Forecasting Tourist Arrivals using a Combination of Long Short-Term Memory and Fourier Series

Ani Shabri[1], Ruhaidah Samsudin[2], Faisal Saeed[3] and Mohammed Al-Sarem[4]

[1]Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81300, Johor, Malaysia.
[2]School Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81300, Johor, Malaysia.
[3]DAAI Research Group, Department of Computing and Data Science, School of Computing and Digital Technology, Birmingham City University,Birmingham B4 7XG, UK.
[4]College of Computers and Engineering, Taibah University, Medina, Saudi Arabia.
ani@utm.my,ruhaidah@utm.my,faisal.saeed@bcu.ac.uk, msarem@taibahu.edu.sa

Abstract. The sector that contributes most to the nation's economy nowadays is tourism. Policymakers, decision-makers, and organisations involved in the tourist sector can use tourism demand forecasting to gather important information for planning and making important decisions. However, it is difficult to produce an accurate forecast because tourism data is critical, especially when a periodic pattern, such as seasonality, trends, and non-linearity, is present in a dataset. In this research, we present a hybrid modelling approach for modelling tourist arrivals time series data that combines the long short-term memory (LSTM) with the Fourier series method. This method is proposed to capture the components of seasonality and trend in the data set. Various single models, such as ARIMA and LSTM, as well as a modified ARIMA model based on Fourier series, are evaluated to confirm the suggested model's accuracy. The efficiency of the proposed model is compared using monthly tourism arrivals data from Langkawi Island, which has a notable pattern and seasonality. The findings reveal that the proposed model is more reliable than the other models in forecasting tourist arrivals series.

Keywords: Fourier series, artificial neural network, long short-term memory, ARIMA, tourist arrival.

1 Introduction

Tourism is a major source of economic activity, employment, and income generation. Accurate visitor arrivals forecasting is essential for stakeholders and researchers to make operational decisions such as providing appropriate operational funds, investing, money planning, and predicting future risks.

For tourist forecasting, various statistical time series prediction approaches have been proposed in the past, including autoregressive integrated-moving average (ARIMA) [1-9], exponential smoothing [1-5], Nave [1-3] and moving average (MA) [4-5]. Because of its high predicted precision and flexibility in representing a wide range of time series, the ARIMA model is one of the most widely used statistical models. However, due to its linearity, the ARIMA model is insufficient for modelling complex real-world time series with significant seasonality [10].

Artificial intelligence (AI) methods including support vector machines [1,11,12], fuzzy time series [14] and artificial neural networks (ANN) [1,3,5,6,8,13] have been more popular in tourist arrival forecasting during the recent year. These AI-based models are capable of explaining non-linear data. Finding the best network model and training algorithms for ANN apps, on the other hand, remains a challenge [15]. In addition, more parameters must be created in order to develop the ANN model, which can lead to overfitting and consequently bad predicting abilities. [16].

Lately, it was discovered that Recurrent Neural Network (RNN) designs are better suited to coping with the intricacy of time series than ANN [17]. However, because to the issue of vanishing gradients in RNN training, new RNN variants, such as long short-term memory (LSTM), have been proposed. Because of this property, the LSTM model has been used to solve forecasting problems [18-20].

Despite the obvious advantages of ARIMA, ANN, and LSTM models, it is not always adequate to use any single model to identify time series data with complexities and seasonal fluctuations. Several previous research have recommended combining several prediction models to improve the performance of an single statistical model or artificial intelligence model. The combined model's main goal is to overcome the inadequacies of single models and establish a predictive synergy. Several studies have also shown that combining prediction models is an excellent approach in time series analysis, such as combining the SVR-ANN [2], ARIMA-ANN [1,9,21], ARIMA-regression [22], ARIMA-deep learning model [23], ANN-Grey–Markov [24], ARIMA-GARCH [21], Grey-Fourier series [25,26] and ARIMA-Fourier series [27,28].

As a result, the primary goal of this research is to assess the efficacy of a combined prediction model based on the LSTM and Fourier series in terms of improving prediction accuracy. A sample of monthly tourism arrivals on Malaysia's Langkawi Island was chosen for a data analysis study to assess the efficacy of the proposed approach.

2 Forecasting Models

This section describes forecasting models such as ARIMA, LSTM, and proposed combining models.

2.1 Arima Model

The ARIMA model was among the most commonly used time series models [29]. The structures of ARIMA(p, d, q)(P, D, Q)s for non-seasonal and seasonal compounds are as follows:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D x_t = \theta_q(B)\Theta_Q(B^s)a_t \tag{1}$$

where d and D denote the degree of non-seasonal and seasonal differentiation,

respectively, s is the season's length and a_t is the errors. $\phi_p(B)$ and $\theta_q(B)$ are the order p and q parameters of the autoregressive and moving average, respectively.; $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are the order P and Q parameters of the seasonal autoregressive and moving average, respectively; The ARIMA was developed in four stages, according to reference [29]: identification, estimating parameter, diagnostic procedures, and forecasting.
The ARIMA models in this study were discovered with R software, which was built on the Forecast package [30]. The seasonal unit root test and the Osborn-Chui-Smith-Birchenhall test [31] are used to determine the number of d and D. The Akaike Information Criterion (AIC) is used to determine the order values of p, q, P, and Q. The best ARIMA model for modelling is the one with the lowest AIC value.

2.2 LSTM Model

Hochreiter and Schmidhuber introduced the LSTM model [32] and has been used to achieve well-known results for many sequential data problems. LSTMs learn long-term dependencies using a mechanism known as gates. These gates can decide which data in a sequence should be kept and which should be disposed. An LSTM's three gates are input, forget, and output. The construction of an LSTM cell is shown in Figure 1[33]. The cell state is the horizontal line between C_{t-1} and C_t. The heart of the LSTM model, where data is added and deleted from memory via pointwise addition and multiplication. The LSTM block's input and forget gates, as well as the output "tanh" activation function, are used to conduct these tasks. The following are the computations inside the LSTM neurons:

Forget gate: $f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f)$ (2)
Input gate: $i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i)$ (3)
Process input: $\tilde{C}_t = tanh(W_C.[h_{t-1}, x_t] + b_C)$ (4)
Cell update: $C_t = f_t \times C_{t-1} + i_t \widetilde{\times C_t})$ (5)
Output gate: $o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o)$ (6)
Output: $h_t = o_t \tanh(C_t)$ (7)

where b_f, b_o, $b_i$and W_f, W_o, W_i are the bias and weight matrices of the forget, output and input gates, respectively. b_C and W_C are the bias and weights of the cell state. h_t calculates the final outputs σ is sigmoid function, the previous cell state's output is h_{t-1}, the current cell state's input is x_t and o_t determines whether part of the cell state should be exported.

The forget gates are formed by combining the present time step input with the prior time step's hidden state, as shown in Eq. 2. Eq. 3 is an input gate that directs cell state adjustments to all neural net cells. Eq. 4 and 5 upgrade the cell states by transferring fresh information and the previous cell state from Eq. 2 and 3. to Eq. 5. Eq. 6 and 7 are being used to generate the current time step's hidden state by combining the hidden and updated cell states from the previous time step from Eq. 5.

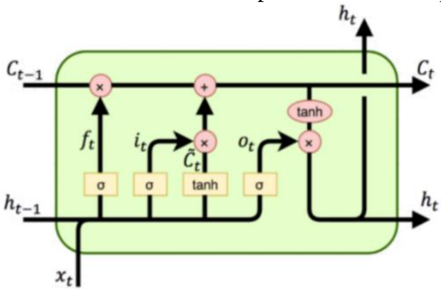

Fig.1. Structure of LSTM model

The loss function was mean squared error, and the "Adam" optimization technique [33] has been used to identify the best weight values for the networks. The data was normalised to a 0 to 1 scale. The function used to normalise the dataset in this investigation is described by Eq. (8):

$$y_t = \frac{x_t}{x_{max}} \tag{8}$$

where x_t is the actual value, y_t is the normalized time series, and x_{max} is the maximum actual value.

2.3 The Fourier LSTM Model

Fourier series have been applied to improve forecast model precision by modifying forecast model residuals. The methodology for developing a modified forecasting method known as the Fourier LSTM (FLSTM) model is as follows:

Step 1: Fit the LSTM model to a data set y_t.

Step 2: Create the residual series,-- $e_t = y_t - \hat{y}_t$ based on the LSTM predicted values $\hat{y}_t$.

Step 3: Fit the Fourier series to a residual e_t as shown below.

$$e_t = \frac{a_0}{2} + \sum_{t=1}^{M} a_t cos\left(\frac{2tk\pi}{n-1}\right) + b_t sin\left(\frac{2tk\pi}{n-1}\right) \tag{9}$$

where $M = \frac{n-2}{2} - 1$, $e = AB$,

$A = (0.5 \quad P_1 \ \ldots \ P_k \ldots \ P_F)$,

$$B = (a_0, a_1, b_1, a_2, b_2, \ldots, a_M, b_M)^T$$

$$A_k = \begin{pmatrix} cos\left(\frac{2\times 2\pi k}{n-1}\right) & sin\left(\frac{2\times 2\pi k}{n-1}\right) \\ cos\left(\frac{3\times 2\pi k}{n-1}\right) & sin\left(\frac{3\times 2\pi k}{n-1}\right) \\ \vdots & \vdots \\ cos\left(\frac{n\times 2\pi k}{n-1}\right) & sin\left(\frac{n\times 2\pi k}{n-1}\right) \end{pmatrix} \tag{10}$$

The ordinary least squares (OLS) method is used to estimate the parameters *B*.

$$B = (A^T A)^{-1} A^T e^T \quad (11)$$

The estimated residual by the Fourier series can be calculated using the following equation.

$$\hat{e}_t = \frac{a_0}{2} + \sum_{t=1}^{M} a_t cos\left(\frac{2tk\pi}{n-1}\right) + b_t sin\left(\frac{2tk\pi}{n-1}\right) \quad (12)$$

Step 4: Finally, the predicted values of the FLSTM model can be calculated by adding the predicted values of the LSTM model to the estimated residual from the Fourier series.

$$\tilde{y}_t = \hat{y}_t + \hat{e}_t \quad (13)$$

3. Examining The Performance

The proposed model's efficiency was measured using the mean absolute percentage error (MAPE) and root mean square error (RMSE).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where y_i, $\hat{y}_i$ and *n* denote the actual, predicted, and overall amount of datasets, respectively. The best forecasting model is the one with the lsmallest MAPE and RMSE values.

3.1 Analysis Data

This paper analysed tourist numbers arrivals to Langkawi Island in Malaysia, as reported on the Langkawi Development Authority (LADA) Malaysia website. This study focuses on examining the monthly tourism arrivals from January 2002 to December 2016. The training data covers the period from January 2002 to December 2015, whereas the testing data covers the period from January 2016 to December 2016. Figure 3 shows the monthly tourism arrivals to Langkawi Island (January 2002-December 2016). The monthly tourism arrivals curve, as seen in Fig. 3, has a trend, is non-stationary, and has a seasonal pattern.

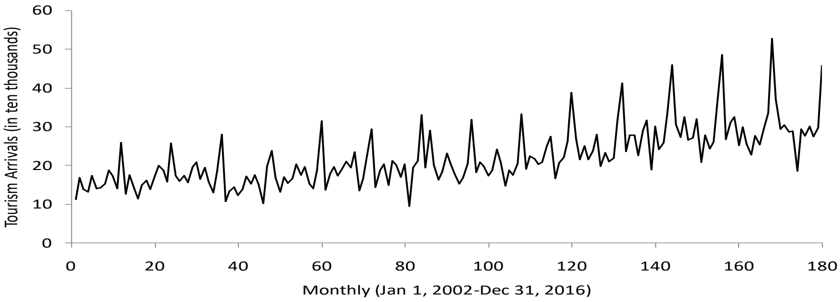

Fig. 2 Monthly visitor arrivals on Malaysia's Langkawi Island from 2002 to 2016.

ARIMA, LSTM, and FARIMA models were used in this study to evaluate the efficacy of the proposed model.

3.2 The ARIMA model fit to the data

The ARIMA model was identified and optimised using R Forecast package [37]. By default, this package specifies the maximum order for *d*, *P* and *Q* to 2, *D* to 1 and *p* and *q* to 5. The ARIMA model parameters were estimated using the maximum likelihood estimate (MLE) and the Ljung-Box (LB) test was used to confirm that the model selected was suitable for the data. The AIC is used to choose values for *p, P, q,* and *Q,* while the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) is applied to choose values for *d* and *D*.

The AIC and p-value of the Ljung Box test for numerous ARIMA models are summarised in Table 1. The five models considered were found to be acceptable by the Ljung-Box test, with ARIMA(1,0,2)(0,1,1)12 having the smallest AIC values. The model can be expressed as follows:

$$x_t = 941.745 + x_{t-12} + 0.771x_{t-1} - 0.771x_{t-13} + (1 - 0.810B + 0.287B^2)(1 - 0.330B^{12})a_t$$

The ARIMA model's AR term is as follows:

$$x_t = f(x_{t-1}, x_{t-12}, x_{t-13})$$

Table 1: AIC and p-values of Ljung-Box test for ARIMA models

No	ARIMA (p,d,q)(P,D,Q)$_{12}$	*p*-value of LB Test			AIC
		lag=12	lag=24	lag= 36	
1	$(2,0,2)(0,1,1)_{12}$	0.977	0.996	0.219	3406.7
1	$(2,0,3)(0,1,1)_{12}$	0.999	0.999	0.281	3405.6
2	$(1,0,2)(0,1,2)_{12}$	0.997	0.996	0.226	3406.4

4	$(1,0,2)(0,1,1)_{12}$	0.997	0.996	0.226	3404.3
5	$(1,0,3)(0,1,1)_{12}$	0.997	0.996	0.207	3406.4

Bold values indicate the smallest AIC

I. Input Time Lag Selection

The selection of significant input variables is one of the most critical processes in the LSTM model generation process. The input variables were determined using three ways in this investigation.

i. Packard et al. [34] suggested in the selection of the input variables, the number of nodes in the input layer equals the number of lag variables. The y_t output is given as

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p})$$

where p is a the number of period for seasonal.

ii. The lagged variables of the ARIMA model are the most significant variables to use as input variables for AI models [35]. The AR term from the ARIMA model is defined in Table 2 as

$$x_t = f(x_{t-1}, x_{t-12}, x_{t-13})$$

iii. Prastyo et al. [36] suggested PACF in determination input variables of AI models.

Fig. 4 and 5 show the decomposition plot and PACF graph. Through observing Fig. 4, it is obvious that the data has a seasonal pattern with seasonal period is 12. Fig. 5 shows that the PACF significance at lags 1, 2, 3, 6, and 11 to 15.

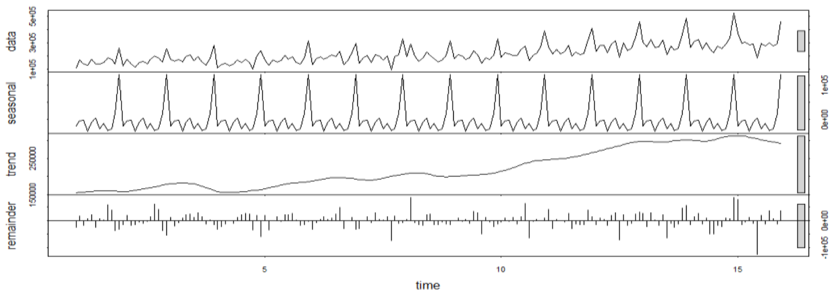

Fig.3: The decomposition plot (Original data (top) and its three additive components (bottom)).

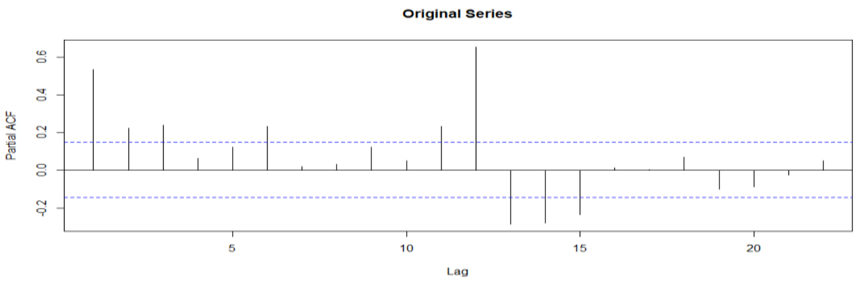

Fig.4: PACF of original data

As a result of the decomposition plot, PACF and AR terms, the LSTM model input may be expressed as shown in Table 2.

Table 2: Input variables for LSTM model

Model	Method	Input
LSTM1	Period 12	$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-12})$
LSTM2	AR	$x_t = f(x_{t-1}, x_{t-12}, x_{t-13})$
LSTM3	PACF	$x_t = f(x_{t-1}, x_{t-2}, x_{t-3}, x_{t-6}, x_{t-11}, x_{t-12}, x_{t-13}, x_{t-14,} x_{t-15})$

3.3 LSTM model fit to the data

During the training phase of this study, we examined all possible combinations to find the best parameters for LSTM1, LSTM2, and LSTM3. The number of epochs is 300, the batch size is 4, and the Adam optimization algorithm [37] with a default-learning rate of 0.01 is used throughout this study. The mean square error is used as the loss function.

The testing dataset gave RMSEs of 44367.33, 61411.02, and 55492.73 for the LSTM1, LSTM2, and LSTM3 models, respectively. LSTM1, LSTM2, and LSTM3 have MAPEs of 11.462, 19.407, and 16.549, respectively. The LSTM2 model has the lowest RMSE and MAPE values and show performs better in forecasting.

3.4 The LSTM and Fourier Series Fit to the Data

The residual series from the best ARIMA and LSTM models were then combined with Fourier series to form the Fourier-ARIMA (FARIMA) and Fourier-LSTM (FLSTM) models, respectively. Table 3 demonstrates the RMSE and MAPE performance of the ARIMA and LSTM models, as well as their Fourier series. In Table 3, the ARIMA, LSTM, FARIMA, and FLSTM models had RMSEs of 49319.52, 44367.35, 34113.20, and 29148.60, respectively. ARIMA, LSTM, FARIMA, and FLSTM models had MAPEs of 14.64%, 11.46%, 10.92%, and 9.28%, respectively.

The forecasting accuracy of the single ARIMA and LSTM models improved when the Fourier residual estimation was used, as shown in Table 3. The FLSTM model had the best overall testing phase performance (RMSE = 29148.60 and MAPE = 9.28%).

This shows that the FLSTM model has fewer errors than single ARIMA and LSTM models, as well as a combination model, the FARIMA model. According to the RMSE and MAPE criteria, the FLSTM model has a strong forecasting ability. According to the findings above, the FLSTM model received the highest ranking of all the models considered.

Table 3 Model performance for monthly tourism arrivals data

Model	RMSE	MAPE (%)
ARIMA	49319.52	14.64
LSTM	44367.35	11.46
FARIMA	34113.20	10.92
FLSTM	29148.60	9.28

The real data versus the ARIMA, FARIMA, LSTM, and FLSTM models' forecasted values are presented in Fig. 5. In comparison to other approaches, the FLSTM model performed excellently and closely followed the true data. The findings suggest that the FLSTM models can accurately anticipate the amount of tourists arrivals. The empirical findings indicate that the FLSTM model may be successful in improving model forecasting accuracy in a series of seasonal monthly tourist arrival to Langkawi Island.

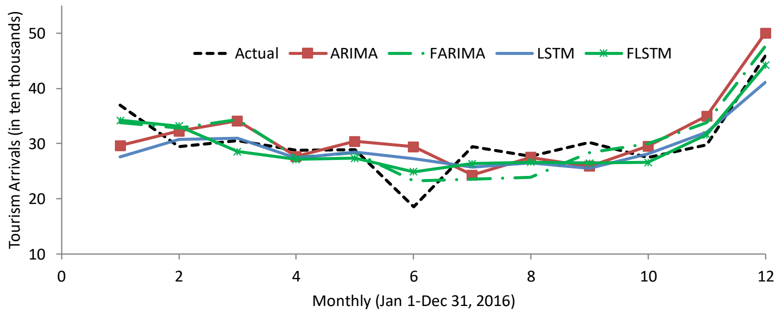

Fig.5 Actual tourism arrivals versus forecasted values

4 Conclusion

In this study, we set out to forecast the number of visitors to the Langkawi region using two traditional methods. To solve the shortcomings of each method, we then combined the two methods in hybrid models utilising Fourier series. The results demonstrated that the LSTM performed better than the ARIMA technique, which produced substantial output values and accurately represented the pattern of the actual data. The hybrid model that combined LSTM and Fourier series was consistently better accurate when compared to other combination technique and the traditional techniques. These findings show that the ARIMA or LSTM model alone cannot be used to forecast visitor arrivals. However, as each technique captures a certain characteristic of the data, researchers should combine different techniques for results with high significance in order to overcome each technique's limitations. More study might examine additional hybrid models, such as those that employ alternative weighting techniques when combining more than two models in an ensemble learning technique. Other research might compare various ensemble learning methodologies when employing mixed-frequency data, which has emerged as the new standard in forecasting tourist demand.

Acknowledgment

The authors are grateful to Universiti Teknologi Malaysia (UTM) and the Ministry of Higher Education Malaysia (MOHE) for their support of this project through the Fundamental Research Grant Scheme (R.J130000.7854.5F271).

References

[1] K.Y. Chen, “Combining linear and nonlinear model in forecasting tourism demand,” Expert Systems with Applications, vol. 38 (8), pp. 10368-10376, 2011.

[2] S. Cang, and H. Yu, “A combination selection algorithm on forecasting,” European Journal of Operational Research, vol. 234 (1), pp. 127-139, 2014.

[3] X. Xu, R. Law, W. Chen, and L. Tang, “Forecasting tourism demand by extracting fuzzy Takagi Sugeno rules from trained SVMs CAAI,” Transactions on Intelligence Technology, vol. 1(1), pp. 30-42, 2016.

[4] I. Önder and U. Gunter, “Forecasting tourism demand with Google Trends for a major European city destination,” Tourism Analysis, vol. 21 (2–3), pp. 203-220, 2016.

[5] H. Hassani, E.S. Silva, N. Antonakakis, G. Filis and R. Gupta. “Forecasting accuracy evaluation of tourist arrivals,” Annals of Tourism Research, vol. 63, pp. 112-127, 2017.

[6] M. A. Díaz and J.R. Nadal, “Forecasting British tourist arrivals in the Balearic Islands using meteorological variables, “ Tourism Economics, vol. 16 (1), pp. 153-168, 2010.

[7] S. Park, J. Lee and W. Song, “Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data,” Journal of Travel & Tourism Marketing, vol. 34(3), pp. 357-368, 2017.

[8] S. Li, T. Chen, L. Wang and C. Ming, “Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index,” Tourism Management, vol. 68, pp. 116-126, 2018.

[9] M. E. Nor, A. I. Nurul and M. S. Rusiman, “A hybrid approach on tourism demand forecasting,” Journal of Physics: Conference Series, vol. 995(1), pp. 1-11, 2018.

[10] P. S. G. Neto, G.D.C. Cavalcanti and F. Madeiro, “Nonlinear combination method of forecasters applied to PM time series. Pattern Recognition Letters, Vol. 95(1), pp. 65-72, 2017.

[11] W.C. Hong, Y. Dong, L.Y. Chen and S.Y. Wei, “SVR with hybrid chaotic genetic algorithms for tourism demand forecasting,” Applied Soft Computing, vol.11 (2), pp. pp. 1881-1890, 2011.

[12] R. Chen, C.Y. Liang, W.C. Hong and D.X. Gu, “Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm,” Applied Soft Computing, vol. 26, pp. 435-44, 2015.

[13] R. Law and N. Au, “A neural network model to forecast Japanese demand for travel to Hong Kong,” Tourism Management, vol. 20(1), pp. 89‑97, 1999.

[14] E. Hadavandi, A. Ghanbari, K. Shahanaghi and S. Abbasian-Naghneh, “Tourist arrival forecasting by evolutionary fuzzy systems,” Tourism Management, vol. 32 (5), pp. 1196-1203, 2011.

[15] H. R. Maier and G. C. Dandy, "Application of artificial neural networks to forecasting of surface water quality variables: issues, applications and challenges" in Artificial Neural Networks in Hydrology, Kluwer Academic Publishers, The Netherlands, 2000.

[16] W. Sun and X. Zhang, “Application of Self-Organizing Combination Forecasting Method in Power Load Forecast,” Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Beijing China, 2007.

[17] T. G. Barbounis, J. B. Teocharis, M. C. Alexiadis, and P. S. Dokopoulos, “Long-term wind speed and power forecasting using local recurrent neural network models,” IEEE Transactions on Energy Conversion, vol. 21, no. 1, pp. 273–284, 2006.

[18] J. Brownlee, “Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” https://machinelearningmastery.com/ time-series-prediction-lstm-recurrent-neural-networks-python-keras/, 2016.

[19] F. A. Gers, J. Schmidhuber, F. Cummins, “Learning to Forget: Continual Prediction with LSTM,” in Neural Computation 12(10): 2451- 2471, 2000.

[20] S. Hochreiter, J. Schmidhuber, “Long Short-Term Memory,” Neural Computation 9(8):1735-1780, 1997.

[21] J. T. Coshall, “Combining volatility and smoothing forecasts of UK demand for international tourism,” Tourism Management,vol. 30(4), pp. 495-511, 2009.

[22] F. L. Chu, “Forecasting tourism: A combined approach. Tourism Management, vol. 19(6), pp. 515-520, 1998.

[23] Y. Chen, K. He and G. K. F. Tso, “Forecasting crude oil prices: a deep learning based model,” Procedia Computer Science, vol. 122, pp. 300-307, 2017.

[24] Y. C. Hu, P. Jiang and P.C. Lee, “Forecasting tourism demand by incorporating neural networks into Grey–Markov models,” Journal of the Operational Research Society, 10.1080/01605682.2017.1418150, 2018.

[25] Z. X. Wang and L. L. Pei, “Forecasting the international trade of Chinese high-tech products using an Fourier Nash nonlinear grey Bernoulli model,” 2014 Seventh International Joint Conference on Computational Sciences and Optimization, 2014.

[26] W. Chia-Nan and P. Van-Thanh, “An Improved Nonlinear Grey Bernoulli Model Combined with Fourier Series,” Mathematical Problems in Engineering, vol. ID 740272, pp. 1-7, 2015.

[27] T.F. Huang, P.J. Chen and T.L. Nguyen, “Forecasting with Fourier Residual Modified ARIMA Model- An Empirical Case of Inbound Tourism Demand in New Zealand. Recent Researches in Applied Economics and Management,” vol.2, pp. 61-65, 2013.

[28] T. L. Nguyen, P. J. Chen, M. H. Shu, B. M. Hsu an Y. C. Lai, “Forecasting with Fourier Residual Modified Arima Model- The Case Of Air Cargo In Taiwan,” Proceeding of 2013 International Conference on Technology Innovation and Industrial Management, Phuket Thailand, 2013.

[29] G. E. P. Box and G.M. Jenkins, “Time series analysis: Forecasting and control,” San Francisco: Holden-Day, 1976.

[30] R. J. Hydman and Y. Khandakar, “Automatic time series forecasting: The forecast package for R,” Journal of Statistical Software, vol. 27(3), pp. 1–22, 2008.

[31] D. R. Osborn, A. P. L. Smith and C. R. Birchenhall, “Seasonality and the order of integration for consumption. Oxford Bulletin of Economics and Statistics,” vol. 50(4), pp. 361–377, 1988.

[32] J. Patterson, Deep Learning: A Practitioner’s Approach, OReilly Media, 2017.

[33] Wegayehu E.B. and F.B. Muluneh. Short-Term Daily Univariate Streamflow Forecasting Using Deep Learning Models. Advances in Meteorology. ID 1860460, 1-21, 2022.

[34] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, Geometry from a time series, Phys. Rew. Lett., 45 (1980), no. 9, 712-716. http://dx.doi.org/10.1103/physrevlett.45.712

[35] S. BuHamra, N. Smaoui and M. Gabr, “The Box-Jenkins analysis and neural network: prediction and time series modeling,” Applied Mathematical Modelling, vol. 27, pp. 805-815, 2003.

[36] Prastyo, D. D., Nabila, F. S., Suhartono, Lee, M. H., Suhermi, N., & Fam, S.-F. VAR and GSTAR-Based Feature Selection in Support Vector Regression for Multivariate Spatio-Temporal Forecasting. International Conference on Soft Computing in Data Science (pp. 46-57). Singapore: Springer. (2019).

[37] Zhang R., Song H., Chen Q., Wang Y., Wang S., and Li Y. Comparison of ARIMA and LSTM for prediction of hemorrhagic fever at different time scales in China. Plos One, 17(1):1-14, 2022.