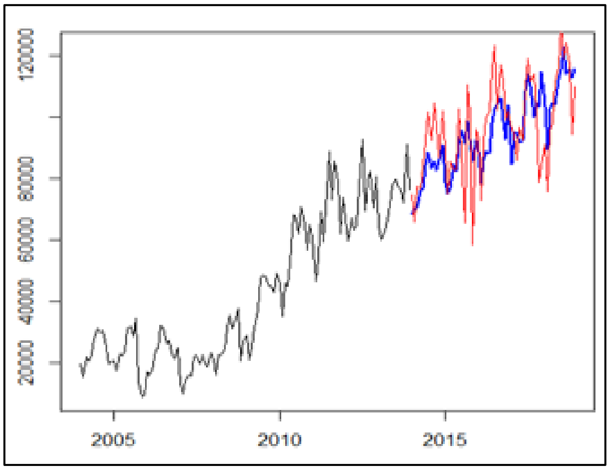

Figure 6. *Forecasting tourist arrival to Bali from the Australia Continent using Thief-MLP hybrid method*

Figure 4 to 6 represent forecast from Thief-MLP for tourist arrivals to Bali from 3 different continents (the Europe continent, the Asia continent, and the Australia Continent) and showed predicted value (a blue line) versus actual value (a red line) using appropriate methods.

IV. CONCLUSIONS

This study can help to show whether the success or failure of particular forecasting methods depends on such factors as the type of data, the length of forecasting horizon, the skill of the analyst is using the method and the numerical methods used to fit a model implement a method and compute predictions. In this paper, we developed Thief-MLP hybrid method for forecasting tourism demand. The forecasting performance of 4 different methods namely Thief-MLP hybrid , Thief-ELM hybrid, TBATS, and Theta methods have been compared under MAPE and MASE criterion. It was found that Thief-MLP hybrid method is the most appropriate method for forecasting tourist arrival to Bali-Indonesia from 3 different continents such as the Australia continent, the Asia continent, and the Europe continent.

REFERENCES

[1] Athanasopoulos, G, Hyndman, R.J., Kourentzes, N. and Petropoulos, F. Forecasting with Temporal Hierarchies, Europe Journal of Operation Research. 262(1), (2017).60-74.

[2] Assimakopoulos, V.& Nikolopoulos, K. The Theta Model: A decomposition approach to Forecasting. International Journal of Forecasting. 16(4), (2000).521-530.

[3] Box, G.E.P & Cox, D.R An analysis of Transformations. Journal of the Royal Statistical Society, Series B, 26(2), (1964). pp 211-252.

[4] Box, G.E.P & Jenkins, G. M Time series analysis: Forecasting and control. San Fransisco: Holden-Day, Revised edn. , (1970).

[5] Box, G.E. P., Jenkins, G.M, and Reinsel, G. Time series analysis: Forecasting and control. 3rd ed., Prectice Hall: New Jersey. C, (1994).

[6] Brockwell, P. J and Davis, R. A., Introduction to time series and forecasting. New York: Springer. (1996).

[7] Burger, C, Dohnal, M., Kathrada, M., and Law, R. a Practitioners guide to time series methods for tourism demand forecasting A case study of Durban, South Africa. Tourism Management, 22, 403-409. (2001).

[8] Chatfield, C.Time-Series Forecasting. Chapman & Hall/CRC:Boca Raton, Florida. (2000).

[9] Cho, V. Acomparison of three different approaches to tourist arrival forecasting. Tourism Management, 24, (2003), 323-330.

[10] Claveria, O., Monte, E., and Torra, S.). Tourism demand forecasting with neural network models: Different Ways of Treating information. International Journal of Tourism research, Int J. Tourism Res, 17: (2015), 492-500.

[11] De Livera, A. M. Hyndman, R. J. and Snydera, R.D. Forecasting time series with complex seasonal patterns using Exponential Smoothing. Journal of the American Statistical Association. 106(496), .(2011). pp.1513-1527.

[12] Fioruci, JA, Pellegrini, TR, Louzada, F, Petropoulos, F & Koehler, AB, Models for optimising the theta method and their relationship to state space models. International Journal of Forecasting, vol. 32, no. 4, (2016) pp.1151-1161.

[13] Gould, P.G., Koehler, A.B., Ord, J.K., Snyder, R.D., Hyndaman. R.J & Vahid-Araghi, F., Forecasting time series with multiple seasonal patterns, Europe Journal of Operational research 191(1), (2008).207-222.

[14] Harvey, A. Forecasting Structural time Series Models and the Kalman Filter. Cambridge Books:Cambridge University Press. (1989).

[15] Harvey, A & Koopman, S.J. Forecasting hourly electricity demand using time varying splines. Journal of the American Statitical Association 88, (1993).1228-1236.

[16] Harvey, A., Koopman, S.J & Riani, M, The modelling and seasonal adjustment of weekly observation. Journal of Business and Economic Statistic, 15, (1997), 354-368.

[17] Haykin, S. Neural Networks: A Comprehensive Foundation. Prentice-Hall, New Jersey, (1999).

[18] Huang, G.B, Zhu, Q.Y & Siew, C.K, Extreme Learning Machine: Theory and Applications. Neurocomputing, 70, (2006). 489-501.

[19] Hyndman, R.J. and Billah, B.. Unmasking the Theta method. International Journal of Forecasting. 19(2), (2003), 287-290.

[20] Kelley, W.G. Difference equation: an introduction with applications, 2nd ed., Academic: London. (2000).

[21] Kon, S.C and Turner W. L. Neural Network Forecasting of tourism demand. Tourism economics, 11, (2005), 301-328.

[22] Makridakis, S., S.C, Wheelwright, & Hyndman. Forecasting Methods an Applications, 3rd Edition, John Wiley & Sons: New York.Ord, J.K. A.B Koehler and R.D, (1998).

[23] Nikolopoulos, K, Assimakopoulos, V., Bougioukos, N.Litsa, A & Petropoulos, F.The theta model: An essential forecasting tool for supply chain planning. In G.Lee (Ed). Lecture notes in electrical engineering: Vol.123. Advances in automation and robotic, Vol.2 (431-437). Berlin, Heidelberg:Springer-Verlag. (2011).

[24] Nikolopoulos., K., Thomakos, D., Petropoulos, F., Litsa, A. and Assimakopoulos. V. Forecasting S&P 500 with the theta model. International Journal of Financial Economics and Econometrics, 4, (2012). 73-78.
[25] Pai. P.F., Hong, W.C and Lin, C.S. Forecasting tourism demnd using a multifactor support vector machine model. In: Hao Y. et al (eds) Computational Intellegence and Security. CIS 2005. Lecture notes in computer Science, vol 3801. (2005) Sringer, Berlin. Heidelberg.
[26] Pedregal, D.J & Young, P.C. Modulated cycles, an approach to modelling periodic components from rapidly sampled data. International journal of Forecasting 22, (2006), 181-194.
[27] Petropoulus, F & Nikolopoulos, K.. Optimizing Theto model for monthly data. In proceedngs of the 5th International conference on agents and artificial intelligence, (2013).
[28] Shen, S., Li., G & Song, H. Combination forecast of international tourism demand.Annal of Tourism Research. 38(1), (2011). 72-89.
[29] Snyder, R, Discussion. International Journal of forecasting, 22(4), (2006), pp. 673-676.
[30] Song, H, Witt, S.F & Li, G. A Review of Recent research, Tourims Management, 29(2), (2008), 203-220.
[31] Taylor, J.W.. Short-term electricity demand forecasting using double seasonal exponential smoothing. Journal of the operational research society, 54, (2003).799-805.
[32] Taylor, J.W & Snyder, R.D Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing (2009). (9/09). URL: http://ideas.repec.org/p/msh/ebswps/2009-9.html.
[33] Thomakos, D., & Nikolopoulas, K. Fathoming the theta method for a unit root process. IMA. Journal of Management Mathematic, 25, (2014), 105-124.
[34] Waciko, K.J & Ismail, B, Forecasting International Tourism in Bali. Zenith International Journal of Multidisciplinary Research. Vol 8. (9). (2018). Pp.21-32.
[35] Zhang, G,. Time series forecasting using a hybrid ARIMA and Neural Network model. Neurocomputing. 50, (2003).pp: 159-175.
[36]