RESEARCH REVIEW JOURNALS
Volume-03
Issue-07
July-2018

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com [UGC Listed Journal]

SARIMA-ELM Hybrid Model for Forecasting Tourist in Nepal

*[1]Kadek Jemmy Waciko & [2]Ismail B

*[1]*Research Scholar, Department of Statistics, Mangalore University, Karnataka (India)*
[2]*Professor, Department of Statistics, Mangalore University, Karnataka (India)*

ARTICLE DETAILS

Article History
Published Online: 05 July 2018

Keywords
Extreme Learning Machine (ELM), SARIMA, SARIMA-ELM hybrid model

***Corresponding Author**
Email: kadekjemmywaciko[at]gmail.com

ABSTRACT

In this study a novel hybrid model has been developed to forecasting tourist arrivals. The main concept is to combine two different forecasting techniques such as SARIMA and Extreme Learning Machine models to produce a new SARIMA-ELM hybrid Model, so as to achieve accuracy in forecasting. Forecasting accuracy for SARIMA, Triple Exponential Smoothing (The Holt-Winter's), Multi Layer Perceptron-Neural Networks (MLP-NN), Extreme Learning Machine (ELM) and SARIMA-ELM hybrid models are computed and compared using criteria like RMSE, MAE, and MAPE. Empirical analysis found that SARIMA-ELM hybrid model has highest forecasting accuracy Thus, SARIMA-ELM hybrid model is the most appropriate model for forecasting tourist arrivals.

1. Introduction

ARIMA and SARIMA proposed by Box and Jenkins (1970) very popular models for forecasting tourism demand as an overview we can see in Loganathan et al (2010), Singh (2013), Baldigara & Mamula (2015) and Petrevska (2017). Nor et al (2016), forecasting tourism demand in Malaysia using SARIMA, Regression and Triple Exponential Smoothing (The Holt-Winter's) models, it was found that Triple Exponential Smoothing (The Holt-Winter's) model gave the most accurate forecast in term of error magnitude.

Some conventional statistical methods such as Triple Exponential Smoothing and SARIMA models are extensively used in building prediction models to predict time series data especially in tourism, but there are many practical situations where even the the SARIMA and Triple Exponential Smoothing models did not produce effective results. Multi Layer Perceptron-Neural Networks (MLP-NN) have emerged in solving the tourism forecasting problem (Koutras et al, 2016). To Forecasting International tourist demand in Catalonia, Claveria et al (2015) used Neural Network Models. The results showed that Multi Layer Perceptron-Neural Networks (MLP-NN) models outperform Elman networks.

The current study of hybrid model for forecasting time series data which combines ARIMA and ANN for forecasting data time series was discussed in Zang (2003). Unlike other researchers, Samsudin et al (2010) combine Group Method of Data Handling (GMDH) and Least Square Support Vector Machine (LSSVM) to forecasting tourism demand in Johor Malaysia. Then the results show that GMDH-LSSVM hybrid model is more effective than GMDH and LSSVM models. SARIMA-Back Propagation hybrid model was used by Tseng et al (2002) to predict seasonal time series data, the results showed that SARIMA- Back Propagation hybrid model is superior to the SARIMA model, the Back Propagation with deseasonalized data, and the Back Propagation with differenced data. Chen and Wang (2007) was used SARIMA-SVM hybrid model for forecasting the seasonal time series data of Taiwan's machinery industry production values. Empirical analysis shows that SARIMA-SVM hybrid model performs better than SARIMA and SVM models. Hua et al (2017) combine SARIMA and Back Propagation Neural Network (BPNN) and they also combine Seasonal and Support Vector Machine (SVM). Empirical analysis shows that the SSVM hybrid model has highest prediction accuracy.

The aim of this study is to propose a novel hybrid approach which adopted SARIMA and Extreme Learning Machine (ELM) models for forecasting tourist arrivals.

2. Methodology

2.1 Seasonal Autoregressive Moving Average (SARIMA)

SARIMA model is used to forecasting tourism demand. Box and Jenkins (1970) generalized the ARIMA model to deal with seasonality, and defined a general multiplicative seasonal ARIMA (SARIMA) model as

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

Where $\emptyset_p, \Phi_P, \theta_q, \Theta_Q$ are polynomial of order p,P,q,Q respectively, S is the seasonal length, B denotes the backward shift operator, ε_t is the estimated residual at time t, d is the number of regular differences, D is the number of seasonal differences, Y_t denotes the observed value at time t, t = 1,2,…,k , $\{\varepsilon(t)\}$ is a White Noise (WN) sequence, $\{\varepsilon(t)\}$ are independently and identically distributed with $E(\varepsilon(t)) = 0$ and Var $(\varepsilon(t)) = \sigma_\varepsilon^2$ (a constant).

$$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$$
$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$
$$\Phi_P(B)^S = 1 - \Phi_1(B)^S - \Phi_2(B)^{2S} - \cdots - \Phi_P(B)^{PS}$$
$$\Theta_Q(B)^S = 1 - \Theta_1(B)^S - \Theta_2(B)^{2S} - \cdots - \Theta_Q(B)^{QS}$$

The above model is called a ARIMA (p,d,q) x (P,D,Q)s or SARIMA model (Nachane, 2006).

2.2. Triple Exponential Smoothing (The Holt-Winter's)

Triple Exponential Smoothing (The Holt-Winter's) model can be used for series including both trend and seasonal components. Triple Exponential Smoothing model is used to forecasting tourism demand. The equations associated with each of these elements are as follows (Frechtling, 2001)

Level: $$L_t = \alpha \frac{A_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1}) \quad (3)$$

Trend: $$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \quad (4)$$

Seasonal: $$S_t = \gamma \frac{A_t}{L_t} + (1-\gamma)S_{t-s} \quad (5)$$

Forecast: $$F_{t+h} = (L_t + hb_t)S_{t-s+h} \quad (6)$$

Where: $L =$ Level of the series; α = Level smoothing constant between 0 and 1; A = actual value; s = Number of seasonal periods in a year; $b =$ Trend of the series; $\beta =$ Seasonal smoothing constant between 0 and 1; S = Seasonal component; γ = Seasonal smoothing constant between 0 and 1; t = Some time period and h = Number of time periods ahead to be forecast.

2.3. Multi Layer Perceptron-Neural Networks (MLP-NN)

Multi Layer Perceptron-Neural Network (MLP-NN) is one of the most popular Artificial Neural Network (ANN) model for forecasting tourism demand. MLP-NN model with a single hidden layer can written following mathematical expression (Haykin, 1999) :

$$z_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j\, g\left(\beta_{oj} + \sum_{i=1}^{p} \beta_{ij} z_{t-i}\right) + \varepsilon_t \quad (7)$$

Where, $z_{t-i}(i = 1,2,\dots,p)$ are the p inputs and $\hat{z}_t$ is the output. The integers p, q are the number of input and hidden nodes respectively. $\alpha_j(j = 0,1,2,\dots,q)$ and $\beta_{ij}(i = 0,1,2,\dots,p; j = 0,1,2,\dots q)$ are the conection weights and ε_t; g is a sigmoid transfer function such as the logistic: $g(x) = \frac{1}{1+e^{-x}}$ is applied as the nonlinier activation function; α_0 and β_{oj} are the bias terms.

2.4. Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) was proposed by Huang (Huang et al. 2006), Figure (1). shows a standar single hidden layer feed forward neural network (SLFN)

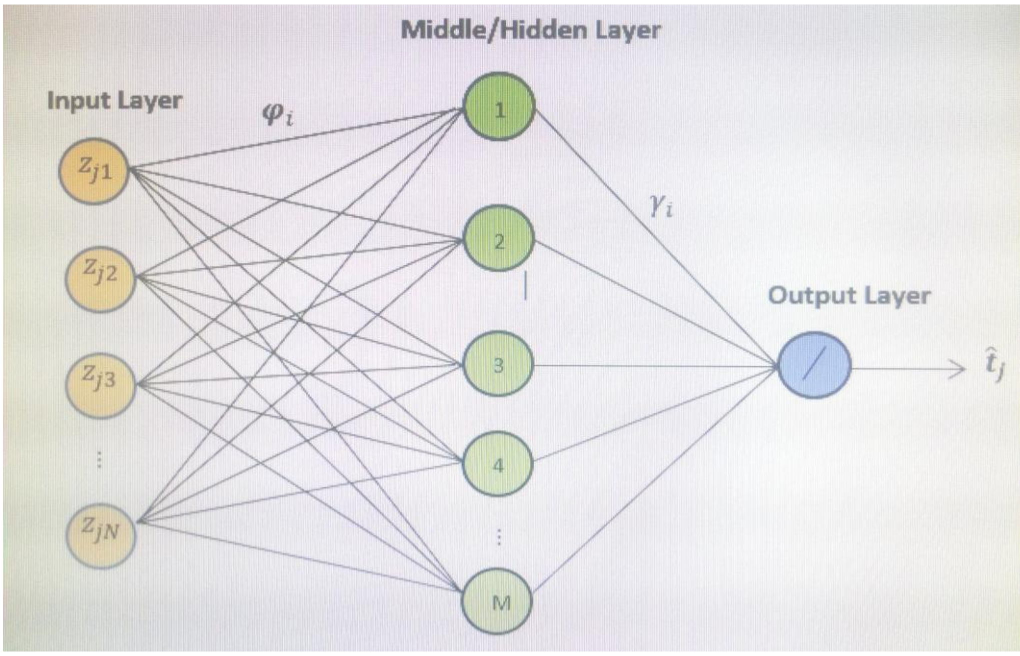


Fig. (1). a standar single hidden layer feed forward neural network (SLFN) in ELM

Consider a set N arbitrary distinct samples (z_i, t_i); $1 \le i \le N$, with $z_i = [z_{i1}, z_{i2}, \dots, z_{in}]^T \in \boldsymbol{R}^N$, $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \boldsymbol{R}^m$, activation function $f(z)$ and $\tilde{N}$ hidden nodes , a standar single hidden layer feed forward neural network (SLFN) in Extreme Learning Machine (ELM) can written following mathematical expression:

$$\sum_{i=1}^{\tilde{N}} \gamma_i f_i(\boldsymbol{z}_j) = \sum_{i=1}^{\tilde{N}} \gamma_i f_i(\boldsymbol{\varphi}_i \cdot \boldsymbol{z}_j + c_i) = \boldsymbol{0}_j, \quad j = 1,\dots,N \quad (8)$$

Where, $\boldsymbol{\varphi}_i \cdot \boldsymbol{z}_j$ represents the inner product of $\boldsymbol{\varphi}_i$ and $\boldsymbol{z}_j$; c_i is the threshold of the ith hidden node. $\boldsymbol{\varphi}_i = [\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{in}]^T$ is the weight vector relating the ith hidden node and the input nodes;

$\gamma_i = [\gamma_{i1}, \gamma_{i2}, \cdots, \gamma_{im}]^T$ is the weight vector relating the ith hidden node and the output nodes. Formula (8) can approximate these N samples with zero error means that $\sum_{j=1}^{\tilde{N}} \|o_j - t_j\|$=0, such that,

$$\sum_{i=1}^{\tilde{N}} \gamma_i f_i (\varphi_i . z_j + c_i) = t_j, \quad j = 1, \ldots, N \tag{9}$$

The above N equations can be written briefly as

$$H\gamma = T \tag{10}$$

Where

$$H(\varphi_i, \ldots, \varphi_{\tilde{N}}, c_1, \ldots, c_{\tilde{N}}, z_1 \ldots, z_N) = \begin{bmatrix} f(\varphi_1 . z_1 + c_1) & \cdots & f(\varphi_{\tilde{N}} . z_1 + c_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ f(\varphi_1 . z_N + c_1) & \cdots & f(\varphi_{\tilde{N}} . z_N + c_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \tag{11}$$

$$\gamma = \begin{bmatrix} \gamma_1^T \\ \vdots \\ \gamma_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad \text{and} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \tag{12}$$

1). Algorithm ELM (Huang et al, 2006)

Assumed a traning set $\aleph = \{(z_i, t_i) \in R^n, t_i \in R^m, 1 \le i \le N\}$, activation function $f(x)$ and $\tilde{N}$ hidden node

Stage 1: Randomly assign input weight φ_i and bias c_i , i = 1,2,…,$\tilde{N}$

Stage 2: Compute the hidden layer output matrix H.

Stage 3: Compute the output weight γ

$$\gamma = H^{\dagger} T \tag{13}$$

Where $T = [t_1, \ldots, t_N]^T$, H is named the hidden layer output matrix of the neural network, the ith column of H is the ith hidden node output with respect to inputs $z_1, z_2, \cdots, z_N$ and $H^{\dagger}$ is represents the Moore–Penrose generalised inverse of matrix H.

2.5. SARIMA-ELM Hybrid Model

In this study SARIMA-ELM Hybrid Model has been proposed that combine linear and non-linear model structure, the observations $\{Y_t, t = 1,2,3, \ldots\}$ has two parts. The first part is linear structure $\{P_t, t = 1,2,3, \ldots\}$ and the second part is Non linear structure $\{W_t, t = 1,2,3, \ldots\}$ it has the representation

$$Y_t = P_t + W_t \tag{14}$$

The part P_t is represents the modeling of SARIMA and the other part W_t is represents the modeling of Extreme Learning Machine (ELM). Basically, the modelling step of SARIMA-ELM hybrid model followed and adopted from Zang (2003) and Hua et al (2017). Some modification have been created especially in formula (9) , the procedure of SARIMA-ELM hybrid model as follows:

Eliminating the seasonal of $\{Y_t, t = 1,2,3, \ldots\}$ using The Seasonal Indeks (S_k) and get linear time series model $\{Z_t, t = 1,2,3, \ldots\}$. Suppose an observation sequence of time series with seasonal effect is $\{M_{ij}, i = 1,2,3 \ldots, L, j = 1,2,3, \ldots, s\}$, here, M_{ij} expresses the ith year (period) and jth month (period) data. Then, the average of months (period) data is

$$\bar{m}_k = \frac{1}{L} \sum_{i=1}^{L} M_{ij}, (j = 1,2, \ldots, s) \tag{15}$$

and the average for the whole period is

$$\bar{m} = \frac{1}{Ls} \sum_{i=1}^{L} \sum_{j=1}^{s} m_{ij} \tag{16}$$

Thus, the seasonal index is

$$S_k = \frac{\bar{m}_k}{\bar{m}}, (k = 1,2, \ldots, s) \tag{17}$$

$$Z_{ij} = \frac{m_{ij}}{S_j}, (i = 1,2, \ldots, L, \text{ and } j = 1,2, \ldots, s) \tag{18}$$

After eliminating the seasonality, we construct and use the linear time series model $\{Z_t, t = 1,2,3, \ldots\}$ and its forecasting. Let the forecasted result by $\hat{P}_t$. We find that the residual series $\{e_t\}$ based on $\{Z_t, t = 1,2,3, \ldots\}$ and $\hat{P}_t$, such that $e_t = Z_t - \hat{P}_t$.

According to Zhang (2003) and Hua et al (2017), the series $\{e_t\}$ contains the Non Linier Structure in the original Sequence, it means that $e_t = f(e_{t-1}, e_{t-2}, \ldots \ldots e_{t-n}) + \varepsilon$, where f is a nonlinier function and ε is a random error term. We construct Extreme Learning Machine (ELM) for the series $\{e_t\}$ which is used in formula (9) to (13) with respect to inputs e_1, $e_2, \ldots, e_n$ where,

$$\sum_{i=1}^{\hat{n}} \gamma_i f_i (\varphi_i . e_j + c_i) = t_j, \quad j = 1, \ldots, n \tag{19}$$

The above n equations can be written compactly as

$H\gamma = T$ where,

$$H(\varphi_i, \ldots, \varphi_{\tilde{n}}, c_1, \ldots, c_{\tilde{n}}, e_1 \ldots, e_n) = \begin{bmatrix} f(\varphi_1 . e_1 + c_1) & \cdots & f(\varphi_{\tilde{n}} . e_1 + c_{\tilde{n}}) \\ \vdots & \cdots & \vdots \\ f(\varphi_1 . e_n + c_1) & \cdots & f(\varphi_{\tilde{n}} . e_N + c_{\tilde{n}}) \end{bmatrix}_{n \times \tilde{n}} \tag{20}$$

$$\gamma = \begin{bmatrix} \gamma_1^T \\ \vdots \\ \gamma_{\tilde{n}}^T \end{bmatrix}_{\tilde{n} \times m} \quad \text{and} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_n^T \end{bmatrix}_{n \times m}$$

After construct ELM for the series $\{e_t\}$, The next step is forecast the sequence from nonlinear model and the forecasted result is $\hat{e}_t = W_t$.

Hybridizing the forecasted result from linear time series model with the forecasted result from Extreme Learning Machine (ELM) model for residuals such that $\hat{Z}_t = \hat{P}_t + \hat{e}_t$, after that, we get the final forecasted result $\hat{Y}_t$ by multiplying S_k.

We describe an algorithm in SARIMA - ELM hybrid model with flow chart in Figure (2)

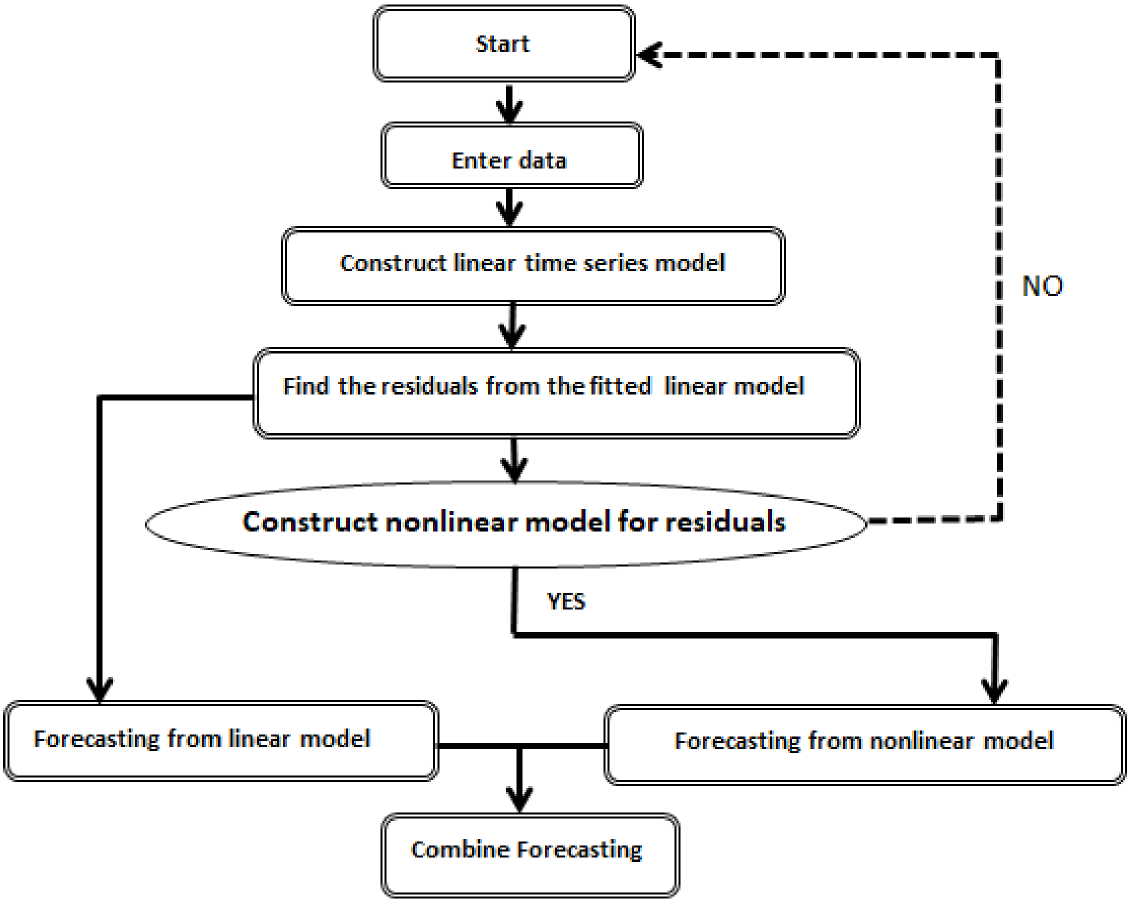


Fig. (2). An algorithm flow chart of SARIMA-ELM hybrid model

2.6. Measures of Forecasting Accuracy

We used Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Scaled Error (MASE) for measures of forecasting accuracy

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \tag{21}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t| \tag{22}$$

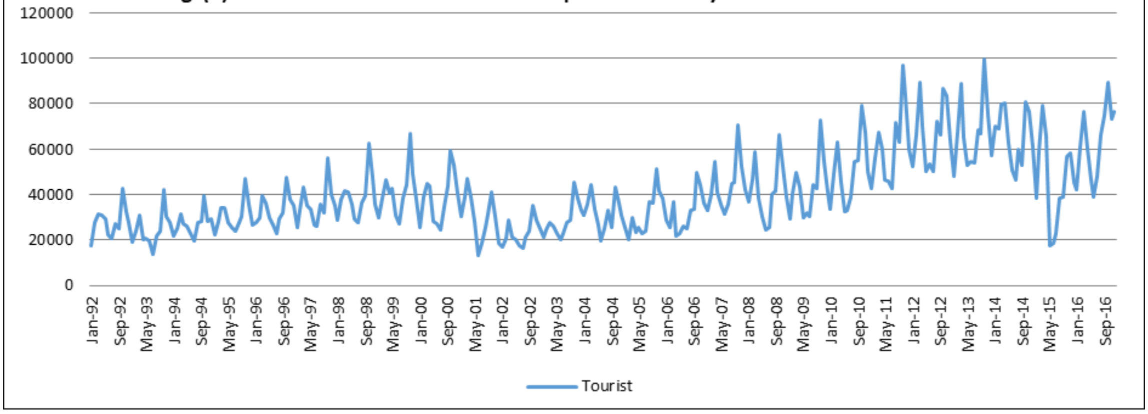
$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right| \times 100 \tag{23}$$

Where $e_t = y_t - \hat{y}_t$ is the forecast error, y_t is the actual value, $\hat{y}_t$ is the forecast value and n is the length of the forecasting horizon/the size of test set. Also $\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$ is the test mean and $\sigma^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t - \bar{y})^2$ is the test variance.

3. Result

The data set used in the empirical experiment is collected from Ministry of Culture, Tourism and Civil Aviation of Nepal (www.tourism.gov.np). Data of International Tourist arrival in Nepal from January 1992 to December 2016 is used to study as show in figure (3). We used data training from January 1992 to December 2011 and data testing is used from January 2012 to December 2016. For time series data analysis we used R software.

Fig. (3). International tourist arrivals in Nepal from January 1992 to December 2016



(Source: www.tourims.gov.np)

Models we have fitted to International tourist arrivals in Nepal series are SARIMA, Triple Exponential Smoothing (The Holt-Winter's), MLP-NN, ELM and SARIMA-ELM Hibrid models.

Showing in figure (4) to (8), Data training from January 1992 to December 2011 and forecasted data points from January 2012 to December 2016 .

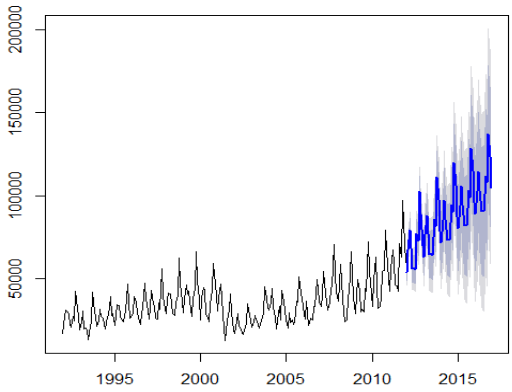


Fig. (4) Forecast from SARIMA

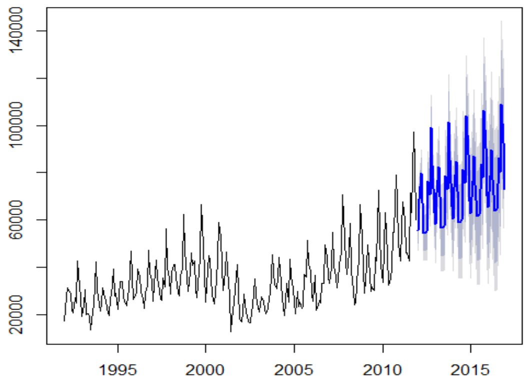


Fig. (5) Forecast from Triple Exponential Smoothing

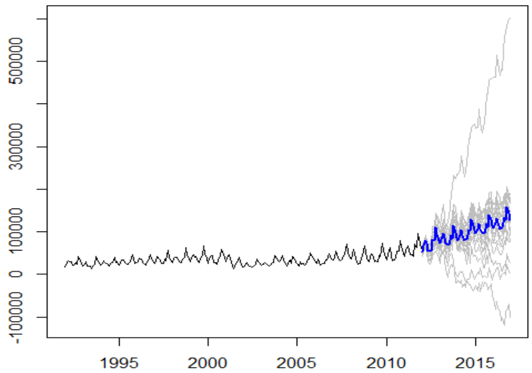


Fig. (6) Forecast from MLP

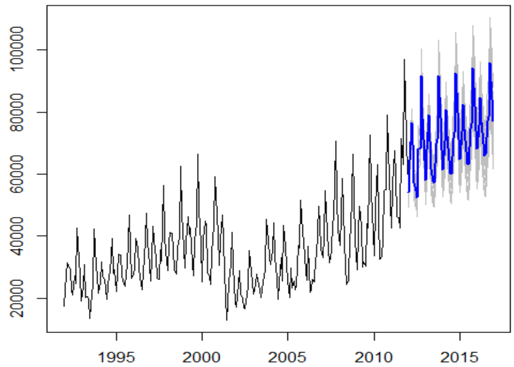


Fig. (7) Forecast from ELM

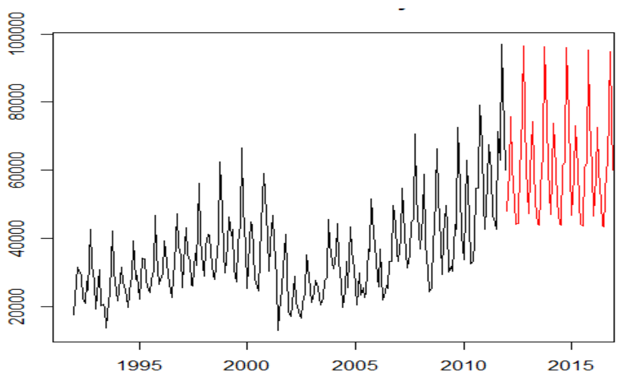


Fig. (8) Forecast from SARIMA-ELM Hybrid

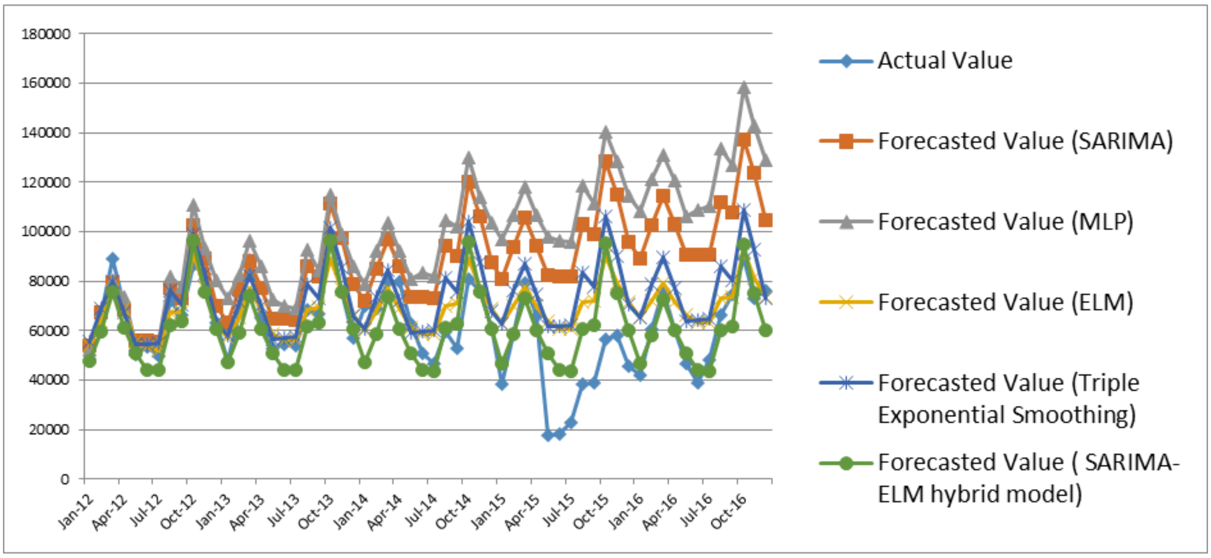


Fig. (9). Comparison between actual value and forecasted value of all models

In Figure (9) shows the comparison between actual value and forecasted value of all models (SARIMA, Triple Exponential Smoothing (The Holt-Winter's), MLP-NN, ELM and SARIMA-ELM Hybrid models), it appears that the closest to the actual value is the forecasted value of SARIMA-ELM Hybrid model. To judge forecast performances of different models, the measures MAE, MSE, RMSE and MAPE are considered. For each dataset, the results are presented in Table (1). SARIMA-ELM hibrid model is the best model for forecasting tourist arrivals as all the measure of accuracy is minimum (RMSE=12127.75; MAE=8975.69 and MAPE = 19.83) for this model.

Table 1. Forecast results for international tourist arrival in Nepal

Models	RMSE	MAE	MAPE
SARIMA	33748.96	27232.75	59.22
Triple Exponential Smoothing (The Holt-Winter's)	18646.35	13909.90	32.57
MLP-NN	45136.81	37415.66	79.27
ELM	17211.85	12208.50	30.70
Sarima-ELM Hybrid	12127.75	8975.69	19.83

Source : Data Processed

4. Conclusions

In this paper we discuss the application of SARIMA-ELM hybrid model for forecasting tourist arrival. SARIMA, Triple Exponential Smoothing (The Holt-Winter's), Multi Layer Perceptron-Neural Networks (MLP-NN), Extreme Learning Machine (ELM) and SARIMA-ELM hybrid models compared using criteria like RMSE, MAE and MAPE. It was found that the most appropriate model for forecasting tourist arrivals is SARIMA-ELM hybrid Model. The future of this research maybe to concentrate on study the statistical properties of proposed model.

References

1. Baldigara, T and Mamula, M. (2015). "Modelling International Tourims demand Using Seasonal ARIMA Models". *Journal of Hospitality and Tourism Management*, 21(1) 19-31.
2. Box, G.E.P. and Jenkins, G.M. (1970). *"Time Series Analysis: Forecasting and Control".* San Fransisco: Holden-Day, Revised edn.
3. Chen, K.Y and Wang, C.H.(2007). "A hybrid SARIMA and Support Vector Machines in Forecasting the Production Values of the Machinery industry in Taiwan", *Expert Systems With Applications*. 32(1), 254-264.
4. Claveria, O., Monte, E., and Torra, S.(2015) "Tourism demand Forecasting with Neural Network Models : Different Ways of Treating Information".*International Journal of Tourism research, Int.J.Tourism Res*, 17: 492-500.
5. Frechtling, D.C.(2001)."*Forecasting Tourism Demand: Methods and Strategies", Butterworth-Heinemann*, Oxford.
6. Haykin, S (1999), *"Neural Networks: A Comprehensive Foundation", Prentice-Hall*, New Jersey.
7. Hua, L., Xing, L and Shuang, W (2017). "Based on SARIMA-BP Hybrid model and SSVM model of international crude oil price prediction research". *Anziam J.*58(E) pp. E143-E161.
8. Huang, G.B., Zhu, Q.Y and Siew, C.K. (2006) "Extreme Learning Machine: Theory and Applications". *Neurocomputing*, 70, 489-501.
9. Loganathan, Nanthakumar and Ibrahim, Y. (2010). "Forecasting International Tourims demand in Malaysia Using Box Jenkins SARIMA Application". *South Asian Journal of Tourism and Heritage*. 3(2), 50-60.
10. Nachane .D.M (2006). *"Econometrics: Theoretical Foundations and Empirical Perspectives"*. Oxford University Press, India.

11. Nor, M.E, Khamis, A, Saharan, S, and Abdullah, M.A.A. (2016). “Malaysia tourism demand forecasting by using time series approaches. *The Social Sciences*.11 (12), 2938-2945
12. Petrevska, B. (2017). “Predicting Tourims Demand by ARIMA Models”.*Economic Research-Ekonomska Istrazivanja.* 30(1), 939-950.
13. Samsudin, R., Saad, P., and Shabri, A. (2010). “Hybridizing GMDH and Least Squares SVM Support Vector Machine for Forecasting Tourism Demand. *IJRRAS* 3(3), 274-279.
14. Singh, E.H. (2013). “Forecasting Tourist Inflow in Bhutan Using Seasonal ARIMA”. *International Journal of Science and Research (IJSR)*. 2(9), 242-245.
15. Tseng, F., M, Yu, H.C.,Tzeng, G.H. (2002). “Combining Neural Network model with seasonal time series ARIMA model. *Technological Forecasting & Social Change*, 66(1),71-87.
16. www.tourism.gov.np
17. Zhang, G. (2003) “Time series forecasting using a hybrid ARIMA and neural network model”, *Neurocomputing* 50, pages: 159–175.