International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2020; 5(2): 01-08

www.mathsjournal.com
Received: 01-01-2020
Accepted: 03-02-2020

Kadek Jemmy Waciko
Research Scholar Department of Statistics, Mangalore University, Mangalagangothri, Karnataka, India

Ismail B
Professor of Statistics, YENEPOYA (Deemed to be University) Deralakatte, Mangalore, Karnataka, India

Corresponding Author:
Kadek Jemmy Waciko
Research Scholar Department of Statistics, Mangalore University, Mangalagangothri, Karnataka, India

SARIMA-ELM hybrid model versus SARIMA-MLP hybrid model

Kadek Jemmy Waciko and Ismail B

Abstract
Advanced quantitative models have been widely developed in academic literature, but practicians and statisticians still have little interest in the design and development of sophisticated hybrid models to produce more accurate predictions. In an empirical study, we compare performance forecast accuracy of the SARIMA-ELM hybrid model and the SARIMA-MLP hybrid model for forecasting tourist arrivals to Bali from 10 different countries such as China, Australia, Japan, India, USA, UK, Canada, South Africa, Malaysia, and Taiwan. Based on the RMSE and MAPE criteria, we found that our novelty model (the SARIMA-ELM hybrid model) performs better than the SARIMA-MLP hybrid model in the aspect of forecast accuracy.

Keywords: SARIMA-ELM hybrid model, SARIMA-MLP hybrid model, RMSE, MAPE

1. Introduction
Over the past few decades, time series models like ARIMA and SARIMA models (Box & Jenkins, 1970) [2] generally have been used especially for forecasting tourism demand. The selection of the forecasting model is the important criteria that will influence the forecasting accuracy (Pai *et al.*, 2005) [12]. Artificial Neural Networks (ANN) also has been developed in tourism forecasting (Song *et al.*, 2008) [13]. As an overview, we can see in (Burger *et al.*, 2001; Cho, 2003; Kon and Turner, 2005; Koutras, 2016) [3, 4, 9, 10]. Multi-Layer Perceptron-Neural Networks (MLP-NN) model have well performed for forecasting tourism demand problem, see in (Waciko & Ismail, 2018) [15] demonstrated that MLP-NN outperformed SARIMA and Extreme Learning Machine (ELM) models. The study of a hybrid model for forecasting seasonal time-series data that combines SARIMA and Back Propagation Neural Network (BPNN) models was used in (Tseng *et al*, 2002) [14]. The empirical analysis found that the SARIMA-BPNN hybrid model outperformed SARIMA, the BPNN with deseasonalized data, and BPNN with differenced data. The ARIMA-ANN hybrid model has been developed in Zhang (2003) [17], where ARIMA-ANN hybrid model outperformed ARIMA and ANN models. On the other hand, Chen *et al* (2007) [5] combined SARIMA and Support Vector Machine (SARIMA-SVM hybrid model), their model is more effective than SARIMA and SVM models. The Seasonal Support Vector Machine (SSVM) has been developed in Hua *et al* (2017) [7]. The empirical analysis found that the Seasonal Support Vector Machine (SSVM) superior to SARIMA and SARIMA-BPNN hybrid model. The recent study, SARIMA-ELM hybrid model has been developed by Waciko & Ismail (2018) [16] to forecasting tourist arrivals in Nepal. They found that the SARIMA-ELM hybrid model outperforms the Holt-Winter's, SARIMA, Extreme Learning Machine (ELM) and Multi-Layer Perceptron-Neural Networks models (MLP-NN) in term of forecast accuracy. Even though many approaches have been implemented to create forecasts for the tourism demand, however, there is a need to develop the method that will improve accuracy in forecasting. In this study, SARIMA-ELM hybrid model and SARIMA-MLP hybrid model are adopted and investigated. We compared the performance forecast accuracy of the SARIMA-ELM hybrid model and the SARIMA-MLP hybrid model and to find an appropriate hybrid model for forecasting tourist arrivals.

2. Material and Methods

2.1 The Framework of the Proposed Study

The framework of the proposed study as shown in Figure 1. Its details are given as follows:

1. From data tourist arrivals to Bali, construct the SARIMA fitted model and find the residuals from the SARIMA fitted model.
2. Construct a nonlinear model (ELM and MLP models) for residuals.
3. Combine forecasting for the SARIMA-ELM hybrid model and for the SARIMA-MLP hybrid model.
4. Make a comparison between the SARIMA-ELM hybrid model and the SARIMA-MLP hybrid model and select the best model that will improve accuracy in forecasting.

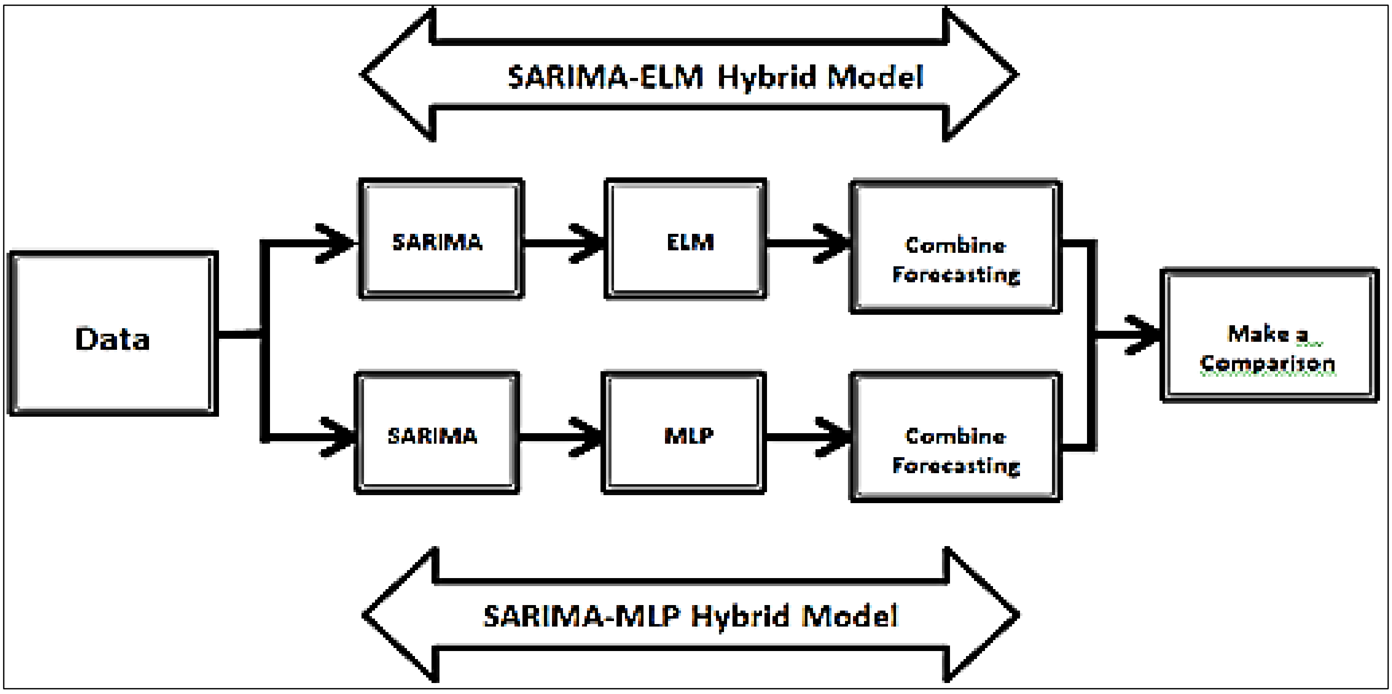


Fig 1: The Framework of the Proposed Study

A. Seasonal Autoregressive Moving Average (SARIMA)

SARIMA $(p,d,q,P,D,Q)_s$ can be written as

$$(B)\Phi_P(B)^S(1-B)^d(1-B^S)^D z_t = \theta_q(B)\Theta_Q(B^S)^S\varepsilon_t \quad (1)$$

$$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

$$\Phi_P(B)^S = 1 - \Phi_1(B)^S - \Phi_2(B)^{2S} - \cdots - \Phi_P(B)^{PS}$$

$$\Theta_Q(B)^S = 1 - \Theta_1(B)^S - \Theta_2(B)^{2S} - \cdots - \Theta_Q(B)^{QS}$$

SARIMA $(p,d,q,P,D,Q)_s$ has two parts: Non Seasonal part (p,d,q) and Seasonal part (P,D,Q) where p = order of non-seasonal AR term; d = order of non-seasonal differencing; q = order of non-seasonal MA term; P = order of seasonal AR term; D = order of seasonal differencing; Q = order of seasonal MA term; and s = seasonal period (Nachane, 2006) [11].

To find SARIMA fitted model we used The Box-Jenkins method (1970) and following 3 steps:

1. Identification (Identification of the time series model that summarizes the data in the best possible way).
2. Estimation (Estimation of the parameters of the model identified in the previous step).
3. Diagnostic Checking (Evaluation of the model for better predictions about future).

2.2 Multi-Layer Perceptron-Neural Networks (MLP-NN)

MLP-NN model with a single hidden layer (see figure 2) is represented as follows:

$$z_t = \beta_0 + \sum_{j=1}^{q} \beta_j f(\gamma_{0j} + \sum_{i=1}^{p} \gamma_{ij} z_{t-i}) + \varepsilon_t \quad (2)$$

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 (Haykin, 1999).

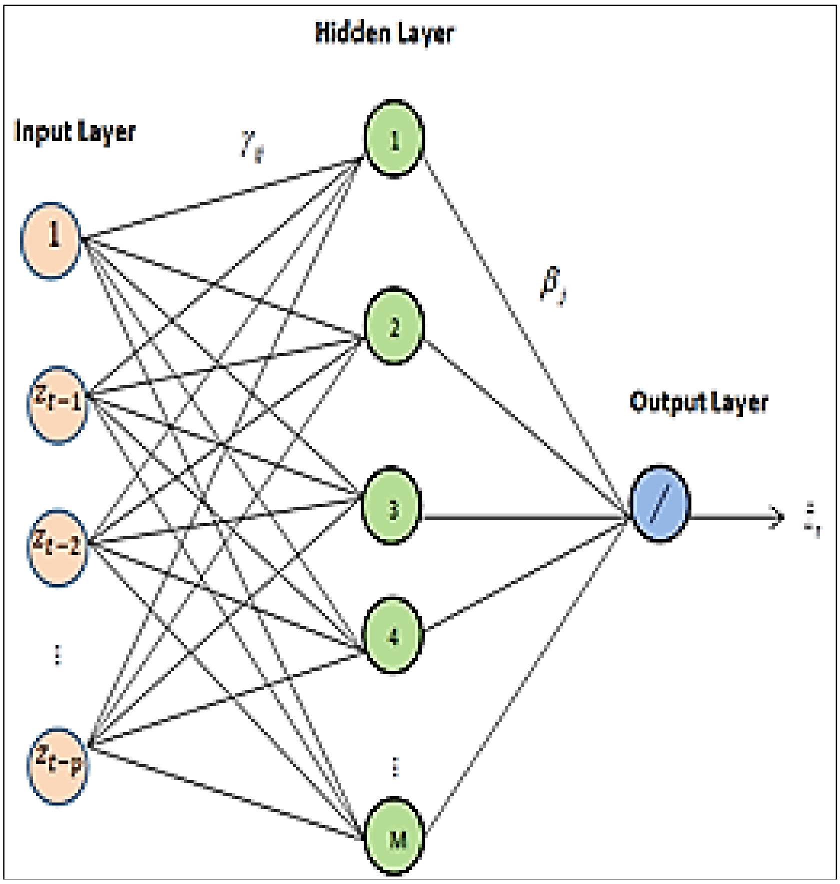


Fig 2: Architectures of MLP-NN with a single hidden layer

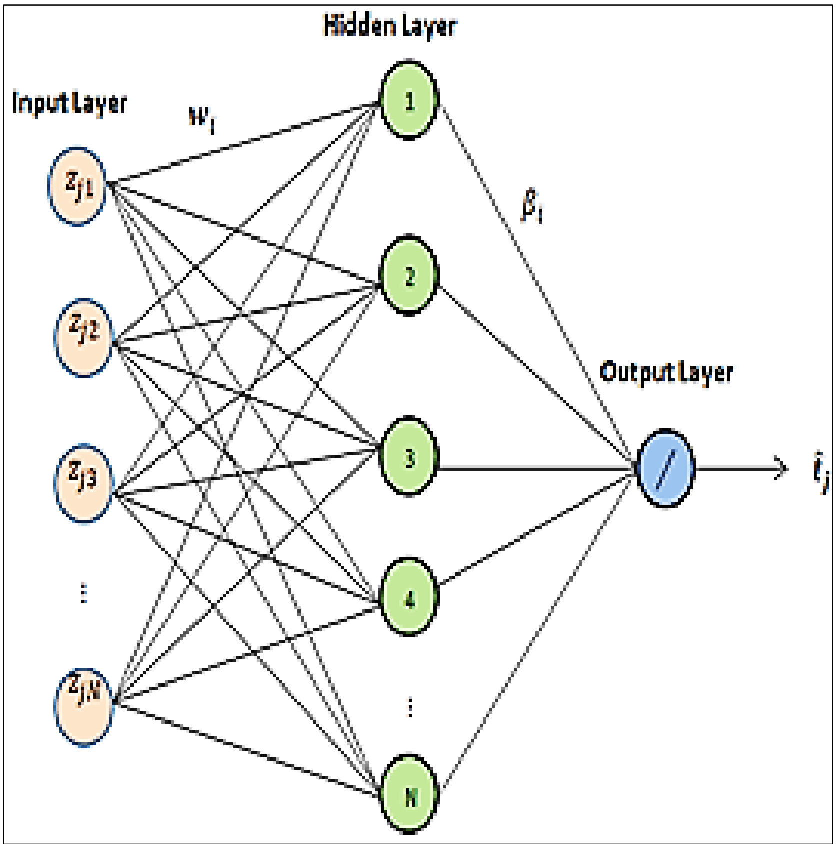


Fig 3: Architectures of ELM with a single hidden layer

2.2.1 Back Propagation Neural Network Algorithm

The Back Propagation method (Bishop, 2006) [1]. can be described as follows:

1. Apply an input vector X_n to the network and forward propagate through the network using $a_j = \sum_{i=1} \omega_{ji} z_i$ and $z_j = h(a_j)$ to find the activations of all the hidden and output units. Where, z_i is the activation of a unit, or input, that sends a connection to unit j, and ω_{ji} is the weight associated with that connection, the sum in $a_j = \sum_{i=1} \omega_{ji} z_i$ is transformed by a nonlinear activation function $h(.)$ to give the activation z_j of unit j.
2. Evaluate the δ_k for all the output units using $\delta_k = y_k - t_k$, where $\delta's$ are often referred to as errors for reasons we shall see shortly. The outputs y_k are a linear combination of the input variables x_i so that $y_j = \sum_{i=1} \omega_{ki} x_i$
3. Backpropagate the $\delta's$ using $\delta_j = h'(a_j) \sum_k \omega_{kj} \delta_k$ to obtain δ_j for each hidden unit in the network and use $\frac{\partial E_n}{\partial \omega_{ji}} = \delta_j z_i$ to evaluate the required derivatives.

2.3 Extreme Learning Machine (ELM)

ELM with a single hidden layer (see figure 3) was developed in (Huang *et al*., 2006) [8]. Assume a set of N different samples (z_i, t_i) with $z_i = [z_{i1}, z_{i2}, \ldots, z_{in}]^T \in R^N$, $t_i = [t_{i1}, t_{i2}, \ldots, t_{in}]^T \in R^m$ then ELM with a single hidden layer is represented as follows:

$$\sum_{\mathrm{i}=1}^{\tilde{\mathrm{N}}} \beta_{\mathrm{i}} f(w_i . z_j + b_i) = 0_j, \; j = 1, 2, \ldots, N \tag{3}$$

These N samples with zero error can be estimated with model (3)

Which means that $\sum_{j=1}^{\tilde{N}} \| o_j - t_j \| = 0,$ such that

$$\sum_{i=1}^{\tilde{N}} \beta_i f(w_i . z_j + b_i) = t_j, \quad j = 1, 2, .. N \tag{4}$$

Model (4) can be modified and written compactly as

$$H\beta = T \tag{5}$$

Where

$$H(w_i, \ldots, w_{\tilde{N}}, b_1, \ldots, b_{\tilde{N}}, z_1 \ldots, z_N)$$

$$= \begin{bmatrix} f(w_1 . z_1 + b_1) & \cdots & f(w_{\tilde{N}} . z_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ f(w_1 . z_N + b_1) & \cdots & f(w_{\tilde{N}} . z_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \tag{6}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \tag{7}$$

2.3.1 ELM Algorithm

Assumed a training set $\aleph = \{(z_i, t_i) \in R^n, t_i \in R^m, i = 1, 2, \ldots . N\}$ with activation function $f(x)$ and $\tilde{N}$ hidden node (Huang *et al*, 2006) [8]:

1. Randomly allocate input weight w_i and biases b_i, i = 1, 2, …,$\tilde{N}$
2. Calculate the hidden layer output matrix H.
3. Calculate the output of weight β, where $\beta = H^{\dagger} T$

2.4 SARIMA-ELM Hybrid Model Versus SARIMA-MLP Hybrid Model

Suppose the observations $\{Y_t, t = 1, 2, 3, \ldots\}$ has two parts, $\{A_t, t = 1, 2, 3, \ldots\}$ denotes the modeling of SARIMA (Linear model) and $\{B_t, t = 1, 2, 3, \ldots\}$ denotes the modeling of ELM (Nonlinear model) and MLP (Nonlinear model), as presented below:

$$Y_t = A_t + B_t \tag{8}$$

Forecasting steps of SARIMA-ELM hybrid model and SARIMA-MLP hybrid model as follows figure 1.

Forecasting steps

1. Get a new time series $\{Y_t, t = 1,2,3 \dots.\}$ after removing the seasonal of $\{Z_t, t = 1,2,3 \dots.\}$ using the seasonal index (S_k). Given a sequence of observations of time series with seasonal effects, $\{M_{ij}, i = 1,2,..L, j = 1,2,..s\}$ M_{ij} represents the ith year (period) and jth month (period) data. Subsequently, the average of months (period) data is shown as follows

$$\bar{m}_k = \frac{1}{L}\sum_{i=1}^{L} M_{ij}, (j = 1,2, \dots, s) \tag{9}$$

and the average for the whole period is

$$\bar{m} = \frac{1}{LS}\sum_{i=1}^{L}\sum_{j=1}^{S} m_{ij} \tag{10}$$

Thus, the seasonal index is

$$S_k = \frac{\bar{m}_k}{\bar{m}}, (k = 1,2, \dots., s) \tag{11}$$

$$Z_{ij} = \frac{m_{ij}}{S_j}, (i = 1,2, \dots, L, \text{and } j = 1,2, \dots, s) \tag{12}$$

Do forecasting using the linear time series model$\{Z_t, t = 1, 2,3, \dots\}$. The forecast result is $\hat{A}_t$ Then found the residual series $\{e_t\}$ by using $\{Z_t, t = 1, 2, \dots\}$ and $\hat{A}_t$, such that $e_t = X_t - \hat{A}_t$. From the original sequence, the series $\{e_t\}$ covers the nonlinear structure, represents that $e_t = f(e_{t-1}, e_{t-2}, \dots \dots e_{t-n}) + \varepsilon$, where f is activation function of ELM and MLP-NN.

Next step Forecasting for SARIMA -ELM

3a) Create ELM for the series $\{e_t\}$ and found the forecast result, namely $\widehat{e_t} = B_t$ by using (3) to (7) with respect to inputse_1, $e_2, \dots, e_n$ such that,

$$\sum_{i=1}^{\tilde{n}} \beta_i f\,(w_i . e_j + b_i = t_j, j = 1, \dots, n \tag{13}$$

Model (13) can be modified and written compactly as $H\beta = T$ where,

$H(w_i, \dots, w_{\tilde{n}}, b_1, \dots, b_{\tilde{n}}, e_1 \dots, e_n)$

$$= \begin{bmatrix} f(w_1 . e_1 + b_1) & \cdots & f(w_{\tilde{n}} . e_1 + b_{\tilde{n}}) \\ \vdots & \cdots & \vdots \\ f(w_1 . e_n + b_1) & \cdots & f(w_{\tilde{n}} . e_N + b_{\tilde{n}}) \end{bmatrix}_{n \times \tilde{n}} \tag{14}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{n}}^T \end{bmatrix}_{\tilde{n} \times m} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_n^T \end{bmatrix}_{n \times m}$$

Next Step Forecasting for SARIMA-MLP

3b). Create Back Propagation Neural Network (Bishop, 2006) [1] for the series $\{e_t\}$ and found the predicted result of Nonlinear time series, namely $\hat{e}_t = B_t$.

Final Step Forecasting for SARIMA-ELM and the SARIMA-MLP Hybrid models

1) Hybridizing the forecast result of the linear time series, with the forecast result of ELM or MLP, such that $\widehat{Z}_t = \widehat{A_t} + \hat{e}_t$. Finally, we can found $\hat{Y}_t$ (The final forecasted result of SARIMA-ELM and SARIMA-MLP hybrid models) by multiplyingS_k.

2.5 Measuring Forecast Accuracy

The Root Mean Squared Error (RMSE) and The Mean Absolute Percentage Error (MAPE) are used to measuring forecast accuracy as presented below:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(z_t - \hat{z}_t)^2} \tag{15}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{z_t - \hat{z}_t}{z_t}\right| \times 100 \tag{16}$$

Where z_t= the actual value, $\hat{z}_t$= the forecasted value, and n = the size of test set.

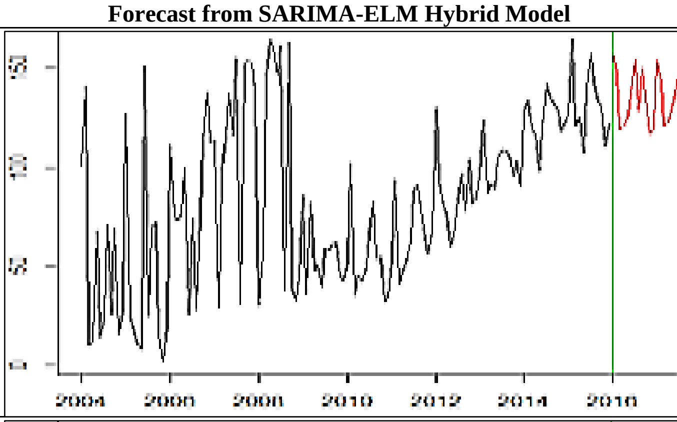
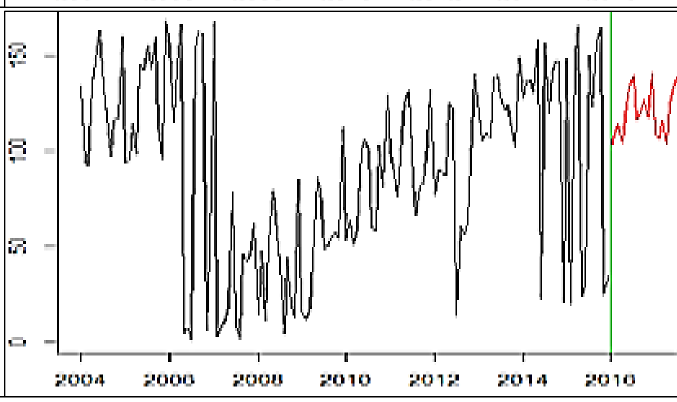
3. Results and Discussion

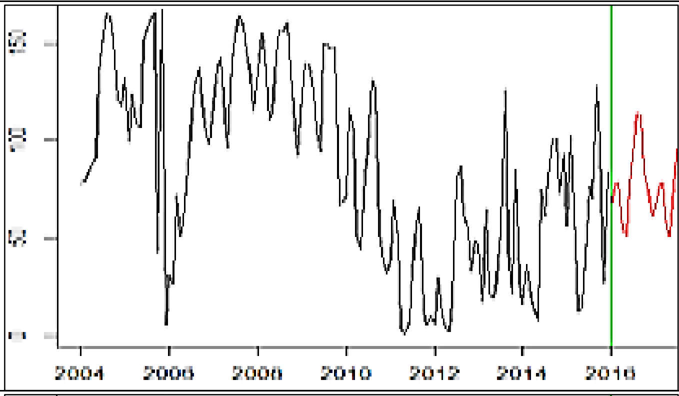
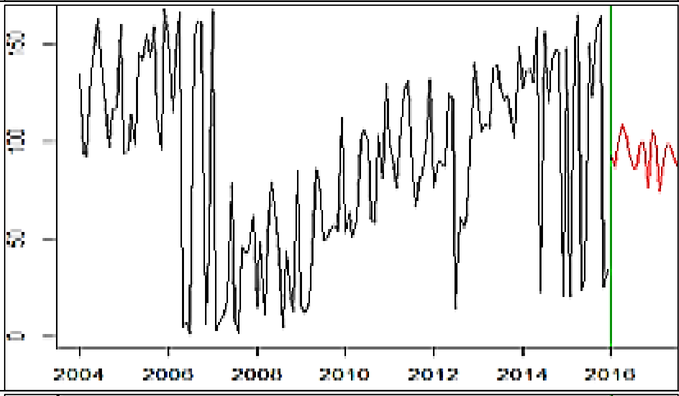
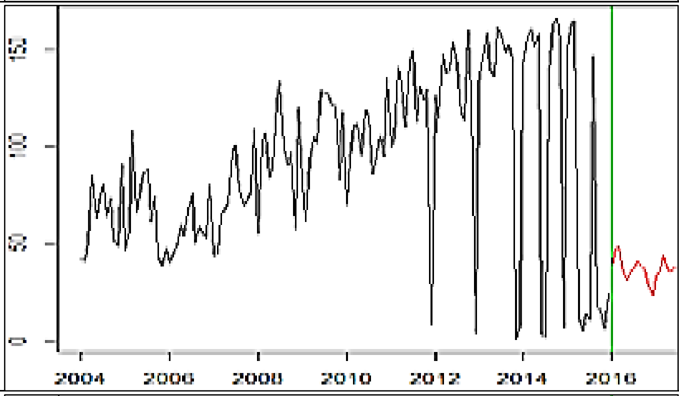
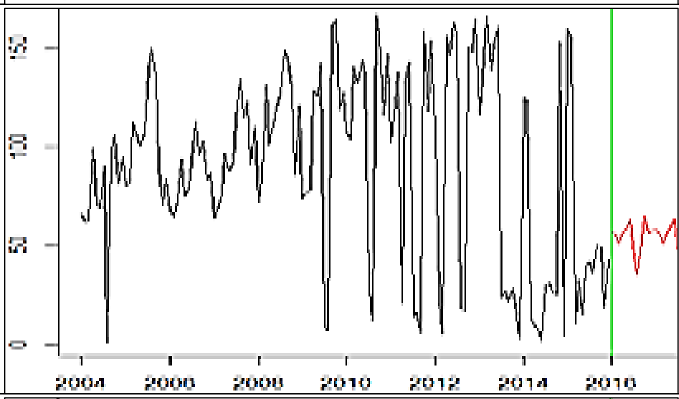
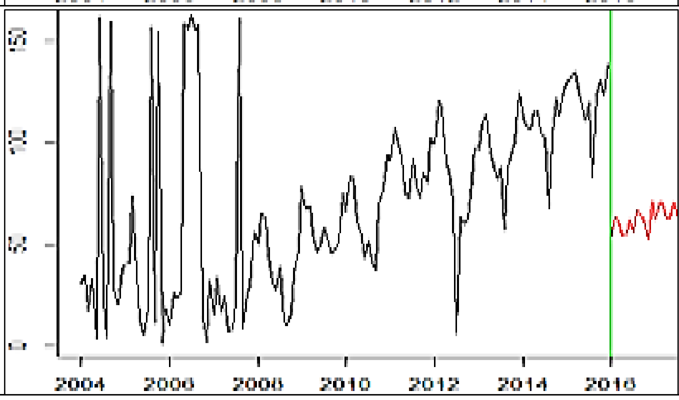
Data of international tourist arrival to Bali from 10 different countries (China, Australia, Japan, India, USA, UK, Canada, South Africa, Malaysia, and Taiwan) have been collected online from Bali Government Tourism Office (www.disparda.baliprov.go.id). In order to compare the forecast performance for the SARIMA-ELM hybrid model and the SARIMA-MLP hybrid model, we consider to used data set from January 2004 to December 2017 where data training (January 2004-December 2015) and data testing (January 2016-December 2017). R Studio 3.5.3 software is used for all processes in data analysis.

Table 1: Measure of forecasting accuracy

No	Country	Model	RMSE	MAPE
1.	China	SARIMA-ELM	94.996	1343.391
		SARIMA-MLP	96.689	1369.658
2.	Australia	SARIMA-ELM	77.303	1273.512
		SARIMA-MLP	77.326	1274.148
3.	Japan	SARIMA-ELM	28.032	72.703
		SARIMA-MLP	28.409	75.533
4.	India	SARIMA-ELM	46.004	116.661
		SARIMA-MLP	49.724	126.508
5.	USA	SARIMA-ELM	14.257	59.295
		SARIMA-MLP	18.737	81.788
6.	UK	SARIMA-ELM	15.216	30.760
		SARIMA-MLP	18.254	38.122
7.	Canada	SARIMA-ELM	78.087	53.659
		SARIMA-MLP	81.143	56.214
8.	South Africa	SARIMA-ELM	61.543	290.335
		SARIMA-MLP	66.560	311.011
9.	Malaysia	SARIMA-ELM	42.295	154.100
		SARIMA-MLP	42.414	154.461
10.	Taiwan	SARIMA-ELM	42.559	345.509
		SARIMA-MLP	42.682	346.460

as shown in table 1. Where all the measures of RMSE and MAPE are lowest for the SARIMA-ELM hybrid model. Based on table (1), the experimental results indicated that the SARIMA-ELM hybrid model is superior to the SARIMA-MLP hybrid model in terms of forecast accuracy. This result also confirms and complements the results of previous studies see (Waciko & Ismail, 2018) [15] related to the SARIMA-ELM hybrid model forecast performance. The forecast from the SARIMA-ELM hybrid model for international tourist arrival to Bali from 10 different countries as shown in figure 4.

Country	Forecast from SARIMA-ELM Hybrid Model
China	
Australia	

Japan	
India	
USA	
UK	
Canada	

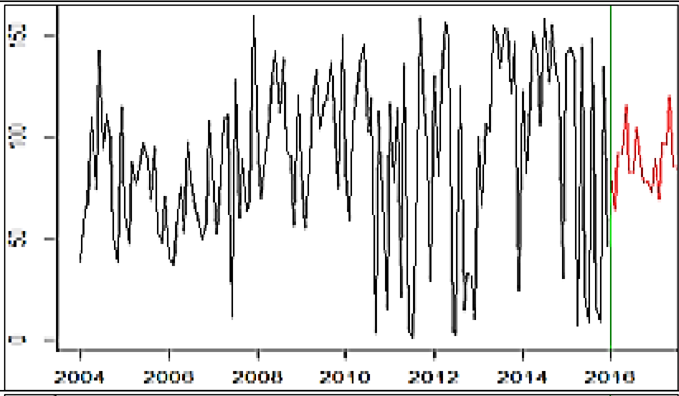
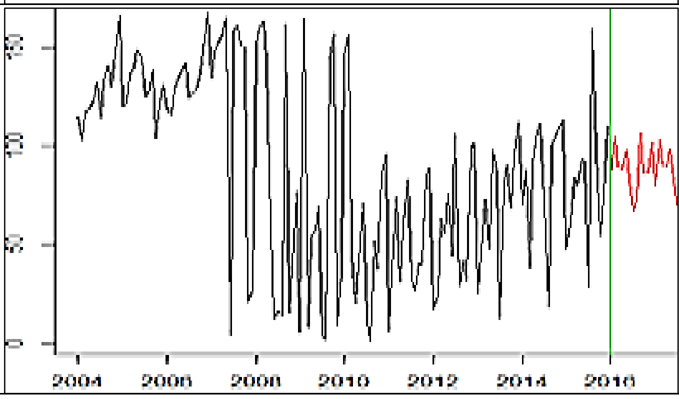
South Africa	
Malaysia	

Fig 4: Forecast Tourist Arrival in Bali from 10 Country using SARIMA-ELM Hybrid Model

Conclusions

The SARIMA-ELM hybrid model was compared with another hybrid model namely the SARIMA-Multi Layer Perceptron Neural Network (SARIMA-MLP) hybrid model. The performance forecast accuracy of these two hybrid models is compared for forecasting tourist arrival to Bali from 10 different countries such as China, Australia, Japan, India, USA, UK, Canada, South Africa, Malaysia, and Taiwan. To choose an appropriate forecasting model, we used the Root Mean Squared Error (RMSE) and the Mean Absolute Percentage Error (MAPE) criteria. Based on RMSE and MAPE criteria, this study clarifies that our novelty the SARIMA-ELM hybrid model outperforms the SARIMA-MLP hybrid model in the aspect of forecast accuracy. However, future research is open to study the comparison performance forecast accuracy of the SARIMA-ELM hybrid model with another hybrid model.

References

1. Bishop CM. Pattern Recognition and Machine learning. New York: Springer Science +Business Media, LLC, 2006.
2. Box GEP, Jenkins GM. Time Series Analysis: Forecasting and Control. San Fransisco: Holden-Day, Revised Edn, 1970.
3. Burger CJSC, Dohnal M, Kathrada M, Law R. A practitioner’s guide to time series methods for tourism demand forecasting a case study of Durban, South Africa. Tourism Management. 2001; 22:403-409.
4. Cho V. A comparison of three different approaches to tourist arrival forecasting. Tourism Management. 2003; 24:323-330.
5. Chen KY, Wang CH. A Hybrid SARIMA and Support Vector Machines in Forecasting the Production Values of the Machinery industry in Taiwan. Expert Systems with Applications. 2007; 32(1):254-264.
6. Haykin S. Neural Networks: A Comprehensive Foundation. Prentice-Hall, New Jersey, 1999.
7. Hua L, Xing L, Shuang W. Based on SARIMA-BP Hybrid model and SSVM model of international crude oil price prediction research. Anziam J. 2017; 58(E):E143-E161.
8. Huang GB, Zhu QY, Siew CK. Extreme Learning Machine: Theory and Applications. Neurocomputing. 2006; 70:489-501.
9. Kon SC, Turner WL. Neural Network Forecasting of Tourism Demand. Tourism Economics. 2005; 11:301-328.
10. Koutras A, Panagopoulos A, Ioannis AN. Forecasting Tourism Demand Using Linear and Nonlinear Prediction Models. Academica Turistica, Year. 2016; 9(1):85-98.
11. Nachane DM. Econometrics: Theoretical Foundations and Empirical Perspectives. Oxford University Press, India, 2006.
12. Pai PF, Hong WC, Lin CS. Forecasting Tourism Demand Using a Multifactor Support Vector Machine Model. In: Hao Y *et al*. (eds) Computational Intelligence and Security. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2005, 3801.
13. Song H, Witt SF, Li G. A Review of Recent Research, Tourism Management. 2008; 29:2003-220.
14. Tseng F, Yu HC, Tzeng GH. Combining Neural Network model with seasonal time series ARIMA model. Technological Forecasting & Social Change. 2002; 66(1):71-87.
15. Waciko KJ, Ismail B. SARIMA-ELM Hybrid Model for Forecasting Tourist in Nepal. RESEARCH REVIEW International Journal of Multidisciplinary. 2018; 3(7):343-349.
16. Waciko KJ, Ismail B. Forecasting International Tourism in Bali. ZENITH International Journal of Multidisciplinary Research. 2018; 8(9):21-32.
17. Zhang G. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model, Neurocomputing. 2003; 50:159-175.