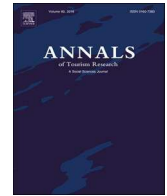


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Annals of Tourism Research

journal homepage: www.elsevier.com/locate/annals

A decomposition-ensemble approach for tourism forecasting

Gang Xie^{a,*}, Yatong Qian^{a,b}, Shouyang Wang^a^a CFS, MDIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China^b University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Associate editor: Haiyan Song

Keywords:

Tourism demand
 Complete ensemble empirical mode decomposition with adaptive noise
 Data characteristic analysis
 Time series forecasting

ABSTRACT

With the frequent occurrence of irregular events in recent years, the tourism industry in some areas, such as Hong Kong, has suffered great volatility. To enhance the predictive accuracy of tourism demand forecasting, a decomposition-ensemble approach is developed based on the complete ensemble empirical mode decomposition with adaptive noise, data characteristic analysis, and the Elman's neural network model. Using Hong Kong tourism demand as an empirical case, this study firstly investigates how data characteristic analysis is used in a decomposition-ensemble approach. The empirical results show that the proposed model outperforms other models in both point and interval forecasts for different prediction horizons, indicating the effectiveness of the proposed approach for forecasting tourism demand, especially for time series with complexity.

Introduction

Tourism has become an emerging sector of economic growth and even a pillar of industry in many regions (Gao, Zhang, Lu, Wu, & Du, 2018; Hassani, Silva, Antonakakis, Filis, & Gupta, 2017). The benefits of tourism are many; they include catering, accommodation, transportation, entertainment, and retailing. Hence, a need exists for substantial investments in infrastructure and promotional activities for tourism (Peng, Song, & Crouch, 2014; Witt & Witt, 1995). Clearly, accurate tourism demand forecasting is a prerequisite for decision-making relating to investments and planning (Chatziantoniou, Degiannakis, Eeckels, & Filis, 2016; Dharmaratne, 1995). Making accurate predictions is particularly critical, given the perishable nature of tourism products. For example, unoccupied hotel rooms and airline seats cannot be stored. With accurate predictions of future tourist arrival numbers, the government could formulate a more appropriate development strategy and provide better service to tourists. Simultaneously, accurate forecasts could help the private sector develop more suitable marketing strategies, thus generating greater revenue. Therefore, developing effective models for forecasting tourism demand is a critical issue in the tourism industry (Shahrabi, Hadavandi, & Asadi, 2013).

The time series of tourist arrivals can be influenced by various underlying factors, such as irregular events, economic crises (Song, Qiu, & Park, 2019) and seasonal variations (Goh & Law, 2002; Hyndman & Athanasopoulos, 2018). For example, the severe acute respiratory syndrome (SARS) epidemic in 2003, and the financial crisis of 2008, both caused structural breaks in the time series of tourist arrivals in many destinations (Cró & Martins, 2017; Narayan, 2005). With the frequent occurrence of irregular events in recent years, the tourism industry in some areas, such as Hong Kong, has suffered great volatility. Consequently, the time series of tourist arrivals exhibits many non-stationary, nonlinear and complex patterns, which makes it difficult to forecast tourism demand.

For a time series with complexity, the “divide and conquer” principle is an effective means to enhance tourism predictive accuracy

* Corresponding author at: 55 Zhongguancun East Road, Haidian District, Beijing 100190, China.

E-mail addresses: gxie@amss.ac.cn (G. Xie), sywang@amss.ac.cn (S. Wang).

(Chen, Lai, & Yeh, 2012). Within this principle, decomposition-ensemble approaches are developed, in order to simplify the difficulty of a forecasting task by dividing it into a number of relatively easier subtasks. Then, the prediction results of the decomposed subtasks are combined into an ensemble forecast (Yu, Wang, & Tang, 2015). In short, the principle consists of three steps: (1) data decomposition, (2) individual prediction, and (3) ensemble prediction.

As a method of the “divide and conquer” principle, the empirical mode decomposition (EMD) method has outstanding advantages in processing various non-stationary and complex time series. Moreover, it can help to analyze the impacts of irregular events on the decomposed time series with different time scales and frequencies, and explain the generation of time series data (Zhang, Yu, Wang, & Lai, 2009). However, EMD also has the problem of mode mixing. Consequently, the ensemble empirical mode decomposition (EEMD) (Wu & Huang, 2009) and the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) (Torres, Colominas, Schlotthauer, & Flandrin, 2011) methods are developed. With CEEMDAN, the rate of reconstruction errors is almost zero, and the cost of calculation is also greatly reduced (Colominas, Schlotthauer, & Torres, 2014). Therefore, the CEEMDAN method may be an effective means with which to deal with the complex time series of tourist arrivals.

This study proposes a decomposition-ensemble approach based on CEEMDAN and data characteristic analysis for tourism demand forecasting. Firstly, the time series of tourist arrivals are decomposed by the CEEMDAN method. Secondly, multiscale permutation entropy is used to measure the complexity of decomposed components, and components are reconstructed. Then, each component is predicted by an Elman's neural network, according to data characteristics of mutability and seasonality. Finally, the individual forecasts of the reconstructed components are aggregated into point forecasts of tourism demand, and interval forecasts are generated, with point forecasts as the means. Using the sample data from the top nine sources pertaining to tourist arrivals at Hong Kong, an empirical study is implemented. Finally, we analyze the related problems and draw a conclusion.

The rest of this study is organized as follows: “Literature review” section reviews the most closely-related studies on forecasting models of tourism demand, decomposition techniques and CEEMDAN-based decomposition-ensemble approaches. “Methodology” section describes the methodology. “Empirical study” section illustrates the proposed model by using an empirical analysis. “Discussion” section discusses a number of related issues. “Conclusions” section gives our conclusions and provides some topics for future research.

Literature review

In this section, we first introduce typical models and decomposition techniques for tourism demand forecasting. Then, current literature on CEEMDAN-based decomposition-ensemble approaches is reviewed. Finally, some problems in related literature are put forward.

Tourism demand forecasting models

Different models have been developed for tourism demand forecasting (Athanasopoulos, Hyndman, Song, & Wu, 2011). There are single econometric models, including exponential smoothing (ETS) models (Lim & McAleer, 2001), autoregressive integrated moving average (ARIMA) models (Cho, 2003; Kulendran & Shan, 2002; Ma, Liu, Li, & Chen, 2016), the autoregressive distributed lag model (Song, Witt, & Jensen, 2003), the vector autoregressive model (Song & Witt, 2006), and piecewise linear models (Chu, 2011). Generally, econometric models are used to model linearly correlated data, while nonlinearly correlated data are captured by machine learning models, such as the artificial neural network (ANN) models (Kon & Turner, 2005; Law & Au, 1999) and the support vector regression model (Chen, Liang, Hong, & Gu, 2015; Hong, Dong, Chen, & Wei, 2011).

Compared with studies relating to single models, relatively few studies have focused on hybrid models for tourism demand forecasting. To predict tourist arrivals in Turkey, Aslanargun, Mammadov, Yazici, and Yolacan (2007) investigated various combinations of ARIMA, linear ANN, multilayer perceptron, and radial basis function network models. Also, Andrawis, Atiya, and El-shishiny (2011) proposed a combination method based on long-term and short-term forecasts for forecasting tourism demand. In addition, Shahrabi et al. (2013) presented a hybrid model, namely the modular genetic-fuzzy forecasting system, to forecast tourism demand. This model uses three main steps: (1) data preprocessing, (2) extracting a fuzzy rule-based system for each cluster, and (3) prediction. Wu and Cao (2016) proposed a novel model by hybridizing a support vector regression model with a fruit fly optimization algorithm and a seasonal index adjustment. To comprehensively evaluate the forecasting performance of this hybrid model, both one-step-ahead and multi-step-ahead predictions were conducted.

Decomposition techniques for tourism demand forecasting

In order to improve predictive accuracy, decomposition techniques are widely used for forecasting tourism demand. Burger, Dohnal, Kathrada, and Law (2001) proposed a classic decomposition method to predict tourism demand from the USA to Durban, South Africa. In this method, a 12-period moving average is created, before calculating a centered moving average and seasonal factors. The actual values are then divided by seasonal factors to give a one-year forecast.

There are also structural time series (STS) models, which are well-known in the tourism demand forecasting literature (Greenidge, 2001; Song, Wen, & Liu, 2019). The basic STS model decomposes a time series into its trend, seasonal, cycle and irregular components, which are regarded as stochastic. Combining the advantages of the STS model and the time-varying parameter regression approach, Song, Li, Witt, and Athanasopoulos (2011) developed a new model, which introduces a time-varying parameter estimation of the explanatory variable coefficients into a causal STS model. This new model is employed for modeling and predicting quarterly

tourist arrivals at Hong Kong. Data are acquired from four main sources: mainland China, Korea, the UK, and the USA. The results indicate that the proposed approach could achieve higher forecasting accuracy than seven alternative models. To forecast monthly tourist arrivals in Turkey from different countries, [Akin \(2015\)](#) proposed a method based on STS models, which are linear Gaussian state-space models based on a decomposition of the time series into a number of components. [Chen, Li, Wu, and Shen \(2019\)](#) proposed a multi-series STS method to improve seasonal tourism demand forecasting. Empirical analysis shows that the newly-proposed approach improves the degree of forecast accuracy, compared to traditional univariate models.

Singular spectrum analysis (SSA) is another decomposition technique that has been used for tourism demand forecasting. There are two steps in the basic SSA process: decomposition and reconstruction. SSA seeks to filter the noise and forecast the signal while the classical time series methods forecast both the signal and noise. [Hassani, Webster, Silva, and Heravi \(2015\)](#) examined the merits of using SSA, and found that the SSA approach could achieve higher accuracy than ARIMA, exponential smoothing or neural networks methods for the prediction of tourist arrivals in the USA. In addition, [Silva, Hassani, Heravi, and Huang \(2019\)](#) used SSA to acquire the denoised tourism demand data that improves the forecasting performance of the neural network autoregressive algorithm. The empirical analysis shows that the proposed models can generally achieve higher predictive accuracy levels than single models.

To deal with tourist arrival time series with complexity, [Chen et al. \(2012\)](#) developed a decomposition-ensemble approach based on both EMD and a back propagation neural network (BPNN). Within this approach, EMD is firstly used to decompose the complicated original data into several intrinsic mode functions and a residue, which are then modeled and predicted by using BPNN models. The final forecasts are obtained by aggregating these prediction results. A model comparison shows that the proposed model outperforms the single BPNN and ARIMA models. Moreover, [Zhang, Wu et al. \(2017\)](#) proposed a decomposition-ensemble approach based on EEMD and ARIMA for predicting daily occupancy of an individual hotel. Result shows that the proposed method can improve forecasting accuracy compared to the ARIMA method, especially for short-term forecasting.

CEEMDAN-based decomposition-ensemble approaches

With its ability to significantly reduce mode mixing in EMD and effectively enhance the predictive accuracy, the CEEMDAN method has been used for decomposition-ensemble approaches for time series forecasting.

Combining CEEMDAN, a flower pollination algorithm with chaotic local search, five ANN models and no negative constraint theory, [Zhang, Qu et al., \(2017\)](#) proposed a decomposition-ensemble approach to short-term wind speed forecasting. Also, [Liu, Mi, and Li \(2018\)](#) proposed a new framework using wavelet packet decomposition, CEEMDAN and the ANN model for multi-step forecasting of wind speed. The experimental results show that CEEMDAN-based models can achieve the highest degree of predictive accuracy.

Using CEEMDAN and a support vector machine optimized by a modified grey wolf optimization algorithm (MGWO-SVM), [Dai, Niu, and Li \(2018\)](#) proposed a novel decomposition-ensemble approach for the time series forecasting of daily peak loads. The investigation suggests that using CEEMDAN can help achieve noise reduction for the non-stationary time series of daily peak loads and make the time series more regular.

To predict the daily closing prices of major global stock indices, [Cao, Li, and Li \(2019\)](#) developed a decomposition-ensemble approach based on CEEMDAN and the long short-term memory (LSTM) model. Similarly, the approach includes three main steps, as follows: (1) First, the time series of a stock index is decomposed into several intrinsic mode functions and one residue by CEEMDAN. (2) A LSTM model is used for each component, in order to obtain all the prediction results. (3) The final forecast result is obtained by aggregating the prediction results of all components. The error analysis indicates that the proposed CEEMDAN-based decomposition-ensemble approach has lower rates of prediction errors than the EMD-based model.

According to a literature review, we can ascertain some problems, as follows: (1) Even though the EMD method has great advantages in non-stationary and complex time series, there is still the potential shortcoming of mode mixing in the EMD-based decomposition-ensemble approach. (2) The data characteristics of decomposed components are generally neglected in CEEMDAN-based decomposition-ensemble approaches. (3) The CEEMDAN method is more effective at eliminating mode mixing, but multiple scales may still exist in the decomposed components. Due to the current deficiency in existing literature, we propose a new decomposition-ensemble approach that can be used to forecast tourism demand.

Methodology

In this section, we first present the procedure of the proposed decomposition-ensemble approach. Then, the decomposition method, data characteristic analysis and component reconstruction are introduced. Finally, the tourism demand forecasting model is described.

Procedure of the proposed approach

Motivated by the “divide and conquer” principle, we developed a novel approach that can be used to overcome the current difficulties associated with the prediction of tourism demand with complexity. The procedure of the proposed approach consists of four main steps, as shown in [Fig. 1](#).

Step 1: Data decomposition. By using the CEEMDAN technique, the sample data x_t ($t = 1, 2, \dots, T$) of tourist arrivals is decomposed

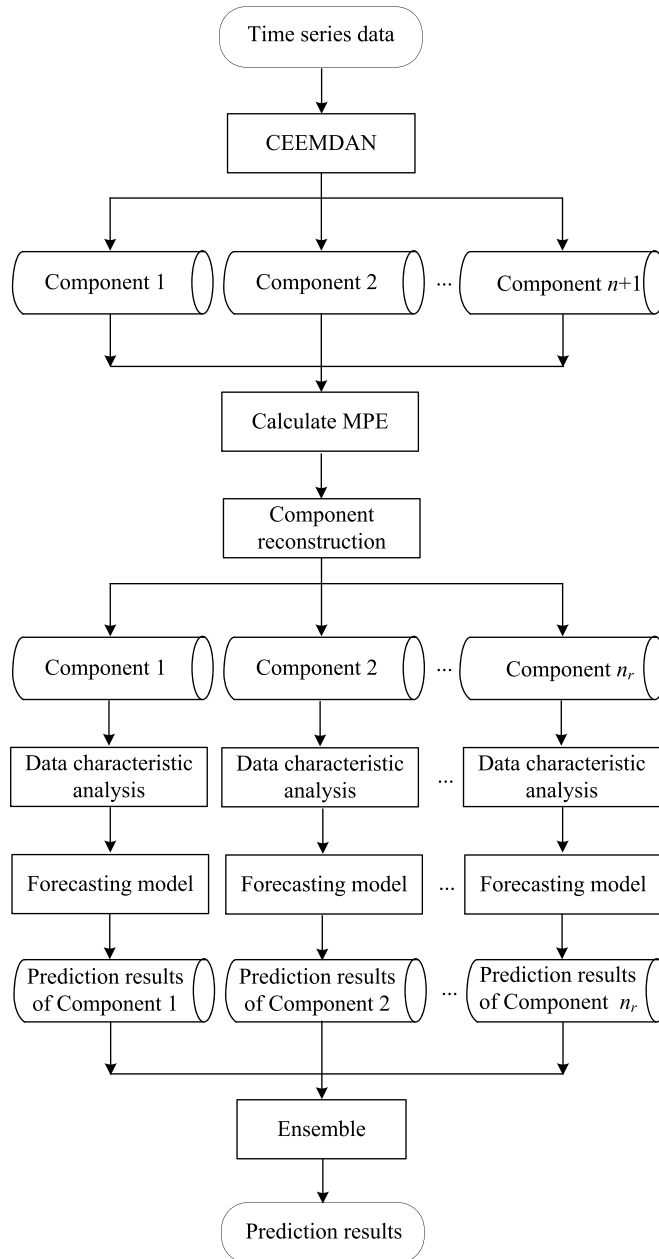


Fig. 1. The procedure of the proposed approach.

into $n + 1$ components (n intrinsic mode functions and a residual series).

Step 2: Component reconstruction. The decomposed

$n + 1$ components are reconstructed into

n_r components in terms of complexity testing results, in order to capture the inner data characteristics and reduce computational costs.

Step 3: Individual forecasts. A single Elman's neural network model is developed, based on data characteristic analysis (including the mutability and seasonality of the components), for the prediction of each reconstructed component.

Step 4: Ensemble forecast. The prediction results for all components are aggregated into point forecasts of tourism demand. Then, interval forecasts are generated, with point forecasts as the means. For more information about interval forecasting, please refer to Li, Wu, Zhou, and Liu (2019).

Complete ensemble empirical mode decomposition with adaptive noise

The methods used in the EMD family can effectively capture various hidden patterns in a complex time series, even without a priori knowledge. Huang, Shen, and Long (1971) proposed the EMD method as a means to decompose a time series into several intrinsic mode functions and a residual series, which are extracted by using an iterative “sifting process”. Within the EMD method, the original time series x_t can be expressed as follows:

$$x_t = \sum_{k=1}^n c_{kt} + r_t, \quad (1)$$

where c_{kt} ($k = 1, 2, \dots, n$) is the k th intrinsic mode function, and r_t is the residue.

To reduce the potential for mode mixing in the EMD method, the EEMD and other extended versions (e.g. the CEEMDAN method), were developed. In the EEMD technique, white noise is added to original data prior to decomposition. The difference ε_{ne} (standard deviation of error) between the input and the corresponding intrinsic mode functions is as follows:

$$\varepsilon_{ne} = \frac{\varepsilon}{\sqrt{NE}}, \quad (2)$$

where ε is the amplitude of the added white noise, and NE is the number of ensemble components.

The CEEMDAN method extends EEMD by adding adaptive Gaussian white noise to smooth pulse interference in data decomposition. This method takes advantage of the zero mean Gaussian white noise to make the data decomposition more complete. In CEEMDAN, the first intrinsic mode function is calculated exactly as in EEMD. Firstly, in I times of CEEMDAN computation, a collection of Gaussian white noise series is added to the original data as follows:

$$s^i(t) = x_t + \varepsilon_i v^i(t), \quad (3)$$

where $s^i(t)$ is the time series with the additional noise in the i th trial ($i = 1, 2, \dots, I$), x_t is the original data, ε_i is the ratio of the additional noise to the original signal, and $v^i(t)$ is the Gaussian white noise series. Then, EMD is used to obtain the first intrinsic mode function $I_1^i(t)$ as:

$$I_1(t) = \left(\sum_{i=1}^I I_1^i(t) \right) / I, \quad (4)$$

where $I_1^i(t)$ is the first intrinsic mode function obtained in the i th trial. Also, the first residue $r_1(t)$ is:

$$r_1(t) = x_t - I_1(t). \quad (5)$$

The k th residue is:

$$r_k(t) = r_{k-1}(t) - I_k(t), \quad (6)$$

where $I_k(t)$ is the k th intrinsic mode function of CEEMDAN. The $I_{k+1}(t)$ of CEEMDAN is:

$$I_{k+1}(t) = \sum_{i=1}^I E_k(r_k(t) + \varepsilon_k E_k(v^i(t))) / I, \quad (7)$$

where $E_k(\cdot)$ is the k th mode of EMD. Then, repeat Eqs. (6) and (7) until all the intrinsic mode functions are found. In order to efficiently capture the inner factors and further enhance the predictive accuracy of tourism demand forecasts, in this study, we develop a decomposition-ensemble approach based on the CEEMDAN method.

Data characteristic analysis

With the proposed model, we first decompose the time series of tourist arrival numbers. Then, we conduct data characteristic analyses, which include tests for the complexity, mutability and seasonality of components.

For a complexity test, a popular method is the sample entropy algorithm (Richman & Moorman, 2000). However, sample entropy may yield inaccurate values for a short time series. For a time series of T observations, m is the length of sequences to be compared (the dimension of the template vector). In order to make the sample entropy algorithm reasonable, it is suggested that T should be within the range of 10^m to 30^m (Liu et al., 2012; Wang, Zhang, Fan, & Ma, 2017). Particularly, when $m = 2$, T should be more than 750 (Costa, Peng, Goldberger, & Hausdorff, 2003). Therefore, sample entropy may be ineffective for a time series that is relatively short.

Compared with sample entropy, the permutation entropy method has the advantages of simplicity, low computational complexity, and robustness, with dynamical and observational noise (Bandt, Keller, & Pompe, 2002; Bandt & Pompe, 2002; Rosso, Larrondo, Martin, Plastino, & Fuentes, 2007). Generally, the traditional permutation entropy method is single-scale and so cannot handle the multiple scales in time series (Costa, Goldberger, & Peng, 2002, 2005). As a consequence, Ouyang, Li, Liu, and Li (2013) proposed multiscale permutation entropy (MPE), which consists of two steps, as follows: (1) “Coarse-graining” is applied to a time series. A given time series $\{x_1, x_2, \dots, x_T\}$ is transformed into a consecutive coarse-grained time series, where the element $y_j^{(s)}$ is calculated by

$$y_j^{(s)} = 1/s \sum_{i=(j-1)s+1}^{js} x_i. \tag{8}$$

Here, s is the scale factor, $1 \leq j \leq T/s$, and the length of the coarse-grained time series is an integer part of T/s . Obviously, when $s = 1$, the coarse-grained data are simply the original time series. (2) The permutation entropy (PE) value is computed for the coarse-grained time series (Bandt & Pompe, 2002).

Researchers have not as yet considered the problem of multiple scales in the time series of decomposed components in CEEMDAN-based decomposition-ensemble approaches. As there may be multiple scales in decomposed components with the CEEMDAN method, MPE is more suitable for complexity tests than the traditional single-scale permutation entropy used in Yu et al. (2015). In complexity tests, the time series of a decomposed component is considered to have high complexity if the PE value is higher than a threshold value of θ . To save on computational costs, the components are reconstructed, based on the complexity of each decomposed component. In other words, we combine the components with low complexity into a new component.

As for mutability testing, a potential breakpoint $k \in \{k_1, k_2, \dots\}$ is determined by the Chow test (Chow, 1960) via the F -test:

$$F = \frac{\frac{SSE - SSE_1 - SSE_2}{m_1 + 1}}{\frac{SSE_1 + SSE_2}{N_1 + N_2 - 2m_1 - 2}} \sim F(m_1 + 1, N_1 + N_2 - 2m_1 - 2), \tag{9}$$

where m_1 is the number of explanatory variables, and N_1 and N_2 are the numbers of observations in the segmented datasets before and after point k . In Eq. (9), the SSE is the sum of the residual squares of the whole time series with $N_1 + N_2 - m_1 - 1$ degrees of freedom, and SSE_1 and SSE_2 are the sums of the datasets before and after point k with degrees of freedom $N_1 - m_1 - 1$ and $N_2 - m_1 - 1$, respectively. When there are breakpoints in the time series of a reconstructed component, the dummy variable $p_{k,t}$ for the structural change at breakpoint k can be denoted as:

$$p_{k,t} = \begin{cases} 0 & t < k \\ 1 & t \geq k \end{cases}. \tag{10}$$

To measure the seasonality of the reconstructed components, we determine the mean period s_1 of the function as follows:

$$s_1 = \sum_{i=1}^L n_i / (L - 1), \tag{11}$$

where L is the number of local maxima in the time series of the reconstructed component, and n_i is the number of observations between the i th and the $i + 1$ th maxima. To extract seasonal patterns with mean period s_1 , we use average seasonal dummies as follows:

$$p_{s,t} = \begin{cases} \beta_1 & t = 1, 1 + s_1, 1 + 2s_1, \dots \\ \beta_2 & t = 2, 2 + s_1, 2 + 2s_1, \dots \\ \dots & \\ \beta_{s_1-1} & t = s_1 - 1, 2s_1 - 1, 3s_1 - 1, \dots \end{cases}, \tag{12}$$

where $\beta_1, \beta_2, \dots, \beta_{s_1-1}$ denote average seasonal dummies.

From the above analysis, the data characteristic analysis in the proposed model can be summarized as follows: MPE is used for complexity tests of component reconstruction, and the dummies in Eq. (10) and Eq. (12) are used in a single model to capture the impacts of breakpoints and seasonality of each reconstructed component. Additionally, there may be outliers caused by one-off events. In order to capture the impacts of these events, one-off event dummies are used in the proposed model.

Forecasting models

For tourism demand forecasting, a recurrent network, Elman's neural network (ENN), is used. In an ENN model, feedback is the output of the hidden layers, so the network is especially suitable for fitting time series (Elman, 1990). An ENN with p input nodes, q hidden neurons, and o output nodes can be defined as follows:

$$S_j(t) = f \left(\sum_{z=1}^q \sum_{j=1}^q U_{jz} S_z(t-1) + \sum_{i=1}^p \sum_{j=1}^q V_{ji} X_i(t) \right), \tag{13}$$

$$O_{s_2}(t) = g \left(\sum_{j=1}^q \sum_{s_2=1}^o W_{s_2j} S_j(t) \right). \tag{14}$$

Here, $S_j(t)$ is the output of the j th hidden node ($j = 1, 2, \dots, q$), $X_i(t)$ is the input to the i th node ($i = 1, 2, \dots, p$), and $O_{s_2}(t)$ is the output ($s_2 = 1, 2, \dots, o$) at period t . Also, U_{jz} is the weight between the z th recurrent connection ($z = 1, 2, \dots, q$) and the j th hidden node, $S_z(t - 1)$ is the output of the z th recurrent connection at period $t - 1$, V_{ji} is the weight of the connection between the i th input node and the j th hidden neuron, and W_{s_2j} is the weight between the j th hidden neuron and the s_2 th output node; f and g are activation functions. The ENN model is supported by multiple studies (Liu, Tian, Liang, & Li, 2015; Ruiz, Rueda, Cuéllar, & Pegalajar, 2018;

Wang, Zhang, Li, Wang, & Dang, 2014), including research involving tourist arrival forecasting (Cho, 2003). In addition, for comparison purposes, the BPNN and the generalized regression neural network (GRNN) are used as single forecasting models.

Using the CEEMDAN method in conjunction with data characteristic analysis, we have developed a new decomposition-ensemble approach for tourism demand forecasting. After data decomposition and data characteristic analysis, each reconstructed component is predicted by using a single ENN model with dummy variables. The ensemble predictions \hat{x}_t without component reconstruction and those with component reconstruction are as follows:

$$\hat{x}_t = \sum_{k=1}^n \hat{c}_{kt} + \hat{r}_t, \tag{15}$$

$$\hat{x}_t = \sum_{k=1}^{n_r} \hat{c}_{kt}, \tag{16}$$

where n_r is the number of reconstructed components, at period t , \hat{c}_{kt} is the predicted value of the k th component, and \hat{r}_t is the predicted value of the residue.

Empirical study

In this section, for the purposes of illustration and verification, we describe our empirical study, which uses the time series of tourist arrivals from the top nine sources of people coming to Hong Kong. The nine sources are mainland China, Korea, Japan, the USA, the Philippines, Singapore, Australia, the UK and Thailand.

Data and experiment design

For this study, the monthly data of tourist arrivals at Hong Kong were collected from the WIND Database (<http://www.wind.com>).

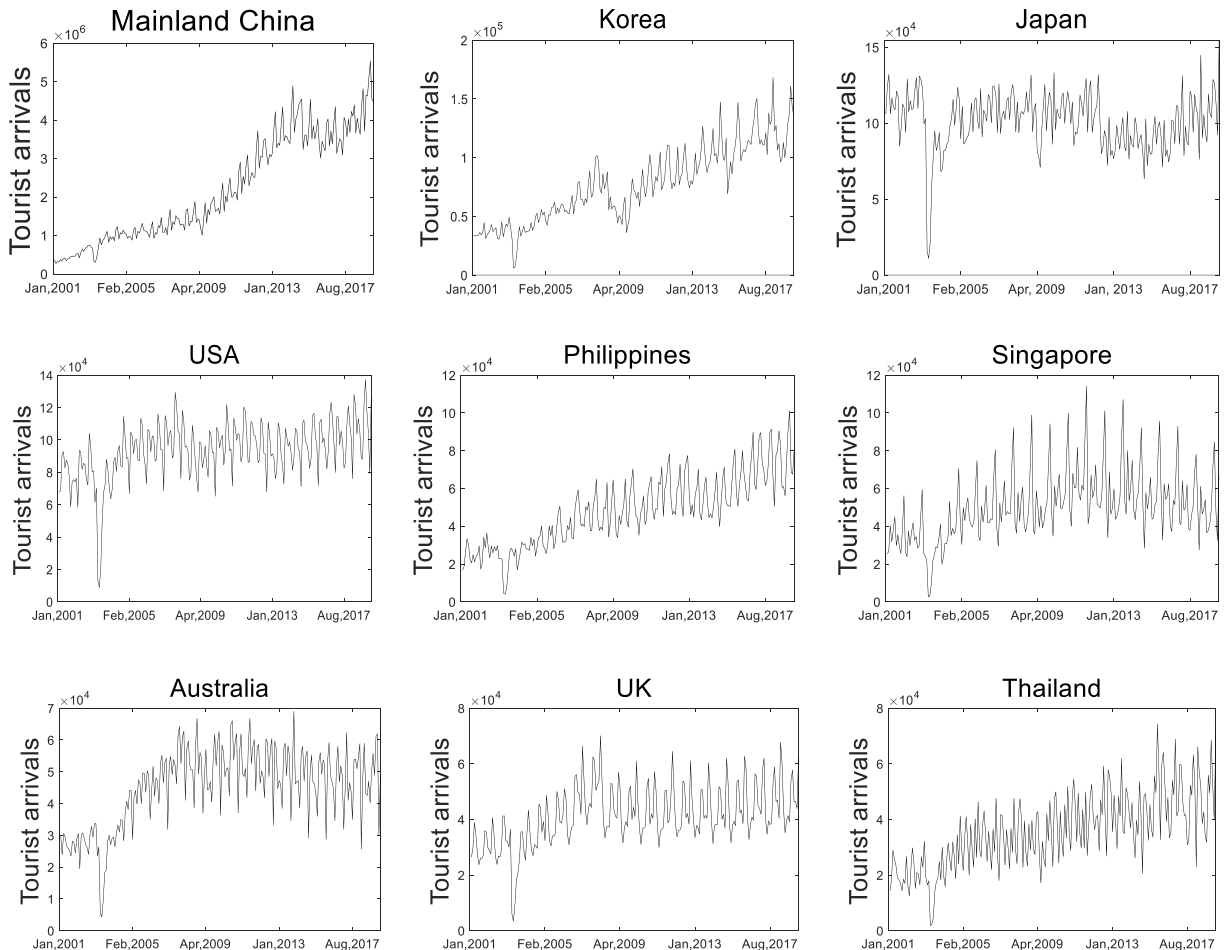


Fig. 2. Monthly tourist arrivals at Hong Kong.

cn/). As shown in Fig. 2, for each source market, the sample data consist of 219 observations, covering the period from January 2001, to March 2019. In the prediction of tourism demand, point forecasts usually do not provide information about the degree of variability or uncertainty related to the forecast, while interval forecasts can provide a range for the estimate with a specific confidence level. As a consequence, the interval provides more useful information for tourism practitioners and policymakers to formulate relevant strategies and policies (Wu, Song, & Shen, 2017). Referring the research in Li et al. (2019), this study produces both point and interval forecasts for tourist arrivals. In point forecasting, the training data sets are the first 80% (175 observations) of the tourist arrival time series from each source market (covering the period from January 2001, to July 2015), while the testing data sets are the latter 20% (44 observations, covering the period from August 2015, to March 2019).

In interval forecasting, the forecast density functions are assumed to follow normal distribution, with point forecasts as the means. Validation data sets (covering the period from August 2015, to May 2017) are used to calculate variances for density forecasts of June 2017, and an expanding data window is used for validation data sets to estimate variances for density forecasts until the variances of residuals (from August 2015 to February 2019) are calculated for the density forecasts of March 2019. Then, the testing data sets (covering the period from June 2017, to March 2019) are used to generate prediction intervals of individual models. In the prediction, we use an expanding data window, and the models are re-estimated. Specifically, the previous data of each out-of-sample observation are used as a training sample, and the models are re-estimated, in order to set the forecasting model for making h -step-ahead forecasts, where h is prediction horizon ($h = 1, 3$ and 6).

Six single models are used as benchmarks. They are naïve, ARIMA, BPNN, GRNN and ENN models, as well as an exponential smoothing model (ETS), i.e. the Holt-Winters model with multiplicative form. Also, five decomposition-ensemble models are used for comparison purposes: 1) a model based on seasonal-trend decomposition using a Loess and ENN (STL-ENN), 2) a model based on EMD and ENN (EMD-ENN), 3) a model based on EEMD and ENN (EEMD-ENN), 4) a model based on the CEEMDAN and ENN (CEEMDAN-ENN) and 5) the proposed model based on the CEEMDAN method, data characteristic analysis and ENN. In this study, naïve, ARIMA and ETS models are generated via the R package; other methods are implemented via a MATLAB software package.

Following the predictive accuracy measurement in Li et al. (2019), we evaluate the models in both point and interval forecasts. In point forecasting, mean absolute percentage error (MAPE) is used as the evaluation criteria:

$$MAPE = 100\% \times \sum_{t=1}^N |1 - \hat{x}_t/x_t|/N, \tag{17}$$

where N is the number of observations in the testing set. Because MAPE is the measure of the deviations between actual and predicted values, a prediction is more accurate when the MAPE value is smaller.

In contrast to point forecast error measurement, interval forecasting often uses coverage rate and interval width to measure predictive accuracy (Wu et al., 2017). However, the coverage rate and interval width may lead to controversial evaluation results (Li et al., 2019). Therefore, the Winkler score is used as a comprehensive measure for interval forecasting. The Winkler score at period t is defined as:

$$WS_t = \begin{cases} W_t & L_t \leq x_t \leq U_t \\ W_t + 2(L_t - x_t)/(1 - p) & L_t > x_t, \\ W_t + 2(x_t - U_t)/(1 - p) & U_t < x_t \end{cases} \tag{18}$$

where W_t is the width of an interval forecast at period t , $W_t = U_t - L_t$, L_t and U_t are the lower and upper limits, respectively, and p is the confidence level. In this study, 80% confidence level is used for each model and each source market. Then, the Winkler score for the interval forecasts is:

$$WS = \sum_{t=1}^N WS_t/N. \tag{19}$$

Just like the research in Li et al. (2019), interval forecasting accuracy is measured based on the natural logarithm form of the tourist arrival data. Similarly, the performance of interval forecasting is better when the Winkler score is lower.

The Diebold-Mariano (DM) statistic is further used to test the statistical significance of different models for point forecasting from a statistical perspective (Diebold & Mariano, 1995). In the DM test, the null hypothesis is that the degree of predictive accuracy across the different models is the same. Mean square prediction error (MSPE) was used as the loss function. The DM statistic S can be defined as:

$$S = \frac{\bar{g}}{\sqrt{\hat{V}_{\bar{g}}/N}}, \tag{20}$$

where $\bar{g} = \sum_{t=1}^N g_t/N$ ($g_t = (x_t - \hat{x}_{te,t})^2 - (x_t - \hat{x}_{re,t})^2$), $\hat{V}_{\bar{g}} = \gamma_0 + 2 \sum_{l=1}^{\infty} \gamma_l$, and $\gamma_l = \text{cov}(g_t, g_{t-l})$. Also, $\hat{x}_{te,t}$ and $\hat{x}_{re,t}$ are the predicted values by the tested te and the reference re methods, respectively, at period t .

Prediction results

The naïve model uses the latest available value as a forecast. The best ARIMA model with a log transformation was determined by analyzing the autocorrelation and partial correlation of the time series. The inputs of the ANN models, as well as the number of

Table 1
Parameters of single models.

Source	ARIMA	BPNN	GRNN	ENN
China	(2,1,0)(1,1,0)	N(3-4-1)	N(3-3-1)	N(3-4-1)
Korea	(2,1,1)(1,1,1)	N(4-4-1)	N(4-4-1)	N(4-4-1)
Japan	(1,0,0)(0,1,0)	N(3-4-1)	N(3-3-1)	N(3-4-1)
USA	(1,0,0)(1,1,0)	N(3-8-1)	N(3-3-1)	N(3-8-1)
Philippines	(3,1,1)(1,1,0)	N(4-10-1)	N(4-4-1)	N(4-10-1)
Singapore	(0,0,0)(1,1,0)	N(2-6-1)	N(2-2-1)	N(2-6-1)
Australia	(0,0,0)(1,1,1)	N(2-6-1)	N(2-2-1)	N(2-6-1)
UK	(0,0,0)(1,1,0)	N(2-4-1)	N(2-2-1)	N(2-4-1)
Thailand	(1,1,1)(1,1,0)	N(3-10-1)	N(3-3-1)	N(3-10-1)

hidden nodes, were determined by trial-and-error testing, in order to minimize the mean square error of training. As for the GRNN models, the number of hidden nodes are the same as the number of input dimensions. The parameters of the single models are listed in Table 1, where N(3-4-1) denotes a neural network with three input nodes, four hidden nodes, and one output neuron.

On the basis of CEEMDAN, the decomposed components of the tourist arrival numbers from China are shown in Fig. 3. In the calculation of the PE value, $m = 4$. Let $\theta = 0.5$. That is, a component with a PE value lower than 0.5 is low complexity, whereas a component with a PE value of no lower than 0.5 is high complexity. The decomposed components with PE values below the threshold value θ were combined into a new reconstructed component.

For each source market, most components remain in an unaltered state of complexity when the scale changes. Relatively few components may be assigned either high complexity or low complexity designations with various scales. For these types of components, the PE value increases at small scales. After reaching the maximum entropy, the PE value becomes stable as the scale factor increases. Related research indicates that PE values are unstable at large scales, because the coarse-graining procedure reduces the

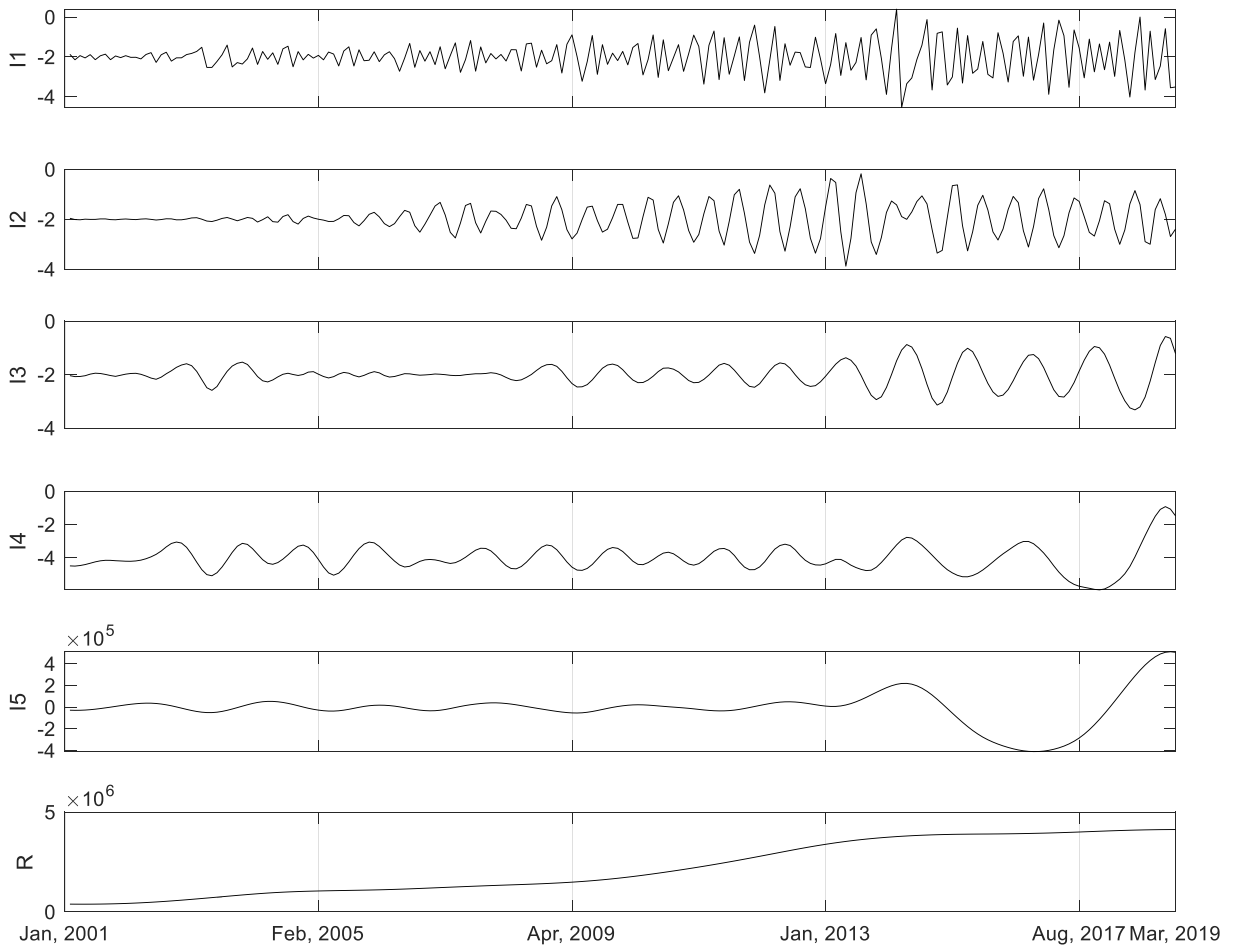


Fig. 3. The decomposed components of tourist arrivals from China.

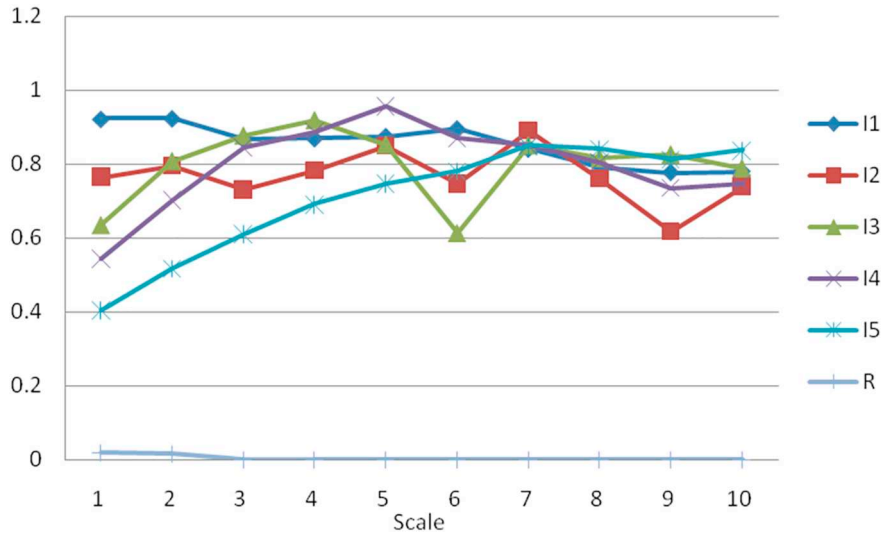


Fig. 4. PE value of components at 10 different scales (China).

sample size of templates (Park, Kim, Kim, Cichocki, & Kim, 2007). In order to allow every possible ordinal pattern of dimension m to occur in a time series with N observations, the condition $N \geq m!$ should be met. In Ouyang et al. (2013), to obtain a more accurate and reliable evaluation of MPE, only scale factors with $s \leq 5$ are considered. In this study, considering the sample size of empirical study, we also consider scale factors with $s \leq 5$.

Taking mainland China as an example, the PE values of decomposed components as a function of scale ($1 \leq s \leq 10$) are shown in Fig. 4, where I_i ($i = 1, 2, \dots, n$) and R are the i th intrinsic mode function and the residue, respectively. With different s , the first four components always have high complexity, and the last component always has low complexity. The fifth component, I_5 , has low complexity only when $s \leq 2$; it has high complexity when $s \geq 3$. As can be seen from Fig. 4, the PE values are more stable when $s = 4$ and 5 than when $s = 1-3$ because the slope of the PE curve decreases with s . Therefore, we select the PE values when $s = 4$ and 5 to evaluate the complexity of I_5 . In this way, we continue to evaluate the complexity of each component of other sources for purposes of component reconstruction.

The aim of component reconstruction is to reduce the computation cost. To save on computational costs, we evaluate the complexity of each decomposed component and then reconstruct the components. Table 2 lists the reconstructed components. For China, Korea, the Philippines, the UK and Thailand, there is no component reconstruction. For Japan, the USA, Singapore, Australia and the last two or three components are reconstructed as a new component.

The breakpoints of reconstructed components are shown in Table 3. Generally, breakpoints cause structure change in a long period, and their impacts are permanent. For any breakpoint, we added a dummy variable with values of 0 and 1, before and after the month of the breakpoint in the forecasting model.

For the time series of tourist arrivals from mainland China, the breakpoints in C_1 , C_2 and C_3 may have been caused by the occupation by protestors of central Hong Kong, middle east respiratory syndrome (MERS) and the adjustment of visa policy for Shenzhen residents, respectively. The breakpoint in C_4 may have been caused by appreciation of Hong Kong dollar and the effects of substitute destinations, while those in C_5 and C_6 should have been caused by some medium and long-term factors such as Hong Kong's reception capacity reaching a limit and economic slowdown in mainland China.

For the time series of tourist arrivals from Korea, the breakpoints in C_1 and C_2 occurred in 2015, the breakpoint in C_3 occurred in 2003, while those in C_4 , C_5 , C_6 and C_7 occurred in 2009. These may have been caused by the MERS, the SARS epidemic in Hong Kong and the global financial crisis of 2008, respectively.

Table 2
Reconstructed components of tourist arrival time series.

Source	C_1	C_2	C_3	C_4	C_5	C_6	C_7
China	I_1	I_2	I_3	I_4	I_5	R	-
Korea	I_1	I_2	I_3	I_4	I_5	I_6	R
Japan	I_1	I_2	I_3	I_4	I_5	I_6 & I_7 & R	-
USA	I_1	I_2	I_3	I_4	I_5	I_6 & I_7 & R	-
Philippines	I_1	I_2	I_3	I_4	I_5	R	-
Singapore	I_1	I_2	I_3	I_4	I_5	I_6 & R	-
Australia	I_1	I_2	I_3	I_4	I_5	I_6 & R	-
UK	I_1	I_2	I_3	I_4	I_5	I_6	R
Thailand	I_1	I_2	I_3	I_4	I_5	R	-

Table 3
Breakpoints of reconstructed components.

Source	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇
China	2014.11	2015.02	2015.04	2016.08	2014.05	2015.06	–
Korea	2015.04	2015.05	2003.05	2009.02	2009.05	2009.07	2009.11
Japan	2011.04	2003.05	2003.05	2003.06	2003.08	2009.09	–
USA	2015.10	2003.04	2003.05	2003.04	2003.04	2002.09	–
Philippines	2013.12	2008.12	2003.05	2003.05	2003.04	2012.04	–
Singapore	2008.09	2003.06	2003.04	2009.07	2003.08	2008.05	–
Australia	2003.05	2015.04	2003.04	2003.04	2003.06	2008.10	–
UK	2003.04	2009.09	2003.04	2003.07	2008.07	2008.10	2008.09
Thailand	2003.07	2012.05	2015.04	2014.06	2014.08	–	–

In terms of tourist arrivals from Japan, the breakpoint in C₁ in 2011 may have been caused by the east Japan earthquake in 2011; the breakpoints in C₂, C₃, C₄ and C₅ in 2003 should have been caused by the SARS epidemic in Hong Kong; the breakpoint of C₆ may have been caused by the global financial crisis of 2008.

For the time series of tourist arrivals from the USA, the breakpoint in C₁, occurred in 2015, may have been caused by MERS in Hong Kong; the breakpoints in C₂, C₃, C₄ and C₅ may have been caused by the SARS epidemic in Hong Kong; the breakpoint in C₆, occurred in 2002, may have been caused by the economic recovery after the 911 Incident.

For the time series of tourist arrivals from the Philippines, the breakpoints in C₁ and C₂ may have been caused by the super typhoon in 2013 and the global financial crisis in 2008, respectively. The breakpoints in C₃, C₄ and C₅, occurred in 2003, should have been caused by the SARS epidemic in Hong Kong, while the breakpoint in C₆ may have been caused by the rapid growth of Philippine economy since 2012.

For the time series of tourist arrivals from Singapore, the breakpoints in C₁, C₄ and C₆ may have been caused by the global financial crisis, while those in C₂, C₃ and C₅ should have been caused by the SARS epidemic in Hong Kong.

For tourist arrivals from Australia, the breakpoints in C₁, C₃, C₄ and C₅, occurred in 2003, may have been caused by the SARS epidemic in Hong Kong; the breakpoint in C₂ may have been caused by MERS in Hong Kong in 2015; the breakpoints in C₆, occurred in 2008, should have been caused by the global financial crisis.

For tourist arrivals from the UK, the breakpoint in C₁, C₃ and C₄, occurred in 2003, should be caused by the SARS epidemic in Hong Kong, while those in C₂, C₅, C₆ and C₇ may have been caused by the global financial crisis.

Finally, for tourist arrivals from Thailand, the breakpoints in C₁ may have been caused by the SARS epidemic in Hong Kong; the breakpoints in C₂, C₄ and C₅ may have been caused by the turbulent political situation in Thailand; the breakpoint in C₃ may have been caused by MERS in Hong Kong.

Besides breakpoints, there may be outliers caused by one-off events. We use function `boxplot.stats()` in R software to find out outliers and analyze what events cause the outliers. In order to reflect the effects of these events, we added one-off event dummies with values of 1 and 0 for outliers and other points, respectively.

The seasonality of reconstructed components for different source markets is shown in Table 4. For each source, the components are from high to low frequency. Taking mainland China as an example, the main time scales of C₁ and C₂ are 3 and 6 months, respectively, while those of C₃, C₄ and C₅ are 12, 24 and 55 months, respectively. No obvious seasonality can be found in C₆. In the prediction, the average seasonal dummies p_s, τ in Eq. (12) are used in a single ENN model, in order to extract seasonal patterns in the time series of each reconstructed component.

The MAPE values for point forecasts and Winkler scores for interval forecasts (80% confidence level) are listed in Table 5 and Table 6, respectively. Additionally, the DM statistics of out-of-sample datasets from different source markets are calculated with different prediction horizons, as shown in Table 7. After implementing the proposed approach by conducting the above-mentioned experiments, we continue to discuss a number of related issues in the next section.

Table 4
Main time scale of reconstructed components.

Source	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇
China	3.00	5.92	12.17	24.33	54.75	219.00	
Korea	3.00	6.08	12.17	24.33	36.50	109.50	219.00
Japan	3.00	6.08	11.53	21.90	36.50	109.50	
USA	3.00	6.08	9.13	15.64	36.50	109.50	
Philippines	6.08	6.08	12.17	21.90	73.00	219.00	
Singapore	3.00	5.92	12.17	36.50	73.00	219.00	
Australia	3.00	5.92	12.17	36.50	73.00	219.00	
UK	3.98	6.08	12.17	21.90	36.50	73.00	109.50
Thailand	3.98	5.92	12.17	43.80	73.00	219.00	

Table 5
MAPE values for point forecasts (%).

Source	<i>h</i>	Proposed model	CEEMDAN-ENN	EEMD-ENN	EMD-ENN	STL-ENN	ETS	ENN	GRNN	BPNN	ARIMA	Naïve
China	1	2.69	3.49	4.09	5.14	5.15	5.31	6.28	6.30	6.45	6.67	11.79
	3	3.50	4.19	5.20	5.38	5.40	5.89	8.83	8.78	9.34	13.08	11.07
	6	4.52	5.77	6.47	6.75	6.60	7.54	10.09	9.92	11.97	15.54	10.94
Korea	1	3.30	4.85	5.65	6.57	5.64	8.62	8.02	8.37	9.04	8.60	10.69
	3	4.11	5.65	7.16	7.50	7.75	10.89	10.19	10.42	10.70	18.55	15.79
	6	4.72	6.34	7.48	8.02	9.56	12.16	10.60	10.83	12.92	15.32	19.76
Japan	1	3.83	5.37	5.59	6.60	6.34	10.24	8.03	8.35	8.48	9.19	18.70
	3	4.20	5.69	7.22	8.86	7.05	13.16	9.24	9.49	10.26	20.67	16.25
	6	4.83	6.72	8.17	9.10	7.82	13.94	11.23	11.46	9.87	19.82	18.98
USA	1	1.55	2.47	2.84	3.32	4.43	4.90	3.67	4.34	4.74	3.74	15.11
	3	1.80	2.78	3.46	4.68	4.68	6.74	5.32	5.95	7.09	18.81	21.69
	6	2.44	3.32	4.37	5.52	6.43	6.41	4.89	5.11	6.01	22.91	11.97
Philippines	1	3.36	4.58	5.34	5.61	7.31	6.36	7.21	7.29	7.73	8.70	15.68
	3	3.78	4.97	5.73	6.31	7.75	6.84	8.78	8.13	9.68	20.32	27.38
	6	4.13	5.52	6.26	6.89	8.76	7.30	8.34	8.42	8.89	23.55	12.75
Singapore	1	3.16	4.70	5.91	6.66	7.30	7.46	8.32	8.86	8.96	9.31	29.92
	3	3.73	4.93	6.42	7.55	7.74	8.11	9.58	9.27	9.09	37.62	36.66
	6	4.41	5.75	6.74	8.18	7.90	8.59	10.03	9.50	10.10	38.32	25.28
Australia	1	2.37	3.72	4.46	4.97	5.49	7.40	5.77	5.94	5.88	6.13	24.90
	3	2.86	4.04	4.73	5.45	6.54	8.10	7.99	8.11	8.64	25.57	22.69
	6	3.40	4.57	5.23	5.74	7.23	8.40	8.36	8.41	9.18	25.71	13.96
UK	1	1.52	2.28	2.61	2.88	3.73	4.17	3.94	4.28	4.11	4.30	16.25
	3	1.94	2.47	2.88	3.18	3.83	4.50	4.34	4.42	4.20	23.27	27.43
	6	2.10	2.65	3.08	3.34	5.49	5.88	4.76	4.54	5.06	24.29	22.92
Thailand	1	4.13	6.60	8.29	9.31	11.53	12.09	10.37	10.71	11.04	12.43	28.00
	3	4.78	7.36	8.70	10.12	13.08	13.64	14.53	14.10	15.37	41.57	31.03
	6	5.03	7.72	9.26	10.55	13.58	15.54	13.72	14.15	15.87	33.12	30.68

Table 6
Winkler scores for interval forecasts (80% confidence level).

Source	<i>h</i>	Proposed model	CEEMDAN-ENN	EEMD-ENN	EMD-ENN	STL-ENN	ETS	ENN	GRNN	BPNN	ARIMA	Naïve
China	1	0.103	0.144	0.167	0.244	0.247	0.251	0.274	0.276	0.315	0.314	0.471
	3	0.161	0.197	0.283	0.317	0.242	0.261	0.493	0.495	0.517	0.544	0.443
	6	0.221	0.297	0.295	0.272	0.341	0.309	0.470	0.540	0.532	0.760	0.579
Korea	1	0.167	0.223	0.214	0.255	0.294	0.343	0.428	0.422	0.470	0.382	0.502
	3	0.180	0.231	0.253	0.285	0.375	0.365	0.496	0.489	0.507	0.732	0.684
	6	0.198	0.244	0.267	0.355	0.416	0.445	0.523	0.515	0.535	0.610	0.861
Japan	1	0.223	0.256	0.259	0.272	0.349	0.458	0.461	0.487	0.530	0.549	0.962
	3	0.243	0.267	0.279	0.381	0.372	0.538	0.503	0.511	0.535	0.725	0.860
	6	0.266	0.281	0.297	0.396	0.378	0.564	0.522	0.529	0.542	0.897	0.760
USA	1	0.080	0.099	0.114	0.135	0.195	0.212	0.215	0.207	0.206	0.182	0.787
	3	0.104	0.116	0.159	0.152	0.247	0.263	0.257	0.261	0.289	0.741	0.879
	6	0.132	0.166	0.233	0.193	0.373	0.284	0.282	0.295	0.314	0.866	0.724
Philippines	1	0.151	0.178	0.217	0.227	0.433	0.283	0.301	0.323	0.348	0.464	0.742
	3	0.186	0.205	0.248	0.241	0.454	0.301	0.412	0.427	0.408	1.081	1.035
	6	0.218	0.233	0.260	0.289	0.479	0.322	0.445	0.475	0.454	0.930	0.539
Singapore	1	0.239	0.252	0.298	0.319	0.434	0.340	0.487	0.485	0.497	0.451	1.310
	3	0.256	0.271	0.315	0.349	0.395	0.317	0.515	0.506	0.512	1.406	1.520
	6	0.287	0.315	0.356	0.436	0.691	0.377	0.490	0.513	0.475	1.544	0.853
Australia	1	0.148	0.151	0.164	0.172	0.256	0.319	0.224	0.267	0.283	0.285	1.281
	3	0.159	0.178	0.191	0.205	0.295	0.317	0.377	0.377	0.421	1.160	0.996
	6	0.172	0.193	0.216	0.234	0.353	0.355	0.382	0.444	0.396	1.336	0.795
UK	1	0.068	0.075	0.111	0.145	0.164	0.188	0.172	0.196	0.190	0.185	0.754
	3	0.077	0.091	0.158	0.183	0.184	0.215	0.226	0.232	0.245	0.905	1.004
	6	0.081	0.127	0.198	0.218	0.296	0.307	0.277	0.295	0.282	0.933	0.973
Thailand	1	0.293	0.328	0.369	0.411	0.647	0.597	0.531	0.674	0.603	0.710	1.461
	3	0.322	0.365	0.396	0.463	0.674	0.580	0.633	0.698	0.726	1.508	1.465
	6	0.364	0.393	0.432	0.497	0.687	0.678	0.694	0.734	0.779	1.557	1.428

Discussion

Based on the experimental results given in “Empirical study” section, we first analyze the forecasting performance of the models. Then, a summary regarding tourism demand forecasting is made, according to the above analysis.

Table 7
DM test results of out-of-sample datasets from different source markets.

Source	h	Reference model									
		CEEMDAN-ENN	EEMD-ENN	EMD-ENN	STL-ENN	ETS	ENN	GRNN	BPNN	ARIMA	Naïve
China	1	-2.322**	-3.334***	-3.606***	-2.946***	-3.419***	-3.584***	-3.582***	-4.269***	-3.012***	-5.094***
	3	-2.039**	-2.255**	-2.561***	-2.615***	-2.516***	-2.331**	-2.281**	-2.602***	-10.84***	-8.754***
	6	-1.785**	-1.981**	-2.618***	-1.974**	-2.071**	-2.044**	-1.742**	-3.012***	-4.691***	-3.689***
Korea	1	-2.984***	-3.979***	-4.574***	-2.726***	-4.861***	-3.186***	-3.345***	-2.92**	-4.512***	-3.937***
	3	-2.548***	-2.481***	-3.159***	-3.493***	-3.443***	-5.915***	-5.221***	-3.486***	-5.609***	-5.165***
	6	-1.828**	-2.523***	-2.643***	-4.959***	-4.075***	-5.063***	-6.247***	-3.810***	-4.678***	-19.19***
Japan	1	-3.107***	-3.414***	-3.444***	-2.518***	-3.96***	-4.336***	-4.396***	-4.409***	-4.594***	-3.835***
	3	-2.912***	-3.282***	-3.421***	-2.713***	-5.445***	-2.762***	-3.143***	-3.335***	-8.025***	-4.414***
	6	-2.768***	-3.166***	-3.389***	-2.928***	-3.989***	-3.436***	-3.551***	-3.574***	-5.294***	-4.663***
USA	1	-2.165**	-2.491***	-3.019***	-4.456***	-4.829***	-2.681***	-3.349***	-4.019***	-3.151***	-4.312***
	3	-1.826**	-2.338**	-2.434***	-3.171***	-4.518***	-3.816***	-3.967***	-3.237***	-15.52***	-7.767***
	6	-1.604*	-1.886*	-1.931**	-2.571***	-5.534***	-2.279**	-2.757***	-3.255***	-25.56***	-10.84***
Philippines	1	-2.536**	-3.502***	-3.708***	-3.211***	-3.781***	-3.639***	-3.921***	-3.593***	-4.005***	-4.843***
	3	-2.506**	-2.737**	-3.818***	-2.421***	-3.958***	-3.473***	-2.768***	-5.035***	-6.854***	-29.74***
	6	-2.091**	-2.577**	-3.472***	-4.715***	-5.737***	-5.086***	-5.021***	-3.351***	-5.663***	-10.72***
Singapore	1	-3.438***	-3.524***	-4.165***	-3.877***	-3.645***	-3.377***	-3.588***	-4.264***	-3.659***	-4.841***
	3	-3.279***	-3.309***	-4.062***	-3.587***	-3.306***	-3.207***	-3.395***	-4.176***	-4.445***	-4.851***
	6	-3.089***	-3.235***	-3.820***	-3.251***	-3.896***	-3.104***	-3.217***	-3.679***	-9.829***	-5.851***
Australia	1	-2.490***	-3.133***	-3.025***	-2.817***	-4.107***	-4.541***	-4.427***	-4.071***	-4.309***	-4.531***
	3	-2.224**	-2.946***	-2.686***	-2.854***	-5.198***	-4.587***	-4.719***	-5.027***	-5.159***	-6.025***
	6	-1.726**	-1.809**	-2.252**	-2.179**	-3.349***	-5.552***	-1.916**	-3.434***	-4.467***	-2.475***
UK	1	-2.933***	-3.251***	-3.282***	-3.332***	-4.330***	-3.421***	-3.668***	-3.576***	-3.685***	-4.384***
	3	-2.027**	-2.158**	-2.205**	-2.319**	-2.622***	-4.527***	-3.022***	-2.562***	-8.613***	-14.45***
	6	-1.806**	-2.057**	-2.164**	-2.075**	-3.340***	-3.064***	-2.462***	-3.868***	-5.385***	-5.144***
Thailand	1	-4.540***	-3.628***	-4.247***	-4.071***	-4.503***	-4.391***	-4.545***	-4.799***	-4.631***	-4.876***
	3	-3.527***	-4.038***	-5.053***	-4.812***	-5.036***	-3.871***	-3.390***	-3.943***	-6.305***	-7.670***
	6	-3.312***	-3.164***	-4.258***	-4.902***	-4.149***	-3.531***	-3.685***	-3.468***	-5.387***	-4.834***

* A rejection of null hypothesis at the 10% significance level.

** A rejection of null hypothesis at the 5% significance level.

*** A rejection of null hypothesis at the 1% significance level.

Analysis of prediction results

According to the results of the nine sources in Table 5 and Table 6, for both point and interval forecasts, the ARIMA models have higher predictive accuracy than naïve models in one-step-ahead prediction. Meanwhile, the ANN models (BPNN, GRNN and ENN) perform better than the ARIMA models in most cases. The ENN model shows greater stability and better forecasting performance than BPNN and GRNN models. The ARIMA model can achieve higher predictive accuracy in the USA and UK time series than in other sources. The reason for this finding may be that the time series of these two sources are more stable and less complex than those of other sources. Comparatively, the MAPE values and Winkler scores of the ARIMA model in Thailand are relatively high. The reason for this finding maybe that the turbulent political situation in Thailand makes the conditions encountered travel abroad more volatile.

Compared with the STL-ENN and single models, the EMD-based models (EMD-ENN, EEMD-ENN and CEEMDAN-ENN) can all achieve higher predictive accuracy in point forecasting. Similarly, the EMD-based models outperform the STL-ENN and single models in interval forecasting in most cases. From the measurement criteria for the forecasting models and the different source markets, the proposed model can achieve the most accurate prediction results in both point and interval forecasting with all different prediction horizons ($h = 1, 3, 6$). The tourism demand forecasting applications relating to tourist arrivals from the nine sources used in this study demonstrate the effectiveness and stability of the proposed decomposition-ensemble approach.

From the DM test results of the out-of-sample datasets from different source markets in Table 7, it is clear that all the DM statistics from the proposed model are less than -1.604 , which corresponds to p -values of less than 0.10. This result clearly indicates that the proposed model significantly outperforms other models, at a 90% confidence level. Except for the CEEMDAN-ENN model and the USA, the DM statistics from the proposed model are less than -1.726 , which corresponds to p -values of less than 0.05. These results indicate that the proposed model significantly outperformed these reference models, at a 95% confidence level. In particular, the proposed model performs better for Singapore and Thailand than for other countries of origin. The reason may be that the time series of tourist arrivals from these two origin countries are more non-stationary and complex, which can be processed well by the proposed model. On the whole, the DM test results suggest that the proposed model can be more capable than other models used for forecasting tourism demand.

Summary

Through a comparison and analysis of predictive accuracy, we can provide a summary of tourism demand forecasting, as follows:

First, the proposed approach improves both point and interval forecasting performance. The results show that the CEEMDAN method is an effective data decomposition technique for tourism demand forecasting, especially for more non-stationary and complex time series of tourist arrivals.

Secondly, due to the nonlinearity and complexity of tourist arrivals, ANN models can achieve higher predictive accuracy than ARIMA models in most cases. Compared to BPNN and GRNN models, the ENN model exhibits greater stability and can usually achieve better forecasting performance.

Multiple scales may exist in the time series of some decomposed components. As a consequence, MPE can be more competent than can single-scale permutation entropy for measuring the complexity of the relevant components.

Finally, the forecasting performance of a model is closely related to the data characteristics of time series. It is helpful to capture the data characteristics for the prediction of reconstructed components. The mutability and seasonality can be effectively described by using dummy variables.

Conclusions

To enhance the predictive accuracy of tourism demand forecasting, a decomposition-ensemble approach is developed. The procedure of the proposed approach consists of four main steps: data decomposition, component reconstruction, individual prediction, and ensemble prediction. The empirical results indicate that the proposed model outperforms other models in both point and interval forecasts for different prediction horizons, especially for more non-stationary and complex time series of tourist arrivals.

For the proposed model, the contributions for tourism demand forecasting are twofold: (1) Previous decomposition techniques for tourism demand forecasting neglect data characteristics of decomposed components, and focus only on point forecasting. Accordingly, this study proposes a new methodology, which is the first to investigate how to enhance both point and interval forecasting performance by combining the CEEMDAN method and data characteristic analysis. (2) New techniques are introduced to decompose sample data and capture the data characteristics of tourism demand. Compared with the original EMD method, the CEEMDAN method reduces the occurrence of mode mixing. As multiple scales may exist in the time series of some decomposed components, this study uses multiscale permutation entropy, rather than the traditional single-scale permutation entropy, as a new efficient technique of component reconstruction. It also investigates the mutability and seasonality of reconstructed components, and analyzes the impacts of irregular events on tourism demand. Thus, it is possible to develop more effective forecasting models, using dummy variables, to predict reconstructed components.

The accurate point and interval forecasts of tourist arrival numbers are critical for all parties in tourism industry. With accurately predicted sizes and directions of future tourist flows, the government can generate a more appropriate tourism strategy and provide a better service to tourists. Accurate forecasts can also benefit industry practitioners for their operations management, including investment, pricing, advertising, promotional activities and revenue management. Particularly, accurate interval forecasts of tourism demand can help to make more comprehensive evidence-based decisions. After knowing the different prediction intervals of tourism demand with different confidence levels, decision makers can develop corresponding plans to deal with uncertainties and risks.

Although the proposed model has achieved a superior standard of forecasting performance, they still have some limitations. First, the decomposition methods mainly focus on the EMD and its extended methods. Other decomposition methods, such as SSA and wavelet packet decomposition, can also be used for tourism forecast purposes. Second, the aim of this study is to validate the effectiveness of the proposed decomposition-ensemble approach, where the ENN is chosen for the prediction of each reconstructed component and exhibits high predictive accuracy. There are various other recurrent neural networks, such as LSTM, which are often found to perform well in time series forecasting. New models can be further developed based on decomposition methods, data characteristic analysis and LSTM. Finally, only MPE is used for complexity tests of component reconstruction. Other component reconstruction methods, including fine-to-coarse techniques, K-means clustering and phase space reconstruction methods, are worthy of investigation in the future.

Statement of contribution

1. What is the contribution to knowledge, theory, policy or practice offered by the paper?

The contribution of this study is that both the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and data characteristic analysis (DCA) methods are used to enhance the degree of prediction accuracy in tourism forecast. New hybrid models are developed to compare with the benchmark models. The results show that the CEEMDAN method is an effective data decomposition technique, and DCA is helpful for tourism demand forecasting. Our investigation has found that multiple time scales may exist in the time series of decomposed components. Hence, multiple scale permutation entropy (MPE) can be more competent than single-scale permutation entropy for measuring the complexity of components with multiple time scales. Moreover, there may be mutability and cyclicity (seasonality) for the reconstructed components. Therefore, dummy variables can be used to capture these data characteristics, in order to achieve higher predictive accuracy.

2. How does the paper offer a social science perspective/approach?

Tourism demand forecasting is important to tourism management. With accurate predictions of future tourist volumes, the government could formulate a more appropriate development strategy and provide better service to tourists. Simultaneously, accurate forecasts could help the private sector develop more suitable marketing strategies, thus generating greater revenue. However, developing effective models for forecasting tourism demand is still a challenge. Our approach actually improves the existing EMD-based models within the “divide and conquer” framework. Due to the complexity of time series of tourist arrivals, a

decomposition-ensemble approach is developed for forecasting tourism demand. Four main steps are involved: (1) Sample data of tourist arrivals are decomposed into several components by using the CEEMDAN technique. (2) The components are reconstructed based on MPE. (3) A neural network model is used for the prediction of each reconstructed component, according to DCA. (4) The forecast results are combined as an aggregated output.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 71771207, 71642006), and the National Center for Mathematics and Interdisciplinary Sciences, CAS.

References

- Akın, M. (2015). A novel approach to model selection in tourism demand modeling. *Tourism Management*, 48, 64–72.
- Andrawis, R. R., Atiya, A. F., & El-shishiny, H. (2011). Combination of long term and short term forecasts, with application to tourism demand forecasting. *International Journal of Forecasting*, 27(3), 870–886.
- Aslanargun, A., Mammadov, M., Yazici, B., & Yolacan, S. (2007). Comparison of ARIMA, neural networks and hybrid models in time series: Tourist arrival forecasting. *Journal of Statistical Computation and Simulation*, 77(1), 29–53.
- Athanasopoulos, G., Hyndman, R. J., Song, H., & Wu, D. C. (2011). The tourism forecasting competition. *International Journal of Forecasting*, 27(3), 822–844.
- Bandt, C., Keller, G., & Pompe, B. (2002). Entropy of interval maps via permutations. *Nonlinearity*, 15, 1595–1602.
- Bandt, C., & Pompe, B. (2002). Permutation entropy: A natural complexity measure for time series. *Physical Review Letters*, 88, 174102.
- Burger, C. J. S. C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting - a case study of Durban, South Africa. *Tourism Management*, 22, 403–409.
- Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 519, 127–139.
- Chatziantoniou, I., Degiannakis, S., Eeckels, B., & Filis, G. (2016). Forecasting tourist arrivals using origin country macroeconomics. *Applied Economics*, 48(27), 2571–2585.
- Chen, C.-F., Lai, M.-C., & Yeh, C.-C. (2012). Forecasting tourism demand based on empirical mode decomposition and neural network. *Knowledge Based Systems*, 26, 281–287.
- Chen, J. L., Li, G., Wu, D. C., & Shen, S. (2019). Forecasting seasonal tourism demand using a multiseries structural time series method. *Journal of Travel Research*, 58(1), 92–103.
- Chen, R., Liang, C.-Y., Hong, W.-C., & Gu, D.-X. (2015). Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm. *Applied Soft Computing*, 26, 435–443.
- Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management*, 24, 323–330.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28, 591–605.
- Chu, F.-L. (2011). A piecewise linear approach to modeling and forecasting demand for Macau tourism. *Tourism Management*, 32, 1414–1420.
- Colominas, M. A., Schlotthauer, G., & Torres, M. E. (2014). Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing & Control*, 14, 19–29.
- Costa, M., Goldberger, A., & Peng, C. (2002). Multiscale entropy analysis of physiologic time series. *Physical Review Letters*, 89, 062102.
- Costa, M., Goldberger, A., & Peng, C. (2005). Multiscale entropy analysis of biological signal. *Physical Review E*, 71, 021906.
- Costa, M., Peng, C. K., Goldberger, A. L., & Hausdorff, J. M. (2003). Multiscale entropy analysis of human gait dynamics. *Physica A*, 330, 53–60.
- Cró, S., & Martins, A. M. (2017). Structural breaks in international tourism demand: Are they caused by crises or disasters? *Tourism Management*, 63, 3–9.
- Dai, S., Niu, D., & Li, Y. (2018). Daily peak load forecasting based on complete ensemble empirical mode decomposition with adaptive noise and support vector machine optimized by modified grey wolf optimization algorithm. *Energies*, 11(1), 163. <https://doi.org/10.3390/en11010163>.
- Dharmaratne, G. S. (1995). Forecasting tourist arrivals in Barbados. *Annals of Tourism Research*, 22(4), 804–818.
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Gao, W., Zhang, Q., Lu, Z., Wu, D., & Du, X. (2018). Modelling and application of fuzzy adaptive minimum spanning tree in tourism agglomeration area division. *Knowledge Based Systems*, 143, 317–326.
- Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23, 499–510.
- Greenidge, K. (2001). Forecasting tourism demand: An STM approach. *Annals of Tourism Research*, 28(1), 98–112.
- Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., & Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, 63, 112–127.
- Hassani, H., Webster, A., Silva, E. S., & Heravi, S. (2015). Forecasting U.S. tourist arrivals using optimal singular Spectrum analysis. *Tourism Management*, 46, 322–335.
- Hong, W.-C., Dong, Y., Chen, L.-Y., & Wei, S.-Y. (2011). SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Applied Soft Computing*, 11, 1881–1890.
- Huang, N. E., Shen, Z., & Long, S. R. (1971). (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proceedings: Mathematical. Physical and Engineering Sciences*, 454, 903–995.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. 2nd ed., OTexts. Available via: <https://otexts.org/fpp2/> (Accessed 25.05.2019).
- Kon, S. C., & Turner, L. W. (2005). Neural network forecasting of tourism demand. *Tourism Economics*, 11(3), 301–328.
- Kulendran, N., & Shan, J. (2002). Forecasting China's monthly inbound travel demand. *Journal of Travel & Tourism Marketing*, 13, 5–19.
- Law, R., & Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, 20, 89–97.
- Li, G., Wu, D. C., Zhou, M., & Liu, A. (2019). The combination of interval forecasts in tourism. *Annals of Tourism Research*, 75, 363–378.
- Lim, C., & McAleer, M. (2001). Forecasting tourist arrivals. *Annals of Tourism Research*, 28(4), 965–977.
- Liu, H., Mi, X., & Li, Y. (2018). Comparison of two new intelligent wind speed forecasting approaches based on wavelet packet decomposition, complete ensemble empirical mode decomposition with adaptive noise and artificial neural networks. *Energy Conversion and Management*, 155, 188–200.
- Liu, H., Tian, H., Liang, X., & Li, Y. (2015). Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. *Applied Energy*, 157, 183–194.
- Liu, Q., Wei, Q., Fan, S.-Z., Lu, C.-W., Lin, T.-Y., Abbod, M. F., & Shieh, J.-S. (2012). Adaptive computation of multiscale entropy and its application in EEG signals for monitoring depth of anesthesia during surgery. *Entropy*, 14(6), 978–992.
- Ma, E., Liu, Y., Li, J., & Chen, S. (2016). Anticipating Chinese tourist arrivals in Australia: A time series analysis. *Tourism Management Perspectives*, 17, 50–58.
- Narayan, P. K. (2005). The structure of tourist expenditure in Fiji: Evidence from unit root structural break tests. *Applied Economics*, 37(10), 1157–1161.
- Ouyang, G., Li, J., Liu, X., & Li, X. (2013). Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis. *Epilepsy Research*, 104, 246–252.
- Park, J. H., Kim, S., Kim, C. H., Cichocki, A., & Kim, K. S. (2007). Multiscale entropy analysis of EEG from patients under different pathological conditions. *Fractals*, 15, 399–404.
- Peng, B., Song, H., & Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tourism Management*, 45, 181–193.

- Richman, J., & Moorman, J. (2000). Physiological time series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory*, 278, 2039–2049.
- Rosso, O., Larrondo, H. A., Martin, M. T., Plastino, A., & Fuentes, M. A. (2007). Distinguishing noise from chaos. *Physical Review Letters*, 99, 154102.
- Ruiz, L. G. B., Rueda, R., Cuéllar, M. P., & Pegalajar, M. C. (2018). Energy consumption forecasting based on Elman neural networks with evolutive optimization. *Expert Systems with Applications*, 92, 380–389.
- Shahrabi, J., Hadavandi, E., & Asadi, S. (2013). Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series. *Knowledge Based Systems*, 43, 112–122.
- Silva, E. S., Hassani, H., Heravi, S., & Huang, X. (2019). Forecasting tourism demand with denoised neural networks. *Annals of Tourism Research*, 74, 134–154.
- Song, H., Li, G., Witt, S. F., & Athanasopoulos, G. (2011). Forecasting tourist arrivals using time-varying parameter structural time series models. *International Journal of Forecasting*, 27, 855–869.
- Song, H., Qiu, R. T. R., & Park, J. (2019). A review of research on tourism demand forecasting: Launching the annals of tourism research curated collection on tourism demand forecasting. *Annals of Tourism Research*, 75, 338–362.
- Song, H., Wen, L., & Liu, C. (2019). Density tourism demand forecasting revisited. *Annals of Tourism Research*, 75, 379–392.
- Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tourism Management*, 27, 214–224.
- Song, H., Witt, S. F., & Jensen, T. C. (2003). Tourism forecasting: Accuracy of alternative econometric models. *International Journal of Forecasting*, 19, 123–141.
- Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 125, 4144–4147.
- Wang, C., Zhang, H., Fan, W., & Ma, P. (2017). A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction. *Energy*, 138, 977–990.
- Wang, J., Zhang, W., Li, Y., Wang, J., & Dang, Z. (2014). Forecasting wind speed using empirical mode decomposition and Elman neural network. *Applied Soft Computing*, 23, 452–459.
- Witt, S. F., & Witt, C. A. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, 11(3), 447–475.
- Wu, D. C., Song, H., & Shen, S. (2017). New developments in tourism and hotel demand modeling and forecasting. *International Journal of Contemporary Hospitality Management*, 29(1), 507–529.
- Wu, L., & Cao, G. (2016). Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow. *Knowledge Based Systems*, 110, 157–166.
- Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1), 1–41.
- Yu, L., Wang, Z., & Tang, L. (2015). A decomposition-ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. *Applied Energy*, 156, 251–267.
- Zhang, G., Wu, J., Pan, B., Li, J., Ma, M., Zhang, M., & Wang, J. (2017). Improving daily occupancy forecasting accuracy for hotels based on EEMD-ARIMA model. *Tourism Economics*, 23(7), 1496–1514.
- Zhang, W., Qu, Z., Zhang, K., Mao, W., Ma, Y., & Fan, X. (2017). A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. *Energy Conversion and Management*, 136, 439–451.
- Zhang, X., Yu, L., Wang, S., & Lai, K. K. (2009). Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method. *Energy Economics*, 31, 768–778.