Forecasting tourist arrivals using dual decomposition strategy and an improved fuzzy time series method

Xiaozhen Liang[1] · Zhikun Wu[1]

Abstract

Tourist arrivals forecasting has become an increasingly hot issue due to its important role in the tourism industry and hence the whole economy of a country. However, owing to the complex characteristics of tourist arrivals series, such as seasonality, randomness, and non-linearity, forecasting tourist arrivals remains a challenging task. In this paper, a hybrid model of dual decomposition and an improved fuzzy time series method is proposed for tourist arrivals forecasting. In the novel model, two stages are mainly involved, i.e., dual decomposition and integrated forecasting. In the first stage, a dual decomposition strategy, which can overcome the potential defects of individual decomposition approaches, is designed to fully extract the main features of the tourist arrivals series and reduce the data complexity. In the second stage, a fuzzy time series method with fuzzy C-means algorithm as the discretization method is developed for prediction. In the empirical study, the proposed model is implemented to predict the monthly tourist arrivals to Hong Kong from USA, UK, and Germany. The results show that our hybrid model can obtain more accurate and more robust prediction results than benchmark models. Relative to the benchmark fuzzy time series models, the hybrid models using traditional decomposition methods and strategies, as well as the traditional single prediction models, our proposed model shows a significant improvement, with the improvement percentages at about 80, 70, and 50%, respectively. Therefore, we can conclude that the proposed model is a very promising tool for forecasting future tourist arrivals or other related fields with complex time series.

Keywords Forecast of tourist arrivals · Dual decomposition strategy · Fuzzy C-means algorithm · Fuzzy time series

1 Introduction

From a worldwide perspective, tourism makes a great contribution to economic growth [1, 2]. Take the case of China, according to China National Tourism Administration, the total revenue of China's tourism industry was 6.63 trillion yuan in 2019, which raises 11 percent compared with 2018 and accounts for more than 11% of China's GDP. Therefore, forecasting tourist arrivals plays an important role in forecasting future economic growth. Moreover, tourist arrivals forecasting can provide valuable reference for subsequent strategic planning and policy formulation [3, 4]. Accurate forecast of tourist arrivals can make the operation of travel agencies more effective and help tourist destinations to be better managed, which is very important to the sustainable development of the whole tourism industry and even the entire economy. In general, the study of tourist arrivals forecasting is of great significance to the whole society, both politically and economically. However, due to the complex characteristics of tourist arrivals series (e.g., seasonality, randomness, and non-linearity), tourist arrivals forecasting is still a difficult problem.

To solve this problem, a growing number of researchers are paying attention to the analysis and prediction of tourist arrivals. Meanwhile, numerous models have been formulated and designed to forecast tourist arrivals. According to related literature [5], single forecasting models that were widely used to forecast tourist arrivals can fall into two main types, i.e., econometric models and artificial intelligence (AI) models. The econometric models, such as

✉ Xiaozhen Liang
liangxz@shu.edu.cn

1 School of Management, Shanghai University,
Shanghai 200444, China

autoregressive moving average (ARMA) [6], autoregressive integrated moving average (ARIMA) [7], exponential smoothing (ES) [8], and generalized autoregressive conditional heteroskedasticity (GARCH) [9], are more suitable for forecasting a relatively stable time series [10]. When forecasting data such as tourist arrivals with nonlinear characteristic and rapid changes, it has been pointed out that econometric models perform poorly in achieving effective prediction results [11]. As for the AI models, the development of AI techniques has greatly promoted their application in various fields, including air quality early warning [12], the prediction of crude oil price [13], and electricity price [14]. The commonly used AI models for forecasting tourist arrivals include artificial neural networks (ANNs) [15], extreme learning machine (ELM) [16], and support vector machine (SVM) [17]. Compared with the econometric models, AI models are more effective due to their strong robustness and fault tolerance. All these forecasting methods have significantly promoted the sustainable development of world tourism industry.

However, almost every single forecasting model has its pros and cons, and even AI models are unlikely to achieve satisfactory performance in all scenarios. For example, due to the poor effect of extrapolation, narrow prediction scale, and high requirement on data quantity and quality, econometric models are unsuitable for data with high fluctuation and noise [18]; for ANNs, the prediction performance of the models will be affected by the initial weights and thresholds which are generated randomly [19]. For this reason, researchers started to turn their attention to developing hybrid forecasting models by incorporating some existing single methods. Numerous studies have shown that hybrid forecasting model can achieve relatively ideal effect and has become the current mainstream forecasting method [20].

In order to develop hybrid models for forecasting, some decomposition methods, such as variational mode decomposition (VMD) [21], empirical mode decomposition (EMD) [22], and wavelet transform (WT) [23], have been employed to extract the main features of raw series. Our previous work [24] has proved that data preprocessing with an effective decomposition method can significantly improve prediction performance. Specifically, data preprocessing strategies can fall into two types. One refers to "decomposition & de-noising" strategy [25]. Under this strategy, the noisy information of the original series is first removed, then the forecasting model is established by using the filtered time series. The other refers to "divide & conquer" strategy [26]. Under this strategy, raw series is first decomposed into several components, which then can be predicted using a determined prediction model respectively, and finally, the predicted values of all components are integrated to get the final results. In terms of tourist arrivals forecasting, Jiang and Ma [27] used fast ensemble EMD (FEEMD) method for data preprocessing to build a hybrid model, which performs well in forecasting future tourist arrivals. Similarly, by using WT for data preprocessing and kernel-based ELM and ARMA for forecasting, Yang et al. [28] developed a hybrid model for daily tourist arrivals forecasting, and the empirical results based on three real tourism markets show that the developed model has good linear and non-linear prediction abilities. In the above studies, the hybrid forecasting models can improve the prediction accuracy and thus perform better than all the considered benchmark models. Nevertheless, data preprocessing only using a single decomposition method in the hybrid model may not be able to fully extract the main features of the tourist arrivals series. Furthermore, inherent defects existed in some data decomposition methods, such as mode mixing and endpoint effect, may also limit their application in feature extraction [29]. In fact, problems such as incomplete data feature extraction and the inherent defects existed in decomposition methods will make it difficult for the hybrid model to achieve satisfactory prediction results. Therefore, to improve the prediction performance, it is worth further improve the data preprocessing techniques in future work.

In addition, there is a problem that the commonly used single forecasting models have a poor interpretation of the prediction results. The fuzzy time series (FTS) model which divides the universe of discourse based on historical data features can solve this problem well. However, most of the traditional FTS models divide the universe of discourse with equal widths and ignore the potential features of the data, which makes the prediction results still unsatisfactory [30]. To address this issue, scholars developed some novel methods for dividing the universe of discourse, such as genetic algorithms and clustering algorithms. Therefore, from the perspective of strengthening the interpretation of the results and improving model accuracy, it is of great value to further explore how to divide the universe of discourse of FTS by fuzzy C-means (FCM) algorithm.

To sum up, the above analysis shows that the existing studies are insufficient to comprehensively improve the forecasting effectiveness. Thus, it is very urgent for sustainable economic and social development to develop a novel forecasting model of tourist arrivals for the tourism industry and significantly improve the forecasting effectiveness.

This paper proposes a novel hybrid forecasting model of tourist arrivals using dual decomposition strategy and an improved fuzzy time series method. Two stages are included in this hybrid model: dual decomposition, and integrated forecasting. In the first stage, the seasonal adjustment method (i.e., X12-ARIMA [31]) is employed to

decompose the tourist arrivals data to extract its significant seasonal characteristics, and then an improved empirical mode decomposition method (i.e., ICEEMDAN [29]) is applied to decompose the remaining component sequences for reducing data complexity. Then in the second stage, the FTS model with the universe of discourse divided by the FCM algorithm, i.e., FCM-FTS method, is used to model and predict each component sequence after the second decomposition, and the predicted values of all the components are linearly summed up to get the final results.

The main contributions of this paper can be summarized as below:

(1) Most importantly, we develop a hybrid forecasting model with high accuracy and high robustness, and its effectiveness has been verified in forecasting Hong Kong's inbound tourist arrivals. According to the experimental results, our hybrid model can decompose and extract the complex features of the raw series, thus obtaining more accurate and more robust prediction results. Hence, it is a very effective tool to predict real tourism markets and can provide valuable reference for tourism decision-making.

(2) Our hybrid forecasting model has two major differences from the traditional hybrid approaches. Firstly, a different strategy for data preprocessing is presented. In most of the former research, individual decomposition approaches have been adopted generally to decompose the raw series, of which the main features may not be fully extracted by such data preprocessing strategy. Therefore, this paper presents a dual decomposition strategy based on X12-ARIMA and ICEEMDAN, which can overcome the drawbacks of the traditional data preprocessing strategies and further improve the prediction performance. Secondly, an effective clustering algorithm, i.e., FCM, is adopted to optimize the domain partition module of FTS model, of which the performance has been successfully improved.

(3) In terms of numerical experiments, this paper not only compares the proposed hybrid model with five commonly used single forecasting models but also compares it with other six hybrid forecasting models using different data preprocessing strategies, which comprehensively demonstrates the superiority of our model. In addition, the benchmark models considered can represent currently popular modeling strategies and ideas, similar to the high-quality papers published in international journals in recent years. On the basis of comparative study, this paper verifies and demonstrates the significance of the components of our hybrid model in detail, such as the validity of X12-ARIMA and ICEEMDAN, as

well as the superiority of dual decomposition strategy and FCM method. Moreover, this paper also verifies the robustness of our hybrid model. To sum up, we finally demonstrate that the developed novel tourist arrivals forecasting model has high superiority and practical values for the real tourism markets.

(4) To verify the model prediction performance, this paper provides a scientific evaluation and an in-depth discussion of the prediction results. We use six typical criteria, including average error (AE), mean absolute percentage error (MAPE), root mean square error (RMSE), mean absolute error (MAE), Theil inequality coefficient (TIC), and index of agreement (IA), to evaluate the performance of the forecasting models. Moreover, we further demonstrate the superiority of the proposed model through an insightful discussion from five aspects: (a) the model robustness according to the prediction performance at different years; (b) the significance of the model from the perspective of statistics; (c) the forecasting effectiveness based on the comparative studies; (d) the improvement percentage relative to the benchmarks; and (e) the grey relational analysis of all the models involved.

The rest of the paper is arranged as follows. Section 2 introduces the main methods involved and the overall framework of the developed hybrid forecasting model. Section 3 mainly describes the data, conducts the comparative experiments, and analyzes the prediction results. Section 4 presents the related discussions. Finally, Sect. 5 concludes the study.

2 Methods

This section presents a hybrid model of dual decomposition and an improved fuzzy time series method for tourist arrivals forecasting. Specifically, Sects. 2.1–2.5 describe the relevant methods for decomposition and prediction respectively, and Sect. 2.6 provides the overall process of our hybrid model. Table 3 in Appendix 1 shows the used nomenclature in this paper.

2.1 X12-ARIMA

X12-ARIMA [31] is a popular seasonal adjustment method developed by the United States Census Bureau, which mainly includes two functional modules: regARIMA module and X-11 seasonal decomposition module. In particular, the regARIMA module can carry out various types of data preprocessing, such as outlier detection and correction, estimation, and elimination of the influence of

calendar factors [32]. The X-11 seasonal decomposition module decomposes the preprocessed data through multiple iterations of moving average method to form a seasonal factors series and a seasonally adjusted series. For the purpose of this paper, we just introduce the basic algorithm for the X-11 seasonal decomposition module.

It is assumed that the monthly series can be decomposed into a seasonal factor (i.e., S), a trend-cycle factor (i.e., TC), and an irregular factor (i.e., I). Two main steps are involved in the X-11 seasonal decomposition module:

Step 1 Estimation of the initial components

Firstly, the 2×12 moving average method is applied to estimate the initial TC component sequences. Then, this TC component is subtracted from the raw time series to obtain the initial estimation of the seasonal-irregular component (i.e., SI). Next, the 3×3 moving average method is applied to estimate the initial seasonal component, which then is normalized by a 2×12 moving average. Finally, the normalized seasonal component is subtracted from the raw series to obtain the initial estimation of the seasonally adjusted series (i.e., SA).

Step 2 Final seasonal adjustment

Firstly, the Henderson moving average method is used to obtain the second estimation of the TC component from the initially estimated SA series. Then, this new TC component is subtracted from the raw series to obtain the second estimation of the SI component. Next, the 3×5 moving average method is applied to estimate a new seasonal component, which then is normalized by a 2×12 moving average. Finally, the normalized seasonal component is removed from the raw series to obtain the final SA series.

It is worth noting that the selection of the number of terms in the moving average is critical in the X-11 seasonal decomposition module. The higher the number of terms, the more irregular factors can be eliminated. But as the number of terms increases, more information is lost. For monthly series that change periodically on a 12-month basis, a centered 12-term moving average can be considered for obtaining the initial TC component and the normalized seasonal component. However, if the series to be decomposed is also an economic flow time series (such as the monthly tourist arrivals), a 2×12 moving average is required to ensure that each element of the newly generated sequence after using the moving average is aligned with that of the raw series. For other parts of the module, the number of terms in the moving average is specified with reference to the standard X-11 procedure [33].

2.2 ICEEMDAN

Traditional empirical mode decomposition methods, including empirical mode decomposition (EMD) [22],

ensemble EMD (EEMD) [34], and complete ensemble EMD with adaptive noise (CEEMDAN) [35], have some problems such as mode mixing, noise, and redundancy, and pseudo components after decomposition. Aiming at these problems, Colominas et al. [29] proposed an improved complete ensemble EMD with adaptive noise (ICEEMDAN), which has a higher ability to extract the components of the complex time series with different time scale features. The following are the main steps and relevant formulas of this algorithm:

Step 1 Calculate the first residue of the original series using the following equation:

$$r_1 = \frac{1}{I} \sum_{i=1}^{I} M\left(x + \beta_1 E_1\left(w^i\right)\right), \tag{1}$$

where $E_k()$ is an operator, which uses EMD method to decompose a series into several intrinsic mode functions (IMFs) and one residual, with the k-th IMF component (i.e., the k-th mode) as output; M() also represents an operator, which produces the local mean (i.e., the mean of the upper and lower envelopes) of a series; x represents the original time series; w^i indicates a realization of white noise, whose mean value is zero and variance is one, $i = 1, 2, ..., I$, and I is the number of times that white noise is added; β_k is the parameter that controls the energy of the white noise in each iteration, $k = 1, 2, ..., K$, and K is the maximum iterations. Mode mixing is defined as either a single IMF consisting of components of widely disparate scales or a component of a similar scale residing in different IMFs [34]. The purpose of including white noise in this equation is to avoid the mode mixing problem so that the components of the complex time series with different time scales can be identified and extracted more accurately.

Step 2 Subtract the first residue from the original series to get the first mode d_1:

$$d_1 = x - r_1. \tag{2}$$

Step 3 Obtain the second residue of the original time series, i.e., r_2, in the same way as in step 1, and finally obtain the second mode d_2 by the following equation:

$$d_2 = r_1 - r_2 = r_1 - \frac{1}{I} \sum_{i=1}^{I} M\left(r_1 + \beta_2 E_2\left(w^i\right)\right). \tag{3}$$

Step 4 Obtain the k-th residue and k-th mode by the following equation:

$$d_k = r_{k-1} - r_k = r_{k-1} - \frac{1}{I} \sum_{i=1}^{I} M\left(r_{k-1} + \beta_k E_k\left(w^i\right)\right). \tag{4}$$

Step 5 Return to step 4 for next k until the residue can no longer be decomposed or K is reached.

2.3 Fuzzy C-means clustering

The fuzzy C-means (FCM) algorithm is one of the commonly used clustering methods [36]. The basic idea of FCM algorithm is to continuously update the cluster centers of all data and the membership degrees of each data point belonging to all cluster centers through iterative calculation, until the dissimilarity index function and the iteration error reach the preset minimum value. The following are the main steps and related formulas of FCM algorithm:

Step 1 Calculate the number of cluster centers:

$$c = \left[(x_{\max} - x_{\min}) \middle/ \frac{\sum_{t=2}^{n_1} |x_t - x_{t-1}|}{n_1 - 1} \right], \quad (5)$$

where $x_t (t = 1, 2, \cdots, n_1) \in R$ is the element of the original series x, and n_1 is the number of elements in x. c is the number of cluster centers, $c \in \{2, 3, ..., n_1 - 1\}$. $x_{\max}$ and $x_{\min}$ represent the maximum and minimum values in the original series, respectively; [] represents the rounding operation.

Step 2 Initialize the cluster centers. Randomly select c samples in x as the initial cluster centers $V(0) = \{v_{01}, v_{02}, ..., v_{0c}\}$.

Step 3 Calculate the membership matrix:

$$u_{ij} = \left(\sum_{r=1}^{c} \frac{d_{ij}}{d_{rj}} \right)^{-1}, \quad (6)$$

where d_{ij} is the Euclidean distance from the element x_j to the cluster center v_i, $i = 1, 2, ..., c$, $j = 1, 2, ..., n_1$.

Step 4 Iterate new cluster centers:

$$v_i = \sum_{j=1}^{n_1} u_{ij}^m x_j \middle/ \sum_{j=1}^{n_1} u_{ij}^m, \quad (7)$$

where m is the weighted index of membership degree, which is used to adjust the fuzzy degree of the clustering results, generally $m = 2$.

Step 5 Repeat steps 3 and 4 iteratively until the condition $\|V(k+1) - V(k)\| < \varepsilon$ is satisfied (ε is the iteration stop threshold) or the maximum iterations are reached.

2.4 Fuzzy time series algorithm

On the basis of the fuzzy set theory and other concepts proposed by Zadeh [37], Song and Chisom [38, 39]

established the fuzzy time series (FTS) model, which was successfully used to predict the enrollment data for the University of Alabama. Subsequently, traditional FTS model and its variants were widely applied in other fields (e.g., temperature, stock index, and network traffic) to perform forecasting and have achieved good forecasting results [40, 41]. The basic definitions of FTS are as below:

Definition 1 It is assumed that U is a given universe of discourse, which can be divided into n_2 subintervals in order, then $U = \{u_1, u_2, \cdots, u_{n_2}\}$. Define A as the fuzzy set on the universe U, expressed as:

$$A = \frac{f_A(u_1)}{u_1} + \frac{f_A(u_2)}{u_2} + \cdots + \frac{f_A(u_{n_2})}{u_{n_2}}, \quad (8)$$

where $f_A( )$ is the membership function of fuzzy set A, $f_A( ) \in [0, 1]$; $f_A(u_i)$ represents the membership degree of the interval $u_i (1 \le i \le n_2)$ with respect to the fuzzy set A.

Definition 2 Let the original time series $Y = \{y_t\} = \{Y(t)\} (t = 1, 2, ...)$ be a subset of the real number field R. Define a set of fuzzy sets $f_i(t)$ ($i = 1, 2, ...$) on the series Y, and the series $F(t) = \{f_1(t), f_2(t), \cdots\}$, then $F = \{F(t)\} (t = 1, 2, ...)$ is a fuzzy time series defined on Y.

Definition 3 Suppose there is a fuzzy logical relationship (FLR), i.e., $R(t, t - 1)$, between $F(t)$ and $F(t-1)$, which satisfies:

$$F(t) = F(t-1) \odot R(t, t-1), \quad (9)$$

then it is said that $F(t)$ is obtained only by $F(t-1)$ ($\odot$ is a combination operator). And set $F(t-1) = A_i$ and $F(t) = A_j$, then the FLR can also be expressed as: $A_i \rightarrow A_j$. Between them, A_i and A_j are called the left-hand side (LHS) and right-hand side (RHS) of the FLR, respectively.

Definition 4 All the single FLRs with the same LHS can be composed into the same fuzzy logical relationship set (FLRS). For example, the three FLRs ($A_l \rightarrow A_{r1}, A_l \rightarrow A_{r2}, A_l \rightarrow A_{r3}$) with the same LHS can be composed into one FLRS, which is expressed as $A_l \rightarrow A_{r1}, A_{r2}, A_{r3}$.

2.5 FCM-FTS model

For fuzzy time series, the unsupervised discretization method was generally used to obtain the equal-width intervals, which is simple and convenient. However, equal-width interval partitioning method is not very interpretable for the intervals and the forecasting results are not accurate enough [42]. The FCM clustering algorithm

partitions the universe of discourse according to data characteristics, which is more objective. Furthermore, this algorithm can explain the actual meaning of each sub-interval by the explanation of the clustering center, which is more scientific and reasonable than the equal-width interval partitioning method. In this paper, the FTS model optimized by Chen [43] with the FCM algorithm partitioning the universe of discourse, i.e., FCM-FTS model, is applied for prediction. The specific steps are as follows:

Step 1 Detect the stationarity of the time series to be predicted by the augmented Dickey-Fuller (ADF) test [44]. If the series is stable, turn to step 2 directly. Otherwise, make the series stable by preprocessing it with the difference method [7].

Step 2 Divide the universe U into n_2 intervals by the FCM clustering algorithm, then $U = \{u_1, u_2, ..., u_{n_2}\}$.

Step 3 Define the fuzzy set for the raw time series by determining the fuzzy membership function. Then, construct fuzzy set A_i based on the intervals. And the fuzzy membership function $f_{A_i}(u_j)$ can be defined as follows [45]:

$$f_{A_i}(u_j) = \begin{cases} 1, & i = j \\ 0.5, & i = j + 1 \\ 0, & \text{others.} \end{cases} \tag{10}$$

Step 4 Fuzzify the actual values. Fuzzify a raw value to A_i when the highest degree of membership of that raw value is in A_i [43].

$$\begin{aligned} fuzzify(actual_t) &= A_i \text{ if } f_{actual_t}(A_i) \\ &= \max[f_{actual_t}(A_z)], \ z = 1,2,...,M, \end{aligned} \tag{11}$$

where $f_{actual_t}(A_z)$ denotes the degree of membership of the actual value at t under A_z, and M denotes the number of the fuzzy sets.

Step 5 Establish and group the FLR. According to the definition 3 and 4 in Sect. 2.4, the first-order FLR and FLRS are constructed for all fuzzy sets of the fuzzy time series.

Step 6 Determine and standardize the weight matrix. The weights can be calculated and standardized based on step 5, and then the centroid defuzzification method can be used to further calculate the defuzzification matrix.

$$\begin{aligned} W\_s(t) &= (W'_1, W'_2, ..., W'_k), \\ W'_i &= W_i / \sum_{i=1}^{k} W_i, \end{aligned} \tag{12}$$

where W_i is the unstandardized weighting matrix element, and W'_i denotes the standardized one. W_s represents the standardized weighting matrix.

Step 7 Obtain the forecasting results. Multiply the defuzzified matrix by standardized weighting matrices to obtain the rudimentary forecasting results:

$$\hat{F}(t) = D(t - 1) \times W\_s(t - 1), \tag{13}$$

where $\hat{F}(t)$ denotes the forecasting result and D denotes the defuzzified matrix.

2.6 Overall process of the proposed model

To forecast tourist arrivals, we propose a novel hybrid model of X12-ARIMA, ICEEMDAN, FCM, and FTS, namely X12-ARIMA-ICEEMDAN-FCM-FTS model. This hybrid model includes two stages, i.e., dual decomposition and integrated forecasting. Figure 1 shows the overall process of our hybrid forecasting model, with four main steps involved as follows:

2.6.1 Stage 1: Dual decomposition

Step1: Considering the seasonal characteristics of the tourist arrivals data, first the original time series is decomposed by X12-ARIMA method, extracting the seasonal component and obtaining the seasonally adjusted series.

Step 2: ICEEMDAN is then used to decompose the seasonally adjusted series into n-1 intrinsic mode functions (IMF_1, IMF_2,…, IMF_{n-1}) with different time scale features and one smooth residual series (Residue), in order to reduce the data complexity.

2.6.2 Stage 2: Integrated forecasting

Step 3 The FCM-FTS method is used to model and predict the seasonal factors series, n-1 IMFs component series, and the residual series, respectively.

Step 4 Finally, the predicted values for all the components, respectively noted as SEA', IMF_1', IMF_2',…, IMF_{n-1}', and Residue', are linearly summed up to get the final prediction results.

3 Experiment

In this section, we used the developed hybrid model to forecast Hong Kong's inbound tourist arrivals from three countries (i.e., USA, UK, and Germany) for illustration and verification purposes. In particular, several related experiments were carried out with multiple control groups set up,

Fig. 1 Overall process of the proposed hybrid forecasting model
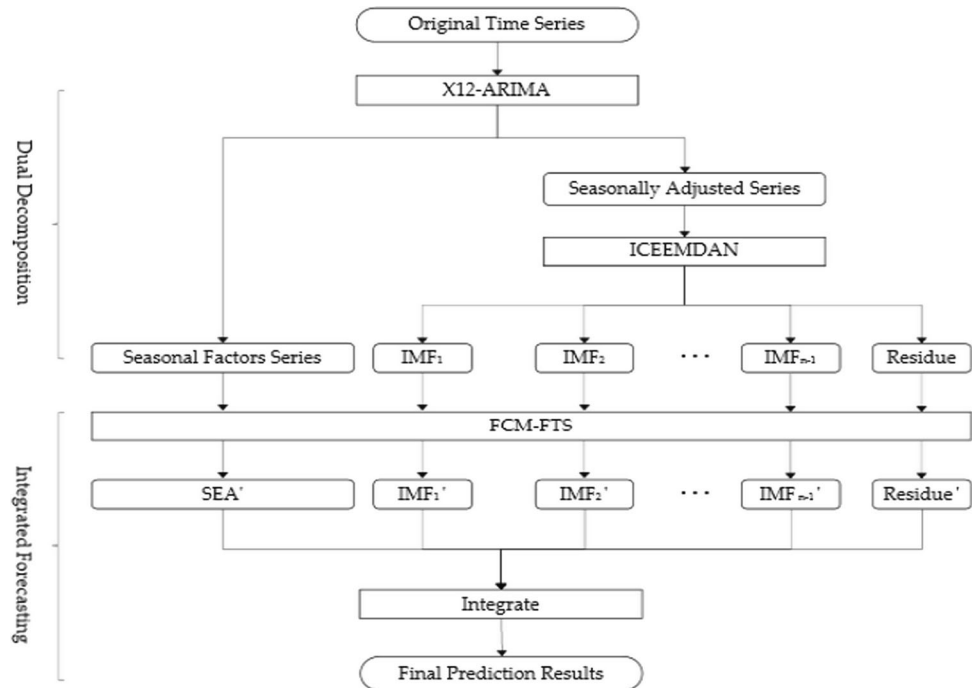
Fig. 2 The monthly tourist arrivals to Hong Kong from USA, UK and GER
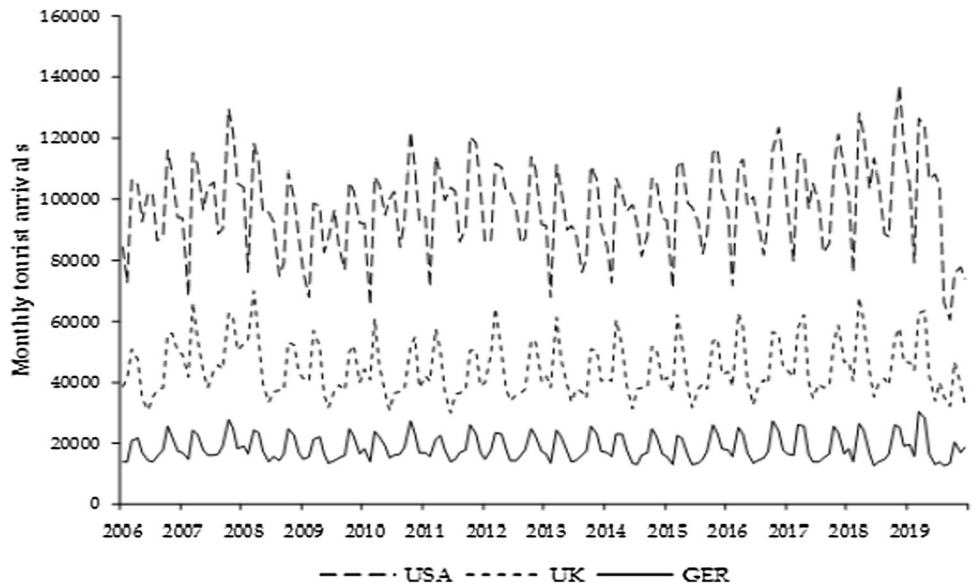
and comparison and analysis were made from various aspects to verify the performance of our proposed model, in which the main parameters involved can be seen in Table 4 (in Appendix 1). Furthermore, final prediction results were taken as the average of 100 runs to avoid the influence of random factors.

3.1 Data description

The monthly tourist arrivals to Hong Kong from USA, UK, and Germany (simply noted as GER) are selected as data samples, as shown in Fig. 2. For each series, there are 168 observations, covering the period from January 2006 to December 2019, which can be obtained from Wind Database (http://www.wind.com.cn/). Meanwhile, to evaluate the model robustness, the samples are rolled backward for one year at a time, thus each sample can produce three subsamples with the same number of observations,

covering the periods from January 2006 to December 2017, January 2007 to December 2018, and January 2008 to December 2019, respectively. The sample data are shown in detail in Table 5 (in Appendix 1). In addition, a link to the supplementary material related to this article (including the data and the code) can be found in Appendix 2.

In addition, the experiments conducted in this paper all perform one-step-ahead predictions. The data of each subsample can be divided into training set for model training and testing set for evaluating model performance. In particular, the data of the preceding 11 years (132 observations) are used as training set, while the following year (12 observations) as testing set. Finally, the monthly tourist arrivals in 2017, 2018, and 2019 are predicted, respectively. According to the results of the three forecasting years, the final prediction performance of the proposed model is evaluated.

3.2 Evaluation criteria

Considering that there is no universally applicable standard for prediction model error evaluation [46], we choose six popular criteria (i.e., AE, MAPE, RMSE, MAE, TIC, and IA) to evaluate the model prediction performance, as listed in Table 1. Obviously, except for the IA criterion, a smaller evaluation criterion means that the prediction is more accurate.

3.3 Experiment design

In this paper, three experiments were designed for comparison purpose. In Experiment I, the proposed model is compared with other six hybrid models based on different decomposition methods to prove the superiority of the proposed dual decomposition strategy. Specifically, the six hybrid models selected as benchmarks are as follows: X12-ARIMA-FCM-FTS, ICEEMDAN-FCM-FTS, CEEMDAN-FCM-FTS, EEMD-FCM-FTS, WD-FCM-FTS, and ICEEMDAN(R)-FCM-FTS. In Experiment II, from a longitudinal perspective, the proposed model is compared with several partial hybrid models which only use some of the single methods involved in our model. On the basis of this experiment, we try to demonstrate the importance of the components of our model, including the effectiveness of X12-ARIMA and ICEEMDAN methods, as well as the superiority of FCM algorithm and dual decomposition strategy. In Experiment III, we further compare the proposed model with some popular single models, such as typical econometric models and ANNs, to prove the superiority of our model.

3.4 Experiment I

To fully verify the forecasting superiority of our proposed dual decomposition strategy, two types of comparative analysis were carried out. In Comparison I, we compare the performance of five hybrid models using different decomposition methods (including X12 -ARIMA, WD, EEMD, CEEMDAN, ICEEMDAN) and the same

Table 1 Evaluation criteria

Criterion	Definition	Equation
AE	Average error of n_3 prediction results	$AE = \frac{1}{n_3} \sum_{t=1}^{n_3} (y_t - \hat{y}_t)$
MAPE	Average of n_3 absolute percentage errors	$MAPE = \frac{1}{n_3} \sum_{t=1}^{n_3} \frac{|y_t - \hat{y}_t|}{y_t} \times 100\%$
RMSE	Square root of average of the error squares	$RMSE = \sqrt{\frac{1}{n_3} \sum_{t=1}^{n_3} (y_t - \hat{y}_t)^2}$
MAE	Mean absolute error of n_3 prediction results	$MAE = \frac{1}{n_3} \sum_{t=1}^{n_3} |y_t - \hat{y}_t|$
TIC	Theil inequality coefficient	$TIC = \sqrt{\frac{1}{n_3} \sum_{t=1}^{n_3} (y_t - \hat{y}_t)^2} \Big/ \left(\sqrt{\frac{1}{n_3} \times \sum_{t=1}^{n_3} y_t^2} + \sqrt{\frac{1}{n_3} \times \sum_{t=1}^{n_3} \hat{y}_t^2} \right)$
IA	Index of agreement	$IA = 1 - \sum_{t=1}^{n_3} (y_t - \hat{y}_t)^2 \Big/ \sum_{t=1}^{n_3} (|\hat{y}_t - \overline{y}| + |y_t - \overline{y}|)^2$

y_t and $\hat{y}_t$ are the actual and predicted values at time t respectively, n_3 is the size of the testing set, and $\overline{y}$ is the mean of the actual values.

forecasting method (i.e., FCM-FTS model) to demonstrate the effectiveness of the individual decomposition methods used in our model. Tables 6 and 7 in Appendix 1 show the main parameters of the compared decomposition methods and the corresponding prediction results, respectively. In Comparison II, we compare our hybrid model with ICEEMDAN-FCM-FTS, X12-ARIMA-FCM-FTS, and ICEEMDAN(R)-FCM-FTS models to further prove the superiority of the proposed dual decomposition strategy, with the corresponding experimental results shown in Table 8 in Appendix 1. In detail, the ICEEMDAN(R)-FCM-FTS model is performed under the decomposition & de-noising strategy [25], while other compared hybrid models in this subsection are performed under the divide & conquer strategy [26]. Generally speaking, the two data preprocessing strategies are widely used for forecasting complex data with high volatility and irregularity and can represent currently popular modeling strategies.

The detailed comparison and analysis are as below:

(1)　In Comparison I, by comparing the prediction performance of four hybrid models, including ICEEMDAN-FCM-FTS, CEEMDAN-FCM-FTS, EEMD-FCM-FTS, and WD-FCM-FTS, we can find that the FCM-FTS forecasting model combined with the ICEEMDAN is superior to that combined with CEEMDAN, EEMD, and WD, which shows the advantages of ICEEMDAN when compared with other traditional decomposition methods. Moreover, it can be found that the X12-ARIMA-FCM-FTS performs better than the above-mentioned four hybrid models in all cases. Taking case 1 as an example, the value of MAPE in the X12-ARIMA-FCM-FTS model is the lowest (9.5122%), in the ICEEMDAN-FCM-FTS model is the second lowest (11.8851%), while in the CEEMDAN-FCM-FTS, EEMD-FCM-FTS and WD-FCM-FTS models are, respectively 12.6937%, 13.2176%, and 14.0263%. Similar results can be obtained in other two cases, which fully prove the superiority of ICEEMDAN and the necessity of seasonal decomposition for tourist arrivals series forecasting. Moreover, for the data (such as Hong Kong's tourist arrivals) with significant seasonal characteristics, using X12-ARIMA for seasonal decomposition can effectively improve the model prediction performance.

(2)　In Comparison II, by comparing the proposed model with other three hybrid models, including ICEEMDAN-FCM-FTS, X12-ARIMA-FCM-FTS, and ICEEMDAN(R)-FCM-FTS, we can see that our

model, i.e., X12-ARIMA-ICEEMDAN-FCM-FTS, has the best prediction performance. In general, the proposed dual decomposition strategy in this paper has more advantages than the above-mentioned traditional decomposition strategies and can achieve better prediction performance. It can be seen from Table 8 (in Appendix 1) that except for the AE criterion in case 2, the X12-ARIMA-ICEEMDAN-FCM-FTS model performs best in all evaluation criteria in all cases. Obviously, the proposed dual decomposition strategy plays a significant role in improving the model prediction performance.

Remark 1 Based on the comparative analysis of the values of AE, MAPE, RMSE, MAE, TIC, and IA criteria, it can be found that compared with other traditional decomposition methods (such as WD and EEMD), ICEEMDAN is more effective when combined with the prediction model, reflecting the superiority of ICEEMDAN. Meanwhile, the X12-ARIMA-FCM-FTS model performs best among all the hybrid models in Comparison I, which again verifies the rationality and necessity of adopting targeted data preprocessing strategy for tourist arrivals data with significant seasonal characteristics. Furthermore, in Comparison II, we also successfully prove the effectiveness of the proposed dual decomposition strategy compared with the traditional data preprocessing strategies. To sum up, by applying the divide & conquer strategy to both the raw series with seasonal patterns and the seasonally adjusted series, the proposed dual decomposition strategy can successfully overcome the potential disadvantages of individual decomposition approaches and plays an important role in improving the model prediction performance.

3.5 Experiment II

Experiment II was designed mainly to verify the effectiveness of the hybrid modeling strategy based on X12-ARIMA, ICEEMDAN, FTS, and FCM. Thus, Experiment II consists of Comparison I and Comparison II for longitudinal comparison purpose. Generally speaking, the discretization method is important to the prediction performance of an FTS method. In the related studies, as the most commonly used unsupervised discretization methods, the equal frequency (EF) and equal width (EW) interval algorithms cannot always achieve satisfactory forecasting results. Therefore, we use FCM algorithm as the discretization method for a fuzzy time series. Accordingly, in Comparison I set in this subsection, the FTS model with FCM method dividing the universe of discourse

(i.e., FCM-FTS model) is compared with that with EW and EF methods dividing the universe of discourse (i.e., EW-FTS and EF-FTS models) to prove the effectiveness of FCM algorithm. In Comparison II, we compare our model, i.e., X12-ARIMA-ICEEMDAN-FCM-FTS, with other three models, including FCM-FTS, X12-ARIMA-FCM-FTS, and ICEEMDAN-FCM-FTS models, which are partial hybrid models only using some of the single methods involved in the proposed model, to fully illustrate the rationality of our hybrid modeling strategy. The detailed comparison and analysis are shown below:

(1)　As demonstrated in Table 9 (in Appendix 1), we can observe that compared with EW-FTS model and EF-FTS model, the FCM-FTS model has almost all the best evaluation criterion values in all three cases. Taking case 1 as an example, except for the AE criterion, the values of MAE, RMSE, MAPE, TIC, and IA of the FCM-FTS model are 16,127.6028, 20,978.7502, 15.8852%, 0.1013, and 0.9910, respectively, which are all smaller than that of the EW-FTS and EF-FTS models. This fully demonstrates that using FCM algorithm as the discretization method can improve the performance of the fuzzy time series model more effectively.

(2)　Table 10 in Appendix 1 shows the experimental results of Comparison II. Obviously, the X12-ARIMA-FCM-FTS and ICEEMDAN-FCM-FTS, models perform better than the FCM-FTS model, which once again shows the importance and necessity of using data preprocessing to the data with complex characteristics. By comparing the proposed hybrid model with the FCM-FTS, X12-ARIMA-FCM-FTS and ICEEMDAN-FCM-FTS models, we can see that our model, i.e., X12-ARIMA-

ICEEMDAN-FCM-FTS, has the best prediction performance. For example, the MAPE values of our proposed model in the three cases are 4.2343%, 3.4946%, and 4.7533% respectively, evidently lower than that of the other three models. Moreover, focusing on the IA criterion, the criterion values of our model in the three cases are all greater than 0.999, while the three compared models' criterion values are all below 0.999.

Remark 2 It is necessary to find a reasonable and effective discretization method for the fuzzy time series model. In Comparison I, the FTS model with FCM as the discretization method performs better than that with EW and EF as the discretization method. In addition, the proposed forecasting model performs best in Comparison II, which intuitively shows that the proposed hybrid modeling strategy can significantly improve the performance of the benchmark FTS model by integrating the advantages of every single method greatly.

3.6 Experiment III

In Experiment III, by taking some commonly used single forecasting models as benchmarks, including traditional econometric models, typical ANNs, and other popular models, we further test the prediction performance of our model. Specifically, the extreme learning machine (ELM) and backpropagation neural network (BPNN) are chosen for comparison as typical ANNs, while seasonal ARIMA (SARIMA) and double exponential smoothing (DES) as typical traditional econometric models. Meanwhile, support vector regression (SVR), which is popular in forecasting, is also chosen for comparison in this experiment.
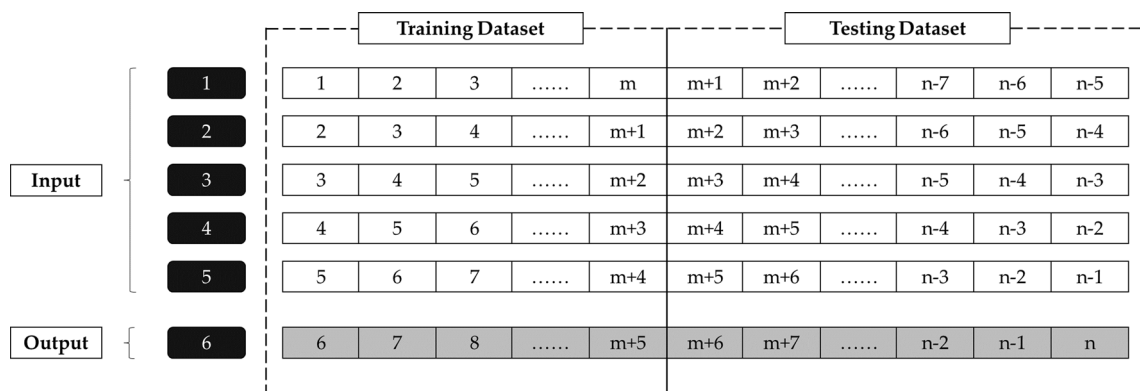

Fig. 3 The training set and testing set of ANN models

Accordingly, two standard models of the ANNs, i.e., BPNN (5–1-1) and ELM (5–1-1) are established respectively, of which the training set and testing set are presented in Fig. 3. For the SARIMA model, the parameters are determined based on a stability test and the Akaike Information Criterion (AIC) [47]. For the SVR model, radial basis function (RBF) as the most popular kernel function is chosen. The main parameters of the comparison models are shown in Table 11 (in Appendix 1), and Table 12 in Appendix 1 presents the corresponding results.

According to Table 12 (in Appendix 1), the detailed experimental results and comparative analysis are shown below:

(1) The proposed X12-ARIMA-ICEEMDAN-FCM-FTS model performs best in all three cases. Since the AE criterion is not sufficient to reflect the prediction accuracy, more attention should be paid to MAPE, RMSE, and MAE criteria when evaluating the model prediction performance. Taking MAPE criterion as an example, the propose forecasting model has the lowest MAPE values in all three cases, even reaching 3.4946% in case 2, which fully reflects its superior prediction performance.

(2) Moreover, the BPNN performs slightly better than the ELM in the ANN models. In the traditional econometric models, the DES model performs worse, even with the MAPE value as high as 25.5657% in case 3. Comparatively speaking, the SARIMA performs better, only worse than our proposed model, which again shows that for data such as tourist arrivals series with significant seasonal characteristics, the targeted use of certain seasonal forecasting methods can achieve better prediction results.

Remark 3 Compared with the five commonly used single forecasting models, our hybrid model performs best. Among all the single models selected, the SARIMA performs best, and the DES performs worst.

4 Discussion

This section presents an in-depth comparative analysis for the prediction results of all models at different years of 2017, 2018, and 2019. Moreover, we further analyze the prediction performance of all the models involved from several different perspectives, including the DM statistics, forecasting effectiveness, improvement percentage, and grey relational degree.

4.1 Forecasting results at different years

As we all know, tourist arrivals data has complex characteristics, and is extremely vulnerable to abnormal events, resulting in abnormal fluctuations, which greatly increases the difficulty in prediction. Therefore, it is particularly important for the model to maintain a stable and great prediction performance when abnormal events occur. In view of this concern, based on three basic evaluation criteria of prediction accuracy, i.e., MAE, RMSE, and MAPE, this subsection compares and analyzes the prediction results of all the involved models at different years of 2017, 2018, and 2019. It should be noted that Hong Kong's tourism industry was severely affected in 2019 due to the outbreak of some social events, which made the data of tourist arrivals to Hong Kong in that year showing a very irregular pattern compared with the previous years. It can also be found from the experimental results that the prediction performances of each model in these three years are very different.

As demonstrated in Table 13 (in Appendix 1), we can observe that: (a) In all three cases, the proposed model performs best in all three years. Taking the prediction of 2017 in case 1 as an example, the MAPE value of our

Table 2 Improvement percentage criteria

Criterion	Definition	Equation
RE_{MAE}	The decreased relative error of MAE	$RE_{MAE} = \frac{MAE_i - MAE_j}{MAE_i} \times 100\%$
RE_{RMSE}	The decreased relative error of RMSE	$RE_{RMSE} = \frac{RMSE_i - RMSE_j}{RMSE_i} \times 100\%$
RE_{MAPE}	The decreased relative error of MAPE	$RE_{MAPE} = \frac{MAPE_i - MAPE_j}{MAPE_i} \times 100\%$

proposed model is as low as 2.30%, showing a very excellent prediction performance. (b) Compared with 2019, most of the prediction models perform better in 2017 and 2018. For example, the SARIMA model has a good performance in predicting the tourist arrivals to Hong Kong in 2017 and 2018 in case 2, where the MAE, RMSE, and MAPE values are 2440.96, 3057.41, and 4.93% in 2017 as well as 1241.60, 1952.32, and 2.44% in 2018, respectively. In 2019, the values of these three criteria are as high as 5997.63, 8297.25, and 15.69% respectively, which are several times the criterion values of the previous two years. This result verifies that the tourist arrivals to Hong Kong in 2019 have undergone extremely irregular changes compared with the previous years due to social events. Therefore, models which are built based on historical data are difficult to capture the trend variation of the data in 2019. (c) Some models show little change in the prediction results of the three years, and even their prediction performances of 2019 are better than that of the previous two years, such as the EEMD-FCM-FTS and WD-FCM-FTS models in case 2 and case 3, which indicates that these models are relatively stable and almost immune from the influence of the abnormal events in 2019. Nevertheless, the comprehensive prediction performances of these models are still poor, and the proposed model performs best among all the comparison models. Even for the data of 2019 which were affected by the abnormal events, our proposed model can still show a superior prediction performance.

4.2 Statistical hypothesis testing: Diebold-Mariano test

From a statistical perspective, Diebold and Marino [48] used the Diebold-Mariano (DM) statistic to test the difference in the significance of prediction performances between different models. For the DM test, the null hypothesis is that there is no significant difference in the prediction performances of two comparison models. If the test result rejects the null hypothesis at a certain level of significance, it indicates that the prediction performances of the two comparison models are significantly different.

For further comparison, the DM test is implemented to test the different significance of the performances between our proposed model and all the fourteen benchmark models involved in the previous designed experiments. Using mean square error as the loss function, the corresponding statistics are shown in Table 14 (in Appendix 1). For all cases and the average level, we can observe that: (a) almost all benchmark models' DM statistics are greater than the

upper limit at the 1% significance level, which once again reflects the remarkable superiority of our model; (b) among all benchmark models, the DM statistics of SARIMA and X12-ARIMA-FCM-FTS models are almost the lowest, showing that for tourist arrivals data with seasonal characteristics, specifically using some seasonal forecasting methods can achieve better forecasting results, and also proving that the seasonal decomposition method chosen in this paper is scientific and reasonable.

4.3 Forecasting effectiveness

Furthermore, we use forecasting effectiveness (FE) [12] to measure the model prediction accuracy. The higher the value of FE, the better is the prediction performance. Specifically, the first-order and second-order FE values are calculated by the expected value of the prediction accuracy sequences as well as the difference between its expected value and standard deviation, respectively. To further compare the prediction accuracy, we calculate the first-order and second-order FE values for all the involved models in this paper, as shown in Table 15 (in Appendix 1). We can observe that: (a) in all the three cases, the FE values of our hybrid model are always the highest, which are 0.957657, 0.965054, and 0.952467 as well as 0.905740, 0.928907, and 0.903511 for the first-order and second-order FE, respectively, meaning that our model performs best; (b) among all the comparison models, the FE values of the EF-FTS model are the lowest, which means that the traditional equal frequency division method (EF) is not suitable for data such as the tourist arrivals to Hong Kong.

4.4 Improvement percentage

In this subsection, we use three criteria to measure the improvement percentages of our hybrid model relative to all the comparison models. Accordingly, the improvement percentage criteria are denoted as RE_{MAE}, RE_{RMSE} and RE_{MAPE}, representing the decreased relative error (RE) of MAE, RMSE and MAPE, respectively, of which the calculation formulas can be seen in Table 2. Table 16 in Appendix 1 shows the corresponding results, and further comparisons and analyses are as below.

As reported in Table 16 in Appendix 1: (a) Our proposed model has the greatest improvement relative to the benchmark FTS models (i.e., EW-FTS, EF-FTS), with the improvement percentages at about 80%. Meanwhile, the improvement percentages of our model relative to the hybrid models using traditional decomposition methods

and strategies are also high, which are around 70%. The above results once again verify the rationality for our model using hybrid modeling strategy and dual decomposition strategy, which greatly improves the model prediction performance by integrating the advantages of each method involved. (b) Relative to the traditional single prediction models, our proposed model shows a significant improvement, with almost all of the values of the three improvement percentage criteria above 50%. Obviously, it can achieve a satisfactory prediction performance in tourist arrivals forecasting.

4.5 Grey relational degree

In this subsection, the correlation degree between the prediction results and the actual time series is measured by the grey relational degree (GRD) [49], of which a higher value means a better prediction performance. The corresponding results are shown in Table 17 (in Appendix 1), with the detailed analysis as follows: (a) the prediction results of our model have the strongest correlation to the actual series, with the values of GRD greater than 0.9 in all three cases, and the average is 0.924274; (b) the SARIMA model and X12-ARIMA-FCM-FTS model also perform very well, with the average GRD being 0.858878 and 0.871148 respectively, which once again reflects the rationality of using seasonal data preprocessing method for tourist arrivals series with significant seasonal characteristics. Therefore, we can reasonably conclude that our hybrid forecasting model has significant differences from the benchmark models in the level of prediction accuracy.

5 Conclusions

Accurate prediction of tourist arrivals is important for the whole tourism industry and also the entire economy. Meanwhile, it is of reference value to both travel agencies and tourist destinations. Unfortunately, due to its complex characteristics, tourist arrivals forecasting remains a challenging task. Thus, a hybrid model using dual decomposition strategy and an improved FTS method is proposed to predict tourist arrivals.

In the empirical study, experiments are designed using Hong Kong's tourist arrivals from USA, UK, and Germany as data samples. The results demonstrate that: (a) The novel dual decomposition strategy based on X12-ARIMA and ICEEMDAN methods proposed in this study can not only overcome the inherent defects of individual decomposition methods, such as mode mixing, noise, and redundancy but also fully extract the complex features of the original time series at different time scales. Compared with traditional decomposition strategies used in hybrid forecasting models, the proposed dual decomposition strategy is more significant and effective in improving the model prediction performance. (b) The combination of the X12-ARIMA-ICEEMDAN decomposition strategy and FCM-FTS forecasting model is very effective, which reasonably integrates the advantages of every single method involved. Our model performs better than benchmark models in all three cases, indicating that it is a promising tool for tourist arrivals forecasting.

The current results have important practical implications. The findings imply that more accurate predicted values of the monthly tourist arrivals to a country or region from other countries or regions can be obtained via the proposed model, which can bring about at least two benefits: (a) Taking the monthly dynamics of the future tourist arrivals as a reference for decision-making, travel agencies and other tourism-related enterprises can operate more effectively, and tourist destinations can also be managed more efficiently. (b) Based on the accurate monthly forecasts of the tourist arrivals, we can monitor the performance of the whole tourism industry and the entire economy in a real-life environment. The findings also imply that the proposed model can maintain a stable and great prediction performance when abnormal events occur. This is important in practical applications, as it is common for the data to be affected by abnormal events. Thus, the proposed model is also applicable to other fields with similar data characteristics to tourist arrivals.

In addition, our proposed hybrid model is still inadequate in qualitative research. Meanwhile, the factors related to tourism demand (such as per capita GDP, the number of air routes opened) can be considered adding into the forecasting model. In future research, qualitative analysis and quantitative prediction can be combined organically so as to achieve better prediction performance.

Appendix 1

See Tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 and 17.

Table 3 Nomenclature

Acronyms/Symbols	Description
ARMA	Autoregressive moving average
ARIMA	Autoregressive integrated moving average
SARIMA	Seasonal ARIMA
X12-ARIMA	A seasonal adjustment method developed by the United States Census Bureau
ES	Exponential smoothing
DES	Double exponential smoothing
GARCH	Generalized autoregressive conditional heteroskedasticity
ANNs	Artificial neural networks
BPNN	Backpropagation neural network
ELM	Extreme learning machine
RBF	Radial basis function
SVM	Support vector machine
SVR	Support vector regression
VMD	Variational mode decomposition
EMD	Empirical mode decomposition
EEMD	Ensemble EMD
FEEMD	Fast ensemble EMD
CEEMDAN	Complete ensemble EMD with adaptive noise
ICEEMDAN	Improved complete ensemble EMD with adaptive noise
WD	Wavelet decomposition
WT	Wavelet transform
FTS	Fuzzy time series
FCM	Fuzzy C-means
EF	Equal frequency
EW	Equal width
FLR	Fuzzy logical relationship
FLRS	Fuzzy logical relationship set
LHS	Left-hand side
RHS	Right-hand side
AE	Average error
MAPE	Mean absolute percentage error
RMSE	Root mean square error
MAE	Mean absolute error
TIC	Theil inequality coefficient
IA	Index of agreement
ADF	Augmented Dickey-Fuller
AIC	Akaike information criterion
DM	Diebold-Mariano
FE	Forecasting effectiveness
RE	Decreased relative error
GRD	Grey relational degree
r_k	The k-th residue
d_k	The k-th mode
$E_k()$	An operator which produces the k-th IMF component obtained by EMD
$M()$	An operator which produces the local mean of a series
$[]$	An operator which represents the rounding operation
$\odot$	Combination operator
x, Y	Original time series
F	Fuzzy time series

Table 3 continued

Acronyms/Symbols	Description
w^i	White noise
I	Number of times that white noise is added
β_k	A parameter that controls the energy of the white noise
K	Maximum iterations
c	Number of cluster centers
x_t, $y(t)$	Element of the original time series
n_1	Number of elements in the original time series
x_{max}, x_{min}	Maximum and minimum values in the original series respectively
v_{0i}, v_i	Initial cluster center and new cluster center respectively
u_{ij}	Element of the membership matrix
d_{ij}	Euclidean distance from the element x_j to the cluster center v_i
m	Weighted index of membership degree
ε	Iteration stop threshold
U	Universe of discourse
n_2	Number of subintervals
u_i	The i-th subinterval
A, $f_i(t)$, A_i	Fuzzy set
$f_A(\ )$	Membership function of fuzzy set A
$R(t, t-1)$	Fuzzy logical relationship
fuzzify()	An operator which fuzzifies a raw value to a fuzzy set
$f_{actual_t}(A_z)$	Degree of membership of the actual value at t under A_z
M	Number of the fuzzy sets
W_i, W_i'	Unstandardized and standardized weighting matrix elements respectively
W_s	Standardized weighting matrix
D	Defuzzified matrix
y_t, $\hat{y}_t$	Actual and predicted values at time t respectively
n_3	Size of the testing set
$\bar{y}$	Mean of the actual values

Table 4 Experimental parameter settings of the proposed model

Method	Parameter	Value
X12-ARIMA	X-11 Method	Additive
	Trend Filter	12
ICEEMDAN	Noise Standard Deviation	0.05
	Number of Realizations	100
	Maximum Iterations	1000
FCM	Membership Weighted Index	2
FTS	Number of Intervals	Number of FCM's cluster centers
	Order	1

Table 5 Detailed description of the three cases used in the experiment

Dataset	Training set	Testing set	Statistical description
Dataset A	Jan,2006-Dec,2016	Jan,2017-Dec,2017	Maximum Value:137,397
Remark: Case 1, USA	Jan,2007-Dec,2017	Jan,2018-Dec,2018	Minimum Value:60,642
	Jan,2008-Dec,2018	Jan,2019-Dec,2019	Mean Value:97,826.77
			Median Value:98,495
			Std. Dev:14,827.75
Dataset B	Jan,2006-Dec,2016	Jan,2017-Dec,2017	Maximum Value:70,102
Remark: Case 2, UK	Jan,2007-Dec,2017	Jan,2018-Dec,2018	Minimum Value:29,933
	Jan,2008-Dec,2018	Jan,2019-Dec,2019	Mean Value:44,722.08
			Median Value:41,454.5
			Std. Dev:9084.64
Dataset C	Jan,2006-Dec,2016	Jan,2017-Dec,2017	Maximum Value:30,158
Remark: Case 3, GER	Jan,2007-Dec,2017	Jan,2018-Dec,2018	Minimum Value:12,447
	Jan,2008-Dec,2018	Jan,2019-Dec,2019	Mean Value:18,513.12
			Median Value:17,060.5
			Std. Dev:4217.83

Table 6 Experimental parameter settings in different decomposition methods

Method	Parameter	Value
EEMD	Noise standard deviation	0.05
	Number of realizations	100
	Maximum iterations	1000
CEEMDAN	Noise standard deviation	0.05
	Number of realizations	100
	Maximum iterations	1000
WD	Wavelet	dmey
	Level	6

Table 7 Forecasting results of Experiment I-Comparison I

Dataset	Model	AE	MAE	RMSE	MAPE	TIC	IA
Case 1	X12-ARIMA-FCM-FTS	1149.4451	8523.0354	11,846.1432	9.5122	0.0584	0.9969
	ICEEMDAN-FCM-FTS	− 1518.9370	11,244.5504	13,726.4787	11.8851	0.0667	0.9959
	CEEMDAN-FCM-FTS	− 300.9934	12,408.9138	15,364.3685	12.6937	0.0754	0.9949
	EEMD-FCM-FTS	− 5.2026	12,927.8342	15,262.0174	13.2176	0.0746	0.9950
	WD-FCM-FTS	− 527.3479	13,881.3243	17,077.9316	14.0263	0.0830	0.9939
Case 2	X12-ARIMA-FCM-FTS	468.5103	3140.3127	4236.1895	7.2118	0.0456	0.9982
	ICEEMDAN-FCM-FTS	− 183.1214	6994.2978	9109.4909	15.4440	0.0975	0.9916
	CEEMDAN-FCM-FTS	− 421.9069	7257.1410	9354.1485	16.0508	0.0998	0.9911
	EEMD-FCM-FTS	− 820.5788	7803.0203	10,133.5340	16.7886	0.1073	0.9898
	WD-FCM-FTS	508.3707	7707.5955	10,090.4242	16.6652	0.1079	0.9903
Case 3	X12-ARIMA-FCM-FTS	95.3236	1146.7366	1763.3747	6.1537	0.0460	0.9982
	ICEEMDAN-FCM-FTS	141.9265	3860.3597	5284.7449	20.5614	0.1378	0.9839
	CEEMDAN-FCM-FTS	− 322.4869	4267.4133	5340.7023	23.6512	0.1385	0.9826
	EEMD-FCM-FTS	− 292.9302	4239.5373	5330.8620	23.4188	0.1383	0.9826
	WD-FCM-FTS	700.3850	5121.4097	6488.1702	25.8458	0.1700	0.9770

The symbol "%" after the values of MAPE criterion is omitted for all the tables

Table 8 Forecasting results of Experiment I-Comparison II

Dataset	Model	AE	MAE	RMSE	MAPE	TIC	IA
Case 1	X12-ARIMA-ICEEMDAN-FCM-FTS	**− 989.4194**	**3721.0220**	**5135.8629**	**4.2343**	**0.0250**	**0.9994**
	ICEEMDAN(R)-FCM-FTS	1119.0920	11,879.2617	14,225.9858	12.5937	0.0701	0.9955
	X12-ARIMA-FCM-FTS	1149.4451	8523.0354	11,846.1432	9.5122	0.0584	0.9969
	ICEEMDAN-FCM-FTS	− 1518.9370	11,244.5504	13,726.4787	11.8851	0.0667	0.9959
Case 2	X12-ARIMA-ICEEMDAN-FCM-FTS	**− 359.7931**	**1579.5172**	**2313.3581**	**3.4946**	**0.0225**	**0.9995**
	ICEEMDAN(R)-FCM-FTS	− 1140.0383	7100.1776	8951.6472	15.9239	0.0949	0.9918
	X12-ARIMA-FCM-FTS	468.5103	3140.3127	4236.1895	7.2118	0.0456	0.9982
	ICEEMDAN-FCM-FTS	− 183.1214	6994.2978	9109.4909	15.4440	0.0975	0.9916
Case 3	X12-ARIMA-ICEEMDAN-FCM-FTS	**− 24.8857**	**828.1133**	**1237.3211**	**4.7533**	**0.0321**	**0.9991**
	ICEEMDAN(R)-FCM-FTS	− 523.9565	4601.4322	5422.4168	25.7785	0.1402	0.9816
	X12-ARIMA-FCM-FTS	95.3236	1146.7366	1763.3747	6.1537	0.0460	0.9982
	ICEEMDAN-FCM-FTS	141.9265	3860.3597	5284.7449	20.5614	0.1378	0.9839

The results of the proposed model are highlighted in boldface for all the tables, indicating that the proposed model performs best among all the comparison models

Table 9 Forecasting results of Experiment II—Comparison I

Dataset	Model	AE	MAE	RMSE	MAPE	TIC	IA
Case 1	FCM-FTS	− 1500.0093	16,127.6028	20,978.7502	15.8852	0.1013	0.9910
	EW-FTS	2976.0361	17,606.6944	23,782.5447	16.8206	0.1166	0.9888
	EF-FTS	− 364.1944	22,522.1944	28,815.3491	22.3950	0.1388	0.9838
Case 2	FCM-FTS	43.8379	9099.4754	11,810.1853	18.9994	0.1254	0.9869
	EW-FTS	2967.2886	10,764.2758	13,585.2486	22.5976	0.1477	0.9833
	EF-FTS	236.2222	10,838.3889	13,522.3921	23.3413	0.1430	0.9832
Case 3	FCM-FTS	− 42.3331	4670.0855	6087.1866	23.5757	0.1565	0.9798
	EW-FTS	1360.6389	5079.5653	6839.1927	24.8739	0.1802	0.9757
	EF-FTS	246.7778	5382.5000	7055.6294	27.3935	0.1809	0.9739

Table 10 Forecasting results of Experiment II—Comparison II

Dataset	Model	AE	MAE	RMSE	MAPE	TIC	IA
Case 1	X12-ARIMA-ICEEMDAN-FCM-FTS	**− 989.4194**	**3721.0220**	**5135.8629**	**4.2343**	**0.0250**	**0.9994**
	X12-ARIMA-FCM-FTS	1149.4451	8523.0354	11,846.1432	9.5122	0.0584	0.9969
	ICEEMDAN-FCM-FTS	− 1598.4213	11,358.3383	13,819.6565	12.0104	0.0671	0.9959
	FCM-FTS	− 1500.0093	16,127.6028	20,978.7502	15.8852	0.1013	0.9910
Case 2	X12-ARIMA-ICEEMDAN-FCM-FTS	**− 359.7931**	**1579.5172**	**2313.3581**	**3.4946**	**0.0225**	**0.9995**
	X12-ARIMA-FCM-FTS	468.5103	3140.3127	4236.1895	7.2118	0.0456	0.9982
	ICEEMDAN-FCM-FTS	− 183.1214	6994.2978	9109.4909	15.4440	0.0975	0.9916
	FCM-FTS	43.8379	9099.4754	11,810.1853	18.9994	0.1254	0.9869
Case 3	X12-ARIMA-ICEEMDAN-FCM-FTS	**− 24.8857**	**828.1133**	**1237.3211**	**4.7533**	**0.0321**	**0.9991**
	X12-ARIMA-FCM-FTS	95.3236	1146.7366	1763.3747	6.1537	0.0460	0.9982
	ICEEMDAN-FCM-FTS	141.9265	3860.3597	5284.7449	20.5614	0.1378	0.9839
	FCM-FTS	− 42.3331	4670.0855	6087.1866	23.5757	0.1565	0.9798

The results of the proposed model are highlighted in boldface for all the tables, indicating that the proposed model performs best among all the comparison models

Table 11 Experimental parameter settings in different models

Model	Parameter	Value
SARIMA	Autoregressive term (p)	3
	Difference times (d)	1
	Moving average number (q)	0/1/1 [a]
	Seasonal autoregressive term (sp)	3
	Seasonal difference times (sd)	1
	Seasonal moving average number (sq)	0/1/0 [a]
	Seasonal order (s)	12
DES	Smoothing coefficient	0.05
BPNN	Maximum iterations	500
	Learning rate	0.001
	Activation function	Relu
	Number of input nodes	5
	Number of hidden nodes	10
	Number of output nodes	1
ELM	Number of input nodes	5
	Number of hidden nodes	10
	Number of output nodes	1
SVR	Number of input nodes	5
	Number of output nodes	1
	Gamma	0.1
	Parameter of epsilon-SVR	3
	Kernel function	RBF

[a]Indicates the values in Dataset A, B and C, respectively

Table 12 Forecasting results of Experiment III

Dataset		Proposed model	ANN		SVR	Econometric model	
			BPNN	ELM		DES	SARIMA
Case 1	AE	**− 989.4194**	1633.2399	3270.1033	− 1300.9627	− 3597.6576	− 4164.2103
	MAE	**3721.0220**	12,138.6848	14,324.6576	12,742.8694	15,520.7834	9331.9107
	RMSE	**5135.8629**	16,362.1426	16,764.6552	19,203.7714	19,193.7148	12,673.9610
	MAPE	**4.2343**	13.1258	15.0937	14.6676	17.3961	10.3133
	TIC	**0.0250**	0.0810	0.0838	0.0854	0.0928	0.0608
	IA	**0.9994**	0.9939	0.9935	0.9921	0.9914	0.9965
Case 2	AE	**− 359.7931**	883.4170	678.3240	− 840.4879	− 1229.5550	− 1829.1168
	MAE	**1579.5172**	5421.2523	6218.0035	4414.8844	8727.1745	3226.7306
	RMSE	**2313.3581**	6705.6725	7884.2968	5601.2438	10,105.3109	5228.2462
	MAPE	**3.4946**	12.0677	13.5149	10.0298	19.6042	7.6860
	TIC	**0.0225**	0.0728	0.0857	0.0596	0.1078	0.0549
	IA	**0.9995**	0.9953	0.9932	0.9968	0.9883	0.9973
Case 3	AE	**− 24.8857**	− 91.5313	4.0849	− 330.2530	− 341.5007	− 340.0379
	MAE	**828.1133**	2692.3110	2952.8525	2000.4840	4600.0830	1583.0440
	RMSE	**1237.3211**	3436.2223	3783.9989	2699.5990	5303.7222	2418.5184
	MAPE	**4.7533**	14.9194	16.0448	11.9964	25.5657	8.8485
	TIC	**0.0321**	0.0896	0.0993	0.0695	0.1387	0.0624
	IA	**0.9991**	0.9929	0.9910	0.9958	0.9807	0.9966

The results of the proposed model are highlighted in boldface for all the tables, indicating that the proposed model performs best among all the comparison models

Table 13 Forecasting results at different years

Dataset	Model	2017			2018			2019		
		MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Case 1	X12-ARIMA-ICEEMDAN-FCM-FTS	**2309.95**	**2751.42**	**2.30**	**3898.40**	**4424.40**	**3.57**	**4954.72**	**7210.11**	**6.83**
	X12-ARIMA-FCM-FTS	2932.12	3539.89	2.92	8294.39	9466.80	7.21	14,342.60	17,856.16	18.41
	ICEEMDAN-FCM-FTS	10,198.42	11,465.65	10.23	11,259.25	14,169.58	10.66	12,275.98	15,264.68	14.76
	CEEMDAN-FCM-FTS	13,050.31	15,370.93	12.97	12,113.62	14,782.26	11.07	12,062.82	15,435.98	14.04
	EEMD-FCM-FTS	13,040.84	15,861.29	12.95	12,146.62	14,804.95	11.12	13,596.04	15,100.34	15.58
	WD-FCM-FTS	15,955.20	19,301.00	15.54	13,085.79	16,681.83	11.78	12,602.98	14,971.82	14.76
	ICEEMDAN(R)-FCM-FTS	9707.06	11,573.97	9.74	12,034.27	14,884.67	11.27	13,896.45	15,862.72	16.77
	FCM-FTS	14,565.35	17,161.76	14.29	18,690.70	24,150.76	16.86	15,126.76	21,036.61	16.50
	EW-FTS	15,838.33	20,814.06	14.91	19,494.87	26,625.64	16.98	17,486.88	23,551.62	18.57
	EF-FTS	17,937.33	22,491.00	17.06	23,804.33	30,304.49	21.25	25,824.92	32,661.39	28.87
	BPNN	7072.14	8937.18	6.78	11,820.86	15,629.30	10.38	17,523.06	21,886.32	22.22
	ELM	8870.38	9958.77	8.69	15,334.78	17,742.55	13.62	18,768.81	20,716.80	22.97
	SVR	8917.88	10,793.87	8.84	9742.28	11,934.61	8.87	19,568.45	29,110.34	26.29
	DES	11,706.40	13,410.92	12.04	14,874.89	18,109.07	13.99	19,981.06	24,441.87	26.16
	SARIMA	9307.55	10,241.46	9.48	5929.86	6864.62	5.20	12,758.32	18,162.53	16.25
Case 2	X12-ARIMA-ICEEMDAN-FCM-FTS	**2012.39**	**2702.36**	**4.08**	**538.96**	**628.39**	**1.11**	**2187.20**	**2890.89**	**5.30**
	X12-ARIMA-FCM-FTS	2304.59	3158.97	4.59	2285.55	2375.06	4.80	4830.80	6181.90	12.24
	ICEEMDAN-FCM-FTS	7534.64	9898.32	16.11	7507.30	9488.63	16.16	5940.96	7806.25	14.06
	CEEMDAN-FCM-FTS	7021.49	8888.37	15.46	7220.90	9419.79	15.03	7529.03	9734.72	17.66
	EEMD-FCM-FTS	7063.85	8968.87	15.60	10,228.39	12,518.07	20.42	6116.83	8421.57	14.34
	WD-FCM-FTS	8012.02	10,914.17	17.68	9517.99	11,639.27	19.11	5592.77	7131.49	13.21
	ICEEMDAN(R)-FCM-FTS	7516.89	8689.47	16.82	7555.30	9796.34	16.21	6228.34	8301.85	14.74
	FCM-FTS	8033.00	10,686.80	17.03	9873.43	12,864.06	19.03	9391.99	11,779.21	20.95
	EW-FTS	10,697.23	12,476.35	22.42	13,349.18	17,017.74	25.92	8246.41	10,412.21	19.45
	EF-FTS	9453.00	12,079.63	20.46	10,897.42	13,759.33	21.96	12,164.75	14,605.77	27.60
	BPNN	4956.52	5882.51	10.78	4417.47	6109.35	8.45	6889.77	7935.36	16.97
	ELM	5359.77	6427.79	11.39	5755.65	7832.13	11.10	7538.59	9155.74	18.06
	SVR	4385.41	5984.61	9.26	4098.46	5006.72	8.89	4760.78	5765.33	11.94
	DES	8168.12	9352.41	17.59	8568.32	10,151.88	17.82	9445.09	10,762.14	23.41
	SARIMA	2440.96	3057.41	4.93	1241.60	1952.32	2.44	5997.63	8297.25	15.69
Case 3	X12-ARIMA-ICEEMDAN-FCM-FTS	**617.38**	**801.76**	**3.59**	**577.68**	**762.85**	**3.27**	**1289.28**	**1835.25**	**7.40**
	X12-ARIMA-FCM-FTS	669.72	911.70	3.49	723.05	886.08	3.91	2047.44	2777.07	11.07
	ICEEMDAN-FCM-FTS	3664.73	5199.38	18.76	4438.67	5781.86	23.29	3477.68	4829.31	19.64
	CEEMDAN-FCM-FTS	4084.57	4455.95	22.78	5066.78	6320.37	27.69	3650.89	5076.09	20.48
	EEMD-FCM-FTS	4060.67	4427.71	22.57	5067.10	6331.78	27.55	3590.84	5055.52	20.14
	WD-FCM-FTS	4157.27	5257.00	20.64	6274.01	7671.39	31.25	4932.94	6308.95	25.65
	ICEEMDAN(R)-FCM-FTS	4575.73	4686.92	25.53	5681.43	6541.26	31.74	3547.14	4842.78	20.06
	FCM-FTS	4584.01	5849.56	22.43	5293.18	6327.57	26.38	4133.06	6075.03	21.92
	EW-FTS	5408.30	6666.93	26.17	6462.28	8120.19	31.34	3368.12	5471.58	17.11
	EF-FTS	5204.42	6600.94	25.39	5663.92	7227.72	29.15	5279.17	7316.65	27.64
	BPNN	1898.54	2316.87	9.99	2254.77	2615.17	12.49	3923.63	4818.29	22.28
	ELM	2228.43	2645.36	11.82	2665.26	3118.67	14.47	3964.86	5121.71	21.85
	SVR	1897.00	2288.27	11.47	1627.44	2187.32	9.63	2477.00	3441.36	14.88
	DES	4470.57	4869.34	23.79	4517.03	5213.66	24.67	4812.64	5787.55	28.24
	SARIMA	920.16	1171.26	4.71	974.97	1240.90	4.97	2854.00	3825.70	16.87

The results of the proposed model are highlighted in boldface for all the tables, indicating that the proposed model performs best among all the comparison models

Table 14 Results of the DM test

Model	Case 1	Case 2	Case 3	Average
X12-ARIMA-ICEEMDAN-FCM-FTS	–	–	–	–
X12-ARIMA-FCM-FTS	2.653934***	2.242439**	1.139425	2.507414**
ICEEMDAN-FCM-FTS	4.670291***	4.118385***	3.914058***	4.214801***
CEEMDAN-FCM-FTS	4.398101***	4.181859***	4.498607***	4.462569***
EEMD-FCM-FTS	4.749956***	3.782570***	4.433012***	4.198911***
WD-FCM-FTS	4.158481***	3.854672***	4.391467***	3.914504***
ICEEMDAN(R)-FCM-FTS	5.484243***	4.552637***	5.114293***	4.993597***
FCM-FTS	3.600406***	4.107028***	3.931071***	3.564915***
EW-FTS	3.504539***	3.697898***	3.927137***	3.370852***
EF-FTS	3.514911***	4.494460***	3.841272***	3.564836***
BPNN	3.717957***	4.432075***	3.408091***	4.214143***
ELM	5.058495***	3.826229***	3.023245***	4.296399***
SVR	2.084516**	4.259447***	2.443427**	2.288896**
DES	4.705974***	5.892422***	5.583836***	5.901837***
SARIMA	2.276890**	1.999582**	1.815357*	2.060152**

*Indicates the 10% significance level. $Z_{0.1/2} = 1.64$

**Indicates the 5% significance level. $Z_{0.05/2} = 1.96$

***Indicates the 1% significance level. $Z_{0.01/2} = 2.58$

Table 15 Forecasting effectiveness of different models

Model	Case 1		Case 2		Case 3	
	1st-Order	2nd-Order	1st-Order	2nd-Order	1st-Order	2nd-Order
X12-ARIMA-ICEEMDAN-FCM-FTS	**0.957657**	**0.905740**	**0.965054**	**0.928907**	**0.952467**	**0.903511**
X12-ARIMA-FCM-FTS	0.904878	0.802250	0.927882	0.857847	0.938463	0.869447
ICEEMDAN-FCM-FTS	0.881149	0.797518	0.845560	0.732314	0.794386	0.651262
CEEMDAN-FCM-FTS	0.873063	0.787227	0.839492	0.730831	0.763488	0.631917
EEMD-FCM-FTS	0.867824	0.794830	0.832114	0.723141	0.765812	0.633374
WD-FCM-FTS	0.859737	0.778830	0.833348	0.722732	0.741542	0.610812
ICEEMDAN(R)-FCM-FTS	0.874063	0.792455	0.840761	0.730092	0.742215	0.610094
FCM-FTS	0.841148	0.735739	0.810006	0.694707	0.764243	0.635458
EW-FTS	0.831794	0.718623	0.774024	0.661364	0.751261	0.608494
EF-FTS	0.776050	0.641721	0.766587	0.640162	0.726065	0.588717
BPNN	0.868742	0.744739	0.879323	0.796119	0.850806	0.744600
ELM	0.849063	0.754211	0.864851	0.776378	0.839552	0.737863
SVR	0.858116	0.688485	0.899702	0.821050	0.880036	0.779034
DES	0.826039	0.689536	0.803958	0.708553	0.744343	0.634295
SARIMA	0.896867	0.794427	0.923140	0.820348	0.911515	0.812901

The results of the proposed model are highlighted in boldface for all the tables, indicating that the proposed model performs best among all the comparison models

Table 16 Improvement percentage of the proposed model relative to the comparison models

Model	Case 1			Case 2			Case 3		
	RE_{MAE}	RE_{RMSE}	RE_{MAPE}	RE_{MAE}	RE_{RMSE}	RE_{MAPE}	RE_{MAE}	RE_{RMSE}	RE_{MAPE}
X12-ARIMA-FCM-FTS	56.34	56.65	55.49	49.70	45.39	51.54	27.79	29.83	22.76
ICEEMDAN-FCM-FTS	66.91	62.58	64.37	77.42	74.60	77.37	78.55	76.59	76.88
CEEMDAN-FCM-FTS	70.01	66.57	66.64	78.23	75.27	78.23	80.59	76.83	79.90
EEMD-FCM-FTS	71.22	66.35	67.96	79.76	77.17	79.18	80.47	76.79	79.70
WD-FCM-FTS	73.19	69.93	69.81	79.51	77.07	79.03	83.83	80.93	81.61
ICEEMDAN(R)-FCM-FTS	68.68	63.90	66.38	77.75	74.16	78.05	82.00	77.18	81.56
FCM-FTS	76.93	75.52	73.34	82.64	80.41	81.61	82.27	79.67	79.84
EW-FTS	78.87	78.40	74.83	85.33	82.97	84.54	83.70	81.91	80.89
EF-FTS	83.48	82.18	81.09	85.43	82.89	85.03	84.61	82.46	82.65
BPNN	69.35	68.61	67.74	70.86	65.50	71.04	69.24	63.99	68.14
ELM	74.02	69.36	71.95	74.60	70.66	74.14	71.96	67.30	70.37
SVR	70.80	73.26	71.13	64.22	58.70	65.16	58.60	54.17	60.38
DES	76.03	73.24	75.66	81.90	77.11	82.17	82.00	76.67	81.41
SARIMA	60.13	59.48	58.94	51.05	55.75	54.53	47.69	48.84	46.28

The symbol "%" after all the values is omitted

Table 17 Grey relational degree of different models

Model	Case 1	Case 2	Case 3	Average
X12-ARIMA-ICEEMDAN-FCM-FTS	**0.918375**	**0.929712**	**0.924735**	**0.924274**
X12-ARIMA-FCM-FTS	0.840965	0.870954	0.901525	0.871148
ICEEMDAN-FCM-FTS	0.794429	0.767043	0.752323	0.771265
CEEMDAN-FCM-FTS	0.780905	0.759847	0.720169	0.753640
EEMD-FCM-FTS	0.768742	0.748393	0.722066	0.746400
WD-FCM-FTS	0.761839	0.751450	0.688271	0.733853
ICEEMDAN(R)-FCM-FTS	0.784866	0.760895	0.693737	0.746500
FCM-FTS	0.743810	0.722874	0.709432	0.725372
EW-FTS	0.735016	0.682705	0.704748	0.707490
EF-FTS	0.677247	0.684495	0.685846	0.682529
BPNN	0.793379	0.798865	0.794739	0.795661
ELM	0.749097	0.780324	0.779895	0.769772
SVR	0.791299	0.829610	0.839304	0.820071
DES	0.744194	0.711482	0.690905	0.715527
SARIMA	0.826558	0.877935	0.872141	0.858878

The results of the proposed model are highlighted in boldface for all the tables, indicating that the proposed model performs best among all the comparison models

Appendix 2: Supplementary material

Supplementary material related to this article (including the data and the code) can be found online at https://github.com/WuZK96/X12-ARIMA-ICEEMDAN-FCM-FTS.

Declarations

References

1. Song HY, Li G, Witt SF, Athanasopoulos G (2011) Forecasting tourist arrivals using time-varying parameter structural time series models. Int J Forecast 27:855–869

2. Gunter U, Onder I (2015) Forecasting international city tourism demand for Paris: accuracy of uni- and multivariate models employing monthly data. Tour Manage 46:123–135

3. Chu FL (2008) Forecasting tourism demand with ARMA-based methods. Tour Manage 30:740–751

4. Hossein H, Emmanuel SS, Nikolaos A, George F, Rangan G (2017) Forecasting accuracy evaluation of tourist arrivals. Ann Tour Res 63:112–127

5. Yao Y, Cao Y, Ding XM, Zhai J, Liu JX, Luo YL, Ma S, Zou KL (2018) A paired neural network model for tourist arrival forecasting. Expert Syst Appl 114:588–614

6. Pan B, Wu CD, Song H (2012) Forecasting hotel room demand using search engine data. J Hosp Tour Technol 3:196–210

7. Chu FL (2011) A piecewise linear approach to modeling and forecasting demand for Macau tourism. Tour Manage 32:1414–1420

8. Christine L, Michael MA (2001) Forecasting tourist arrivals. Ann Tour Res 28:965–977

9. Liang YH (2014) Forecasting models for Taiwanese tourism demand after allowance for Mainland China tourists visiting Taiwan. Comput Ind Eng 74:111–119

10. Wu LJ, Cao GH (2016) Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow. Knowledge-Based Syst 110:157–166

11. Law R, Li G, Fong DKC, Han X (2019) Tourism demand forecasting: a deep learning approach. Ann Touris Res 75:410–423

12. Xu Y, Yang W, Wang J (2017) Air quality early-warning system for cities in China. Atmos Environ 148:239–257

13. Li JC, Zhu SW, Wu QQ (2019) Monthly crude oil spot price forecasting using variational mode decomposition. Energy Econ 83:240–253

14. Wang J, Liu F, Song Y, Zhao J (2016) A novel model: Dynamic choice artificial neural network (DCANN) for an electricity price forecasting system. Appl Soft Comput 48:281–297

15. Song H, Li G (2008) Tourism demand modelling and forecasting-A review of recent research. Tour Manage 29:203–220

16. Sun SL, Wei YJ, Tsui KL, Wang SY (2019) Forecasting tourist arrivals with machine learning and internet search index. TourManage 70:1–10

17. Claveria O, Monte E, Torra S (2016) Combination forecasts of tourism demand with machine learning models. Appl Econ Lett 23:428–431

18. Yu L, Zhao YQ, Tang L, Yang ZB (2019) Online big data-driven oil consumption forecasting with Google trends. Int J Forecast 35:213–223

19. Xiao L, Shao W, Wang C, Zhang K, Lu H (2016) Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting. Appl Energy 180:213–233

20. Ma XJ, Jin Y, Dong QL (2017) A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting. Appl Soft Comput 54:296–312

21. Bisoi R, Dash PK, Mishra SP (2019) Modes decomposition method in fusion with robust random vector functional link network for crude oil price forecasting. Appl Soft Comput 80:475–493

22. Fan CD, Ding CK, Zheng JH, Xiao LY, Ai ZY (2020) Empirical mode decomposition based multi-objective deep belief network for short-term power load forecasting. Neurocomputing 388:110–123

23. Bento PMR, Pombo JAN, Calado MRA, Mariano SJPS (2019) Optimization of neural network with wavelet transform and improved data selection using bat algorithm for short-term load forecasting. Neurocomputing 358:53–71

24. Liang XZ, Guo ZK, Zhang QW, Yang MG, Wang SY (2020) An analysis and decomposition ensemble prediction model for air passenger demand based on singular spectrum analysis. Syst Eng Theor Pract 40:1844–1855

25. Wang JZ, Yang WD, Du P, Li YF (2018) Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. Energy 148:59–78

26. Yaslan Y, Bican BR (2017) Empirical mode decomposition based denoising method with support vector regression for time series prediction: a case study for electricity load forecasting. Measurement 103:52–61

27. Jiang P, Ma XJ (2016) A hybrid forecasting approach applied in the electrical power system based on data preprocessing, optimization and artificial intelligence algorithms. Appl Math Model 40:10631–10649

28. Yang Z, Ce L, Lian L (2017) Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. Appl Energy 190:291–305

29. Colominas MA, Schlotthauer G, Torres ME (2014) Improved complete ensemble EMD: a suitable tool for biomedical signal processing. Biomed Signal Process Control 14:19–29

30. Jiang P, Yang HF, Heng JN (2019) A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting. Appl Energy 235:786–801

31. U.S. Census Bureau. X-12-ARIMA reference manual version 0.3. http://www.census.gov/srd/www/sapaper/.

32. Dominique L, Benoit Q (2004) Implementation issues on shrinkage estimators for seasonal factors within the X-11 seasonal adjustment method. Int J Forecast 20:557–560

33. Shiskin, J.; Young, A. H.; Musgrave, J. C. The X-11 variant of the census method II seasonal adjustment program. *Technical Paper No.15, US Department of Commerce, Bureau of the Census*, **1967**.

34. Wu ZH, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1:1–41

35. Lin H, Sun Q, Chen SQ (2020) Reducing exchange rate risks in international trade: a hybrid forecasting approach of CEEMDAN and multilayer LSTM. Sustainability 12:2451

36. Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Adv Appl Pattern Recognit 22:203–239

37. Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353

38. Song Q, Chissom BS (1993) Forecasting enrollments with fuzzy time series—Part I. Fuzzy Sets Syst 54:1–9

39. Song Q, Chissom BS (1994) Forecasting enrollments with fuzzy time series—Part II. Fuzzy Sets Syst 62:1–8

40. Cheng CH, Yang JH (2018) Fuzzy time-series model based on rough set rule induction for forecasting stock price. Neurocomputing 302:33–45

41. Fan XS, Wang YN, Zhang MY (2020) Network traffic forecasting model based on long-term intuitionistic fuzzy time series. Inf Sci 506:131–147

42. Bose MH, Mali K (2019) Designing fuzzy time series forecasting models: a survey. Int J Approx Reason 111:78–99

43. Chen SM (1996) Forecasting enrollments based on fuzzy time series. Fuzzy Sets Syst 81:311–319

44. Said SE, Dickey DA (1984) Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71:599–607

45. Wang JZ, Li HM, Lu HY (2018) Application of a novel early warning system based on fuzzy time series in urban air quality forecasting in China. Appl Soft Comput 71:783–799

46. Tian CS, Hao Y, Hu J (2018) A novel wind speed forecasting system based on hybrid data preprocessing and multi-objective optimization. Appl Energy 231:301–319

47. Barbeito I, Zaragoza S, Tarrío SJ, Naya S (2017) Assessing thermal comfort and energy efficiency in buildings by statistical quality control for autocorrelated data. Appl Energy 190:1–17

48. Diebold FX, Mariano RS (1995) Comparing predictive accuracy. J Bus Econom Statist 13:253–263

49. Sun S, Qiao H, Wei Y, Wang S (2017) A new dynamic integrated approach for wind speed forecasting. Appl Energy 197:151–162