Forecasting Tourism Demand by a Novel Multi-Factors Fusion Approach

HONGWEI WANG AND WENZHENG LIU
School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China
Corresponding author: Wenzheng Liu (liuwz03@163.com)

ABSTRACT

The volatility of tourism demand is often caused by some irregular events in recent years. Typically, inbound tourists are quite sensitive to various factors, including the exchange rate fluctuation, consumer price index, personal or household income or consumption expenditure. We combine these multivariate time series data onto an ingenious multi-factor fusion strategy to contribute to precise tourism demand forecasting. A novel hybrid deep learning forecasting approach is developed by integrating several modules such as improved complete ensemble empirical mode decomposition with adaptive noise, intrinsic mode functions classification, multi-factors fusion and predictors matching. The monthly tourist flow data of Shanghai inbounding from USA, Korea and Japan are conducted to verify the performance of the proposed approach, which outperforms all benchmark models for different prediction horizons. The experimental results show that introducing external influencing factors can improve the prediction accuracy significantly, and therefore confirm the rationality and validity of the proposed approach.

INDEX TERMS
Tourism demand forecasting, intrinsic mode functions classification, multi-factor fusion, bidirectional gated recurrent unit, support vector regression.

I. INTRODUCTION

In the last few decades, these have been a surge of interest in the research of tourism demand among academics and practitioners. Although the tourism flow is growing steadily in the medium and long term, its rapid fluctuation is usually caused by various uncertain obstacles in the short term, such as financial crisis, regional conflict and epidemic disease [1], [2]. These obstacles pose a great challenge to short term forecasting. Moreover, the perishable nature of tourism products, including vacant hotel rooms, unsold tourism services and surplus employees, can lead to an inevitable financial loss without the rational market demand guidance [3], [4]. From a macro perspective, scientific tourism demand forecasting can provide a reliable reference for tourism related departments to formulate tourist policies and sustainable development strategic plans; from a micro perspective, it can guide relevant tourism enterprises to make development plans and allocate the tourism resources rationally [5]. Therefore, to achieve reasonable analyses and credible forecasting of short term tourism demand has always been a goal strive for [6].

As we all know, the change of tourism demand is affected by many factors. As stated by Song et al. [7], the numerous influence factors with respect to both origin and destinations, such as tourist income level, tourism price in a destination or the competing destinations, exchange rate, are the main factors corresponding to tourism demand analysis, and this view is also supported by Jiao et al. [8]. Furthermore, many researchers also believe that some crucial influence factors can be used to establish potential map relationships to make up for the insufficiency of predictors, so as to achieve the purpose of improving the prediction accuracy [9], [10]. Notably, the prices of competing European city destinations and tourist income [1], the exchange rate and tourism price in competing destinations [11] and other similar influencing factors have been employed to forecast tourism demand. Typically, Ulrich et al. [1] introduced the price of competing European city destinations and tourist income as the influencing factor and combined them with the Bayesian VAR model. Jiao et al. [11] used the exchange rate and tourism price in competing destinations to forecast the inbound visitor arrivals.
to 26 European cities. However, due to the low availability and time difference of influence factors data, the extraction and utilization efficiency is poor. Many existing researches combined with influence factors analyze mainly the annual changes of tourism demand, and most predictors used are relatively simple and less precision. Therefore, this study constructs a set of composite influence factors based on exchange rate, consumer price index (CPI) and others to analyze monthly fluctuations of tourist arrivals. The aim is to mine and capture the correlation between the inherent components of historical data and various influence factors. Based on this, the following problem is to regulate the difficulty of prediction task. Undoubtedly, introducing exogenous variables is necessary for improving the prediction accuracy, but it is bound to increase the systematic errors of prediction approach. How to solve this technical problem of multi-factors fusion and balance the relationship between them is the primary goal of this paper.

In general, for univariate data (i.e. historical tourism demand data), most of the current researches attempt to denoise the tourist flow time series and transform irregular fluctuation into a smooth form, which is easy to capture and understand, for the purpose of improving the forecasting performance [12], [13]. In the field of short-term time series prediction, one of the popular data preprocessing technologies is extracting the empirical mode features according to different time scales by empirical mode decomposition (EMD) [14]. Then, its related derivative technologies came into being, such as the wavelet decomposition (WD) [15], ensemble EMD (EEMD) [16], [17] and complete EEMD with adaptive noise (CEEMDAN) [2]. Their main theoretical principle is extracting intrinsic mode functions (IMFs) according to the frequencies and time scale of data, which can approximately represent the law and trend of data at different stages. For a single tourist historical time series, its processed goal is to explore the potential periodic rules, trends and complex changes within the data. However, some emergencies and external factors often have a significant impact on the tourism market that can not be ignored. And these uncertain events (white noise) adversely affect the internal development of the tourism industry, and bring difficulties in grasping the future trend of tourist flow [6]. Although introducing influence factors into the decomposition strategy to explain these outliers (i.e. emergencies) can effectively grasp the dynamic of tourists flow and improve the accuracy of prediction results, a new issue subsequently arises on how to fuse the influence factors into tourism data by the preprocessing methods.

Generally speaking, the tourism demand data contains regular (trends, cycles, and seasons) and irregular components. These irregular components of the influence factors (such as the change in tourist income, exchange rate, and travel intention) are the main factors causing the fluctuation of tourist arrivals. To solve the issue of multi-source data fusion, this study innovatively proposes a decomposition ensemble learning framework integrating the influence factors into the irregular components of tourism demand data. Usually, forecasting the most complex feature components is regarded as the most difficult challenge and as the key to improving forecasting accuracy. Existing studies show that the irregularity contained in these IMFs is generally caused by economic development, policy orientation, international situation, and other factors [18]. Given the importance and availability of these factors in prediction, this paper investigates the impact of certain economic factors on tourism passenger flow. First of all, the fluctuation of the exchange rate produces the most direct impact on the travel cost, especially for overseas tourists [19]. For example, the currency devaluation leads to an increase in the cost of studying and shopping abroad, but for foreign tourists, the cost of international travel is reduced and their willingness to journey is enhanced. Secondly, the consumer price index (CPI) is a macroeconomic indicator reflecting changes in household consumption levels over a certain period. It reflects not only the economic growth capacity of a country or region, but also the willingness of residents to consume and travel [20]. The third factor to be considered is the personal or household consumption expenditure or income, which plays a similar role as CPI and includes a large proportion of spending on transport and travel [21]. It reflects short-term intentions of consumers without delay, while there is a time lag in CPI. Considering these three factors causing short-term tourism demand to fluctuate, this study combines them with the IMFs containing the complex irregularity, thereby excavating their internal correlation by machine learning (ML).

At present, the most popular ML is the deep learning model, which mainly adopts large-scale distributed parallel processing and extends the shallow neural network to more than two nonlinear mapping layers, by imitating the functioning of human brain [22]. And its effectiveness in various situations has been verified by many researchers [23], [24]. Thereinto, the bidirectional gated recurrent unit (BiGRU) can obtain a stronger self-learning function bases on solving the problem of gradient disappearance under long-term information retention and be considered the best predictor for mining potential dependence on historical data [25]. Hence, the superior BiGRU model is employed to explore the relationship between potential factors, which is used as a bridge for multifactor fusion in this study. Usually, for each IMF, various ML models can be adopted as predictors, such as auto-regressive integrated moving average (ARIMA) [16], nonlinear autoregressive neural network with exogenous input (NARX) [15], and back-propagation neural network (BPNN) [14]. But with these developments, each predictor assessing the tourism demand has its advantages and drawbacks. Although some of them have been widely used in the empirical analysis according to their respective strengths, no single model can deal with all types of IMF components. Therefore, it is necessary to design a model matching strategy for selecting appropriate predictors based on the characteristics of each component [3], [8]. For the low-complexity components, some studies have found that they show two characteristics, namely trend and
ARIMA models are employed to forecast the remaining two components (LCLC) and low-complexity high-correlation (LCHC) components, respectively. Finally, the final result is obtained by summing up all the individual predictions. To verify the effectiveness of the proposed approach, an empirical experiment investigating the monthly tourist arrivals data from three main origins (USA, Korea and Japan) to Shanghai is conducted to analyze the related issues.

II. METHODOLOGY

In this section, we first introduce the decomposition algorithm, the classification method, and the prediction model. Then, the procedure of the proposed ensemble learning approach is detailedly presented. Finally, the tourism demand forecasting model is described.

A. IMPROVED COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE (ICEEMDAN)

As a derivative of empirical mode decomposition (EMD), the improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) is developed by the Colominas et al. [30]. The EMD is an empirical signal decomposition method introduced to analyze nonlinear and nonstationary signals, which decomposes a signal \( x(t) \) into a quantity of intrinsic mode functions (IMFs). As an IMF, a signal \( x(t) \) must satisfy two conditions: (1) the number of extrema and the number of zero crossing must be equal or differ at most by one; and (2) the mean value of the upper and lower envelope is zero everywhere [31]. But there exists the challenge of mode mixing, which means the existence of similar oscillations in different modes or the presence of disparate oscillations in a same mode. To overcome these defects, ensemble empirical mode decomposition (EEMD), an upgraded version of EMD, is performed with a noise-assisted signal analysis method based on the EMD framework [32]. The IMF component of the EEMD is defined as the mean of the corresponding IMFs obtained via EMD over a series of trials, produced by adding different realizations of white noise with finite variance to the raw signal \( x(t) = \mu[n] + \omega[n] \). Yet the very problem with it is that the reconstruction of signals contaminated by different noises may lead to different numbers of modes, and the reconstructed signal contains residual noise. The CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise), as a variation of the EEMD algorithm, is proposed to address these issues, which can achieve a negligible reconstruction error and solving the problem of different number of modes for different realizations of signal plus noise. In spite of that, the CEEMDAN still has some problems about residual noise and spurious modes. The ICEEMDAN is proposed to obtain an exact reconstruction of the original signal and a better spectral separation of the modes by introducing the operators \( E_k(\cdot) \) (producing the \( k \)th mode by using EMD) and \( M(\cdot) \) (producing the local mean of the applied signal). For a new signal \( x(t) \), two operators meet the conditions of \( E_k(x(t)) = x - M(x) \). The primary steps of ICEEMDAN are described as follows:
1) Get N new signal \( x^{(n)}(t) = x + \beta_0E_1(w^{(n)}(t)) \) by decomposing \( n(1, \cdots, N) \) Gaussian white noise \( w^{(n)} \) using EMD, where \( \beta_0 = \epsilon_0 \text{std}(x)/\text{std}(E_1(w^{(n)})) \). And the first residue \( r_1 \) is obtained by using operator \( M(\cdot) \) and performing the action of averaging \( \langle \cdot \rangle \) on \( N \) new signal:

\[
    r_1 = \langle M(x^{(n)}) \rangle \quad (1)
\]

2) When \( k = 1 \), the first decomposed mode \( \text{imf}_1 \) is calculated:

\[
    \text{imf}_1 = x - r_1 \quad (2)
\]

3) When \( k = 2 \), the second residue is evaluated by averaging the local mean of \( r_1 + \beta_1E_2(w^{(n)}) \), where the adaptive adjustment parameter of Gaussian white noise is \( \beta_k = \epsilon_k \text{std}(r_k), \ k = 1, \cdots, K \). The second decomposed mode is defined as follows:

\[
    \text{imf}_2 = r_1 - r_2 = r_1 - \langle M(r_1 + \beta_1E_2(w^{(n)})) \rangle \quad (3)
\]

4) The rest residue \( r_k(k = 3, \cdots, K) \) can be calculated in the same manner:

\[
    r_k = \langle M(r_{k-1} + \beta_{k-1}E_k(w^{(n)})) \rangle \quad (4)
\]

5) The \( k \)th decomposed mode \( \text{imf}_k \) is obtained by:

\[
    \text{imf}_k = r_{k-1} - r_k = r_{k-1} - \langle M(r_{k-1} + \beta_{k-1}E_k(w^{(n)})) \rangle \quad (5)
\]

6) Return to step 4 until all the intrinsic mode functions conditions are met. The final residue satisfies:

\[
    r_n = x - \sum_{k=1}^{K} \text{imf}_k \quad (6)
\]

Therefore, a signal \( x \) can be expressed as \( x = \sum_{k=1}^{K} \text{imf}_k + r_n \). In the field of time series, the ICEEMDAN has achieved a better spectral separation of the IMFs, and at the same time, it can mine effectively the intrinsic information with different features and reduce the difficulty of prediction tasks. Therefore, the ICEEMDAN, as a decomposition tool, is used to decompose the tourism demand time series in this study.

**B. INTRINSIC MODE FUNCTIONS (IMFs) CLASSIFICATION METHOD**

For the different intrinsic mode functions (IMFs), we first employ range entropy (RE) to evaluate the complexity of time series. Then, the least absolute shrinkage and selection operator (LASSO) is used to filter both the high-complexity components and the preselected influence factors, and the correlation coefficients between the low-complexity components and the original time series are calculated by the maximal information coefficient (MIC) method for further partition. The application and introduction of specific methods are as follows.

1) RANGE ENTROPY (RE)

Approximate entropy (AE) and sample entropy (SE) are the most commonly used methods for measuring complexity in modern signal processing (especially time series analysis). In fact, SE is derived from AE, which can overcome some limitations of the former, limitations including the inconsistency of the embedding dimension \( m \), tolerance \( r \) and the strong dependence on the length of input signal. However, AE and SE are susceptible to changes in signal amplitude. Therefore, RE is proposed to solve the problem of insensitivity and alleviate the need for amplitude correction simultaneously [33]. In this study, RE is used to test the complexity of time series, and the embedding dimension and tolerance are set to 5 and 0.8, respectively.

2) LEASTABSOLUTE SHRINKAGE AND SELECTION OPERATOR (LASSO)

The LASSO is a compression estimation technique for dimensionality reduction of high-dimensional data [34]. Its core principle is to construct a penalty function that forces the absolute value of the parameter to be estimated within a preset threshold, thus achieving the effect of compression model regression coefficients. A smaller threshold can compress some regression coefficients to 0 and eliminate such variables with a coefficient of 0, thereby realizing the effect of variable filter. The LASSO method is equivalent to adding the penalty term \((L_1 \text{ norm})\) to the ordinary linear model:

\[
    \hat{\beta}_{\text{LASSO}} = \arg \min_{\beta \in \mathbb{R}^d} \left( ||Y - X\beta||^2 + \lambda \sum_{j=1}^{d} |\beta_j| \right)
\]

where the concomitant variable \( X = (X^{(1)}, \cdots, X^{(n)}) \), the response variable \( Y = (y_1, \cdots, y_n)^T \), and the regression coefficient \( \beta = (\beta_1, \cdots, \beta_n) \).

Equivalently,

\[
    \hat{\beta}_{\text{LASSO}} = \arg \min_{\beta \in \mathbb{R}^d} \left( ||Y - X\beta||^2 \right)
\]

s.t. \( \sum_{j=1}^{d} |\beta_j| \leq t \)  \( (8) \)

where \( t \) and \( \lambda \) are adjustment coefficients.

Let \( t_0 = \sum_{j=1}^{d} |\beta(\text{OLS})| \), when \( t < t_0 \), the regression coefficients of variables \( X \) will be compressed to 0, resulting in sparseness and dimension reduction. Therefore, the adjustment coefficient affects the effect of variable filtering. In this study, we use a combination of grid search method, LASSO with SVR to determine the optimal adjustment coefficient according to the optimal principle of the training set effect. The grid search and LASSO methods are performed by an interface from the Python library, Scikit-learn (0.23.2).

3) MAXIMAL INFORMATION COEFFICIENT (MIC)

In regard to the calculation of correlation, MIC is used to analyze or measure the linear or nonlinear strength between each IMF and the original series in this paper. The MIC
can not only find the linear function relationship between variables, but also get the nonlinear function relationship (such as exponent, period). The larger the MIC value, the higher the correlation is. The paired data \(\{X_m; Y\}\) consisting of the \(m\)-th decomposed mode \(X_m\) and the original data \(Y\), \(MIC(X_m; Y)\) is defined as:

\[
MIC(X_m; Y) = \max_{|X_m||Y|<B} \frac{I(X_m; Y)}{\log_2(\min(|X_m|, |Y|))}
\]

where \(I(X_m; Y)\) is the value of mutual information between the two sequences. The detailed algorithm is also available in Reshef et al. [35].

C. FORECASTING MODELS

Each prediction model has its advantages and disadvantages. To improve the forecasting accuracy of tourism demand, it is necessary to select a suitable predictor for each component based on the strength of different models and the data feature. In this paper, bidirectional gated recurrent unit (BiGRU) is used to forecast the high-complexity components, while ARIMA and SVR models are applied to fit the future trend of low-complexity low-correlation (LCLC) and low-complexity high-correlation (LCHC) components, respectively.

1) AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

The ARIMA model is one of the most common regression models in the time series forecasting. ARIMA\((p, d, q)\) is developed by combing the autoregressive (AR\((p)\)), moving average (MA\((q)\)) and difference degree \((d)\). For a stationary time series \(x_t(t = 1, \ldots, T)\), the mathematical formula of the ARIMA model can be described as follows:

\[
x_t = \alpha_1 x_{t-1} + \cdots + \alpha_p x_{t-p} + \delta + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q}
\]

where \(\alpha\) and \(\beta\) are the lag parameters of AR and MA part of the model, respectively, \(\epsilon_i\) represents the error term, and \(\delta\) is the constant term.

There are three iterative stages in building an ARIMA model. At the first stage, the time series needs to be stabilized. For an unstable time series, it usually requires to be stabilized through difference disposal of \(d\) times. At the second stage, the model parameters are estimated based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC). The last stage covers the diagnostic checking of residual analysis, time series prediction and evaluation. In this study, the ARIMA model is performed by using a Python library, Statsmodels (v0.12.1).

2) SUPPORT VECTOR REGRESSION (SVR)

The SVR model, as a supervised learning technique, is developed by Sain et al. [36]. Based on the principle of structural risk minimization, SVR uses a kernel function to detect the best regression hyperplane in a high-dimensional feature space that can balance learning ability and generalization ability. Compared with other models, SVR has the advantages of fast convergence and multidimensional function estimation. For a set of learning sample \((D = x_i, y_i, \text{where input variables are } x_i \in \mathbb{R}^m, \text{output value is } y_i \in \mathbb{R}^m, \text{and } m \text{ is the number of samples})\), its function relationship \(f(x)\) can be presented as:

\[
f(x) = \omega^T \phi(x_i) + b
\]

where \(\omega\) and \(b\) denote the weight coefficient vectors and a bias constant, respectively. The identity function \(\phi(x)\) can map \(x_i\) to a high-dimensional feature space.

Typically, the kernel function of SVR is employed to dispose the mapping relationship in high dimensional feature space. In this study, the linear function \(\langle x, x' \rangle\) is the most suitable kernel function for the decomposed components of tourism demand time series due to its high computational efficiency. Therefore, the original problem can be equivalently expressed as:

\[
\min_{\omega, b} \frac{1}{2} \omega^T \omega + C \sum_{i=1} \max(0, |y_i - (\omega^T \phi(x_i) + b)| + \epsilon)
\]

where \(C\) is the penalty parameter. The samples that differ from the actual target will be penalized by at least \(\epsilon\). The fitting accuracy and generalization performance of SVR are generally affected by \(C\) and \(\epsilon\). Hence, the cross-validation grid-search is used to optimize the two parameters of these methods over a parameter grid. Specifically, a Python library, Scikit-learn (0.23.2) is adopted to implement SVR model, grid search and cross validation methods.

3) BIDIRECTIONAL GATED RECURRENT UNIT (BiGRU)

As a variant of the long short-term memory (LSTM), the gated recurrent unit (GRU) has fewer parameters and faster convergence while maintaining the good learning performance of LSTM. Differing from LSTM, GRU replaces the input gate and forgetting gate of LSTM with update gate, and its internal structure consists of update gate and reset gate (Figure 1). Among them, the update gate \((r_t)\) controls the influence degree of output hidden layer at the previous moment on the current moment. The larger the update gate value, the higher the influence level. The reset gate \((z_t)\) decides the neglect degree of output information of hidden layer at the previous moment, the less information is ignored with this larger value. At \(t\) time, the hidden state \((h_t)\) of the GRU can be calculated as follows:

\[
r_t = \sigma(W_r x_t + U_r h_{t-1})
\]

\[
z_t = \sigma(W_z x_t + U_z h_{t-1})
\]

\[
\tilde{h}_t = \tanh(W x_t + U h_{t-1})
\]

\[
h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t
\]

where \(\sigma\) and \(\tanh\) are the sigmoid and the tanh activation function, respectively. \(W_r, W_z, W_x, U_r, U_z, \text{and } U\) are the weight matrices of training sample. \(\tilde{h}_t\) is named the candidate vector, it is used as a modulation operation to control the degree of receiving new input information in the cell state.
BiGRU is developed based on GRU concept in the forward and backward transmission of hidden layers, so that the output layer includes the information from both previous and following states. Its basic structure is shown in Figure 2. The idea of BiGRU is that the current hidden state \( h_t \) is determined by the forward hidden output \( \hat{h}_t \) in positive time direction, the reverse hidden output \( \hat{h}_t^{-1} \) in negative time direction and the current input information \( x_t \). It can be described as follows:

\[
\begin{align*}
\hat{h}_t &= GRU(x_t, \hat{h}_{t-1}) \\
\hat{h}_t^{-1} &= GRU(x_t, \hat{h}_{t-1}) \\
h_t &= \alpha_t \hat{h}_t + \beta_t \hat{h}_t^{-1} + b
\end{align*}
\]

where \( \alpha_t \) and \( \beta_t \) represent the weight of \( \hat{h}_t \) and \( \hat{h}_t^{-1} \) corresponding to BiGRU at \( t \) time, respectively. \( b \) is the bias of \( h_t \) at \( t \) time. This study constructs the BiGRU network architecture based on the Python library, Pytorch (1.6.0).

D. PROCEDURE OF THE PROPOSED MODEL

Generally, exchange rate, tourist income, consumer price index (CPI) can cause short-term changes in tourists traveling intention. To study the changes of overseas tourists arrival in Shanghai, a novel multi-factor hybrid ensemble learning framework is established to forecast the monthly tourist overnights flow from three major origins (the USA, Japan, and Korea). This hybrid approach integrates decomposition ensemble strategy and multi-factors fusion to achieve the purpose of capturing short-term fluctuations. The overall framework is shown in Figure 3. It consists of five fundamental steps: data decomposition, intrinsic mode functions classification, multi-factors fusion, individual prediction and ensemble results.

Step 1:

1) Historical data decomposition. By using ICEEMDAN, the tourism demand time series are decomposed into several intrinsic mode functions with different frequencies and characteristics (denoted as \( IMF_i \), \( i = 1, 2, \cdots, n \)).

2) Intrinsic mode functions (IMFs) classification. The obtained IMFs by ICEEMDAN are firstly classified into the high-complexity (HC) and low-complexity (LC) components according to range entropy (RE) testing results. Then, by calculating the correlation between each low-complexity component and the original time series by Maximal information coefficient (MIC), LC components are divided into two categories, namely, low-complexity and high-correlation (LCHC) components and low-complexity and low-correlation (LCLC) components.

3) Multi-factors fusion. The HC components \( Y \) are usually caused by various short-term change of influence factors. Here, the preselected three influence factors \( X \) including exchange rate, consumer price index (CPI) and personal or household consumption expenditure or income are combined with each HC component to generate an information matrix with 4 rows and \( n \) columns, where 4 represents the number of HC and \( X \), \( n \) represents the length of each training set series. Meanwhile, the same lag length is set for each column of information matrix. To eliminate the negative influence of redundant components in each column of time series, LASSO, a feature extraction method, is used to screen out the efficient component from each row of the matrix, thereby determining the reasonable input structure for each predictor simultaneously, where a combination of grid search method, LASSO with SVR is developed to determine the optimal adjustment coefficient. In detail, for each row vector of the matrix, the extracted components are remained, otherwise it is taken as 0. Then, all the row vectors processed by the above way are arranged into a time series from the first row to the last row.

4) Individual prediction. Based on the previous results, three individual predictors (BiGRU, SVR, and ARIMA) are selected to forecast each time series. In detail, BiGRU is adopted to forecast the time series fusing the influence factor \( X \) and HC components \( Y \), while the SVR and ARIMA are applied to fit the future trend of LCHC and LCLC components, respectively.

5) Ensemble results. The final result is obtained by summing the prediction results (PR) of individual prediction for each time series in Step 4.

III. EMPIRICAL STUDY

In this section, the inbound overnight tourist volume from three major origins to Shanghai is selected to verify the effectiveness of our approach. First, the time series of three
In this paper, all sample data are divided into training set and testing set. Specifically, we define the USA-SH and Korea-SH data as follows: training data set covers the period from January 2004 to October 2016 (154 observations), and testing data set consists of 38 observations, covering the period from November 2016 to December 2019; to Japan-SH data: training data set (January 2005 - October 2016, 142 observations) and testing data set (November 2016 - December 2019, 38 observations). In the h-step-ahead forecasting without influence factors, the prediction horizon (h) is set to 1, 3, and 6, its mapping function \( f(\cdot) \) can be described as follows:

\[
y_{t+h-1} = f(y_{t-1}, y_{t-2}, \ldots, y_{t-d})
\]

where \( y_t \) is the time series of tourism flow, and \( d \) represents the lag order with \( d = 10 \). When the influence factors are considered, the function relationship \( f(\cdot) \) of the model can be written as follows:

\[
y_{t+h-1} = f(\text{LASSO}(y_{t-1}, y_{t-2}, \ldots, y_{t-d}), p_{t-1}^{d}, p_{t-2}^{d}, \ldots, p_{t-d}^{d}))
\]

where \( p_t^d \) is the time series of ith influence factors, and \( \text{LASSO}(\cdot) \) is the operator of feature extraction using the LASSO method.

The developed approach, six single models, namely ARIMA, SARIMA, MLP, SVR, GRU and BiGRU models, are applied as benchmarks. Also, several hybrid ensemble models are employed for comparison aims. (1) Based on the different decomposition methods (ICEEMDAN, CEEMDAN, EEMD, and EMD) and individual predictors (ARIMA, MLP, SVR, GRU, and BiGRU), 20 combinations of technology modules are constructed as the models, such as ICEEMDAN-ARIMA, CEEMDAN-SVR, and EMD-BiGRU. (2) Some of the hybrid models include ICEEMDAN-BSA, CEEMDAN-BSA, EEMD-BSA, and EMDLXBSA, where BSA are the BiGRU, SVR and ARIMA predictor, respectively, involving various decomposition methods and decomposed components classification method without LASSO method and influence factors. (3) Other hybrid models consist of CEEMDAN-LXBSA, EEMD-LXBSA, and EMDLXBSA, where L and X represent the LASSO method and influence factors, respectively. (4) The proposed approach is denoted as ICEEMDAN-LXBSA.

In this study, five evaluation metrics, namely, normalized root-mean-square error (NRMSE), mean absolute percentage error (MAPE), Theil coefficient (TC), direction accuracy (DA) and coefficient of determinate (R²), are adopted to...


TABLE 1. The statistical descriptions of sample data.

| Variable     | Observations | Mean     | Median    | Maximum   | Minimum   | Std. Dev. | Skewness | Kurtosis | Jarque-Bera |
|--------------|--------------|----------|-----------|-----------|-----------|-----------|----------|----------|-------------|
| USA-SH       | 192          | 51194.16 | 50212.00  | 87998.00  | 19747.00  | 14883.77  | 0.10     | 2.30     | 4.26        |
| CPI(USA)     | 192          | 224.63   | 226.78    | 257.35    | 185.20    | 19.36     | -0.23    | 2.06     | 8.72        |
| PCE(USA)     | 192          | 11104.03 | 10817.15  | 14796.30  | 7987.40   | 1832.80   | 0.33     | 2.09     | 10.11       |
| USD/CNY      | 192          | 6.93     | 6.82      | 8.28      | 6.05      | 0.69      | 0.81     | 2.37     | 23.95       |
| Korea-SH     | 192          | 46985.33 | 47499.50  | 77369.00  | 18896.00  | 11982.78  | -0.08    | 2.65     | 1.20        |
| H/H(Korea)   | 192          | 1329748.00 | 1309394.00 | 1745020.00 | 904992.60 | 224273.60 | 0.36     | -1.02    | 9.45        |
| CPI(Korea)   | 192          | 92.80    | 95.99     | 105.65    | 75.03     | 9.42      | -0.40    | 1.76     | 17.47       |
| CNY/KRW      | 192          | 161.15   | 168.74    | 212.46    | 117.41    | 22.87     | -0.63    | 2.35     | 16.09       |
| Japan-SH     | 180          | 84149.99 | 82459.50  | 136180.00 | 41109.00  | 18017.82  | 0.28     | 2.78     | 2.68        |
| HCE(Japan)   | 180          | 290415.70 | 286067.00 | 351667.00 | 260644.00 | 186374.48 | 1.01     | 3.81     | 35.45       |
| CPI(Japan)   | 180          | 98.35    | 97.50     | 102.30    | 95.70     | 1.93      | 0.45     | 1.77     | 17.43       |
| 100JPY/CNY   | 180          | 15.19    | 15.25     | 19.94     | 12.02     | 1.96      | 0.28     | 2.46     | 4.47        |

FIGURE 4. Monthly tourist flow in Shanghai and their influence factors.

assess the performance of our approach. In more detail, the first three (i.e. NRMSE, MAPE, and TC) are used to take into account the differences between the target and the predicted output. Then, the DA is employed to reflect the consistency of change trend between the predicted value and the real value, and $R^2$ indicates the fit degree between the original data and the predicted value, which depends largely on the trend analysis and is denoted as a value between 0 and 1. The closer the value is to 1, the better it fits.

The details of the above evaluation criteria are summarized in Table 2, where $M$ is the total number of the exchange rate out-of-samples series; $\hat{y}_m$ and $y_m$ represent the forecasted value and the observed value, respectively; and $\overline{y}$ is the mean value of the observed values.

Obviously, the test approach is to select the predicted value with fewer errors based on one of the error measurements described in the aforementioned evaluation criteria. However, we need to take a step to determine whether this difference is significant (for predictive purposes). For this reason, the Diebold-Mariano (DM) and the Kolmogorov-Smirnov predictive accuracy (KSPA) tests are introduced from a statistical perspective.

Here, DM test detects the differences of prediction direction selection between target model ($\text{mod}0$) and benchmark model ($\text{mod}1$). The null hypothesis is that the two models have equal forecasting accuracy. In this study, the loss function ($l$) is defined as:

$$l_t = (y_{\text{mod}0,t} - y_t)^2 - (y_{\text{mod}1,t} - y_t)^2$$

(22)

where $y_{\text{mod}0,t}$ and $y_{\text{mod}1,t}$ denote the estimates for $y_t$ calculated by the $\text{mod}0$ and $\text{mod}1$, respectively, in period $t$. Thus, the DM statistic ($H$) can be described as:

$$H = \frac{\bar{l}}{\sqrt{\bar{v}_t}/M}$$

(23)

where $\bar{l} = 1/M \sum_{t=1}^{M} l_t$ and $\bar{v}_t = y_0 + 2 \sum_{q=1}^{\infty} \gamma_q$, ($\gamma_q = \text{cov}(l_t, l_{t-q})$). Other, please refer to Diebold et al. [37] for the details of DM test.

Another point is that the KSPA test is designed to ascertain whether there is a statistical difference among the distributions of prediction errors. It estimates whether a target with a lower error reflects a smaller random error relative to the model. The one-side KSPA test is applied in this study, and its null hypothesis is that the empirical cumulative distribution function of tested model error is smaller than that of competition model. When calculating the statistical value of KSPA test, the absolute value of errors is used as the loss function. For more, please refer to Hassani et al. [38] regarding the strict mathematical derivation of the KSPA statistical test.

C. DECOMPOSITION AND CLASSIFICATION RESULTS

The ICEEMDAN decomposition results of tourism flow from three major origins to Shanghai and their RE and MIC results
TABLE 2. The performance metric rules.

| Metric | Definition | Equation |
|--------|------------|----------|
| NRMSE  | The normalized root-mean-square error | \( NRMSE = \frac{100}{M} \sqrt{\frac{1}{M} \sum_{m=1}^{M} (y_m - \hat{y}_m)^2} \) |
| MAPE   | The mean absolute percentage error | \( MAPE = \frac{1}{M} \sum_{m=1}^{M} \frac{|y_m - \hat{y}_m|}{y_m} \times 100\% \) |
| TC     | The Theil coefficient | \( TC = \sqrt{\frac{\sum_{m=1}^{M} (y_m - \hat{y}_m)^2}{\sum_{m=1}^{M} (y_m)^2}} \) |
| DA     | The direction accuracy | \( DA = \frac{1}{M} \sum_{m=1}^{M} d(m) \) |
| R²     | The coefficient of determinate | \( R^2 = \frac{\sum_{m=1}^{M} (y_m - \hat{y}_m)^2}{\sum_{m=1}^{M} (y_m - \bar{y})^2} \) |

Note: If \( y_m \leq \hat{y}_m \), return \( d(m) = 1 \); otherwise, return 0.

TABLE 3. The RE and MIC results of tourism flow time series.

| Source | Component | Index | Original data | IMF₁ | IMF₂ | IMF₃ | IMF₄ | IMF₅ | IMF₆ | IMF₇ | IMF₈ |
|--------|-----------|------|---------------|------|------|------|------|------|------|------|------|
| USA    | ICEEMD    | RE   | 0.1592        | 0.1655 | 0.0762 | 0.1250 | 0.1357 | 0.0552 | 0.0251 | 0.0006 |
|        |           | MIC  | -              | 0.1136 | 0.2387 | 0.1121 | 0.0324 | 0.1106 | 0.1379 | 0.3164 |
|        | CEEMD     | RE   | 0.1592        | 0.1673 | 0.1356 | 0.1777 | 0.1205 | 0.1317 | 0.0551 | 0.0204 | 0.0000 |
|        |           | MIC  | -              | 0.1078 | 0.2348 | 0.2215 | 0.1744 | 0.0364 | 0.1106 | 0.2870 | 0.3164 |
|        | EEEMD     | RE   | 0.1592        | 0.1677 | 0.1075 | 0.1321 | 0.1367 | 0.0452 | 0.0007 | 0.0000 |
|        |           | MIC  | -              | 0.1027 | 0.2791 | 0.1024 | 0.0558 | 0.1046 | 0.3164 |
|        | EMD       | RE   | 0.1592        | 0.1719 | 0.1010 | 0.1271 | 0.1179 | 0.0404 | 0.0003 |
|        |           | MIC  | -              | 0.0910 | 0.1959 | 0.0899 | 0.0822 | 0.0987 | 0.3164 |
| Korea  | ICEEMD    | RE   | 0.1179        | 0.1672 | 0.1396 | 0.1522 | 0.1050 | 0.0692 | 0.0026 |
|        |           | MIC  | -              | 0.0852 | 0.1016 | 0.0713 | 0.1689 | 0.1120 | 0.2658 |
|        | CEEMD     | RE   | 0.1179        | 0.1631 | 0.1687 | 0.1367 | 0.1552 | 0.1092 | 0.0716 | 0.0025 |
|        |           | MIC  | -              | 0.0803 | 0.0888 | 0.1058 | 0.0822 | 0.1757 | 0.1253 | 0.2658 |
|        | EEEMD     | RE   | 0.1179        | 0.1677 | 0.1404 | 0.1366 | 0.0732 | 0.0539 | 0.0178 |
|        |           | MIC  | -              | 0.0755 | 0.0722 | 0.0430 | 0.2465 | 0.0500 | 0.2402 | 0.2658 |
|        | EMD       | RE   | 0.1179        | 0.1684 | 0.1468 | 0.1421 | 0.0729 | 0.0433 | 0.0186 |
|        |           | MIC  | -              | 0.0732 | 0.0703 | 0.0375 | 0.1557 | 0.0806 | 0.2402 | 0.2658 |
| Japan  | ICEEMD    | RE   | 0.1657        | 0.1728 | 0.1310 | 0.1389 | 0.1257 | 0.0446 | 0.0250 |
|        |           | MIC  | -              | 0.1522 | 0.0855 | 0.0649 | 0.1259 | 0.1774 | 0.2114 | 0.3315 |
|        | CEEMD     | RE   | 0.1657        | 0.1588 | 0.1904 | 0.1408 | 0.1400 | 0.1153 | 0.0465 | 0.0276 | 0.0009 |
|        |           | MIC  | -              | 0.1322 | 0.1067 | 0.0825 | 0.0987 | 0.1259 | 0.1678 | 0.1961 | 0.3315 |
|        | EEEMD     | RE   | 0.1657        | 0.1721 | 0.1516 | 0.1442 | 0.1269 | 0.0484 | 0.0202 | 0.0003 |
|        |           | MIC  | -              | 0.1068 | 0.0825 | 0.0640 | 0.1528 | 0.1133 | 0.2710 | 0.3315 |
|        | EMD       | RE   | 0.1657        | 0.1757 | 0.1304 | 0.1301 | 0.1113 | 0.0375 | 0.0102 |
|        |           | MIC  | -              | 0.0910 | 0.0861 | 0.0741 | 0.0825 | 0.1763 | 0.3315 |

TABLE 4. The intrinsic mode functions classification results.

| Source | Components | ICEEMDAN | CEEMDAN | EEEMD | EMD |
|--------|------------|----------|---------|-------|-----|
| USA    | ICEEMD     | IMF₁     | IMF₁    | IMF₁  | IMF₁ |
|        | CEEMD      | IMF₅     | IMF₅    | IMF₅  | IMF₅ |
|        | LCLC       | IMF₂, IMF₃, IMF₄, IMF₆, IMF₇ | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃ |
|        | LCHC       | IMF₂, IMF₃, IMF₄, IMF₆, IMF₇, IMF₉ | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃ |
| Korea  | ICEEMD     | IMF₁, IMF₂, IMF₃ | IMF₁, IMF₂, IMF₃, IMF₄ | IMF₁, IMF₂, IMF₃ | IMF₁, IMF₂, IMF₃ |
|        | CEEMD      | IMF₅     | IMF₅    | IMF₅  | IMF₅ |
|        | LCLC       | IMF₂, IMF₃, IMF₄, IMF₆, IMF₇, IMF₉ | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃ |
|        | LCHC       | IMF₂, IMF₃, IMF₄, IMF₆, IMF₇, IMF₉, IMF₁₀ | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃ |
| Japan  | ICEEMD     | IMF₁     | IMF₁    | IMF₁  | IMF₁ |
|        | CEEMD      | IMF₅     | IMF₅    | IMF₅  | IMF₅ |
|        | LCLC       | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃ |
|        | LCHC       | IMF₂, IMF₃, IMF₄, IMF₆, IMF₇, IMF₉, IMF₁₀ | IMF₂, IMF₃, IMF₄ | IMF₂, IMF₃ |

Notes: HC represents high-complexity components, LCLC and LCHC represent low-complexity low-correlation and low-complexity high-correlation components, respectively.

The decomposed components of different source markets represent the changes caused by different influence factors.
TABLE 5. The adjustment coefficient of high-complexity components by LASSO.

| Source | Component | ICEEMDAN | CEEMDAN | EMD |
|--------|-----------|----------|---------|-----|
| USA    | IMF₁      | 0.0020   | 0.0037  | 0.0093 | 0.0024 | 0.0010 | 0.0085 | 0.0019 | 0.0004 | 0.0002 |
|        | IMF₂      | 0.0002   | 0.0021  | 0.0099 | 0.0022 | 0.0083 | 0.0009 | 0.0050 | 0.0078 | 0.0104 |
|        | IMF₃      | 0.0097   | 0.0016  | 0.0002 | 0.0003 | 0.0021 | 0.0016 | 0.0013 | 0.0014 | 0.0037 |
|        | IMF₄      | 0.0099   | 0.0056  | 0.0082 | 0.0090 | 0.0061 | 0.0011 | 0.0048 | 0.0075 | 0.0011 |
| Korea  | IMF₁      | 0.0064   | 0.0093  | 0.0099 | 0.0058 | 0.0034 | 0.0082 | 0.0061 | 0.0079 | 0.0060 |
|        | IMF₂      | -        | -       | -      | -      | -      | -      | -      | -      | -      |
| Japan  | IMF₁      | 0.0064   | 0.0093  | 0.0099 | 0.0058 | 0.0034 | 0.0082 | 0.0061 | 0.0079 | 0.0060 |

TABLE 6. The optimal parameters of SVR of each low-complexity and low-correlation component.

| Source | Component | Parameter | ICEEMDAN | CEEMDAN | EMD |
|--------|-----------|-----------|----------|---------|-----|
| USA    | IMF₃      | C         | -        | -       | -   |
|        |           | ε         | -        | -       | -   |
|        |           | C         | 7.7573   | 7.7573  | 9.7961 |
|        |           | ε         | 0.0003   | 0.0002  | 0.0010 |
|        | IMF₅      | C         | -        | -       | -   |
|        |           | ε         | -        | -       | -   |
|        | Korea     | IMF₅      | C         | -        | -   |
|        |           | ε         | -        | -       | -   |
|        | IMF₃      | C         | 8.3690   | 0.2139  | 0.0100 |
|        |           | ε         | 0.0002   | 0.0008  | 0.0005 |
| Japan  | IMF₄      | C         | -        | -       | -   |
|        |           | ε         | -        | -       | -   |

TABLE 7. The optimal forms of ARIMA(p, d, q) of low-complexity and high-correlation components.

| Source | Method   | ICEEMDAN | CEEMDAN | EEMD | EMD |
|--------|----------|----------|---------|------|-----|
| USA    | IMF₁     | (2, 0, 2) | (2, 0, 1) | -    | (1, 0, 2) |
|        | IMF₂     | (2, 0, 5) | (4, 0, 4) | (2, 0, 1) | -    | (4, 0, 2) |
|        | IMF₃     | (3, 0, 4) | (5, 0, 5) | (4, 1, 3) | (3, 2, 3) | |
|        | IMF₄     | (2, 0, 1) | -        | (4, 0, 3) | (1, 0, 3) | (4, 2, 0) |
| Korea  | IMF₁     | -        | -        | (4, 0, 2) | (1, 0, 2) | (4, 2, 0) |
|        | IMF₂     | -        | -        | (5, 0, 0) | (1, 0, 2) | (2, 2, 3) |
|        | IMF₃     | -        | -        | (3, 1, 2) | (2, 2, 3) | (1, 2, 0) |
|        | IMF₄     | -        | -        | (5, 1, 0) | (1, 2, 1) | (1, 2, 0) |
| Japan  | IMF₁     | -        | -        | (4, 0, 2) | (4, 1, 3) | (3, 2, 0) | (4, 2, 0) |
|        | IMF₂     | -        | -        | (4, 0, 2) | (3, 1, 3) | (2, 2, 0) | (2, 2, 4) |
|        | IMF₃     | -        | -        | (4, 0, 2) | (1, 0, 3) | (2, 3, 2) | (3, 2, 2) |
|        | IMF₄     | -        | -        | (3, 1, 2) | (3, 2, 4) | -       | -       |

in tourist flow. The fluctuation of the first few components means that the short-term tourism volume is affected by certain influence factors, such as weather conditions, holiday length, consumer disposable income and exchange rate movement. Therefore, for short-term fluctuation components with high-complexity, we select CPI, exchange rate and personal or household expenditure or income as the auxiliary variables to assist the forecasting of these components. The changes in the middle few components represent that the emergencies can cause a certain impact on tourism demand within a period of time. For example, due to the global financial crisis in 2008, the international tourist flow from the USA, Korea and Japan to Shanghai decreased simultaneously, and it did not rebound until 2010. Specifically, all IMF₄ of three time series show a significantly downward trend during the financial crisis. In Sep. 2016, the terminal high altitude area defense (THAAD) incident caused tensions in international relations between China and South Korea, and also resulted in a decline of tourist volumes. The IMF₄ of Korea-SH time series displays a downward trend from Sep. 2016 to Oct. 2017. Since the Diaoyu island incident in 2012, the tourist volume from Japan to Shanghai has dropped rapidly, which could also be seen from the IMF₄ of Japan-SH time series. The long-term fluctuation corresponds to last few decomposition components, they are generally affected by local infrastructure, tourism policies, international conditions, and so on. From the last one component of three time series, it can be seen that the tourist flow from USA and Korea to Shanghai has
shown a steady growth year by year, while the tourist arrivals from Japan have shown a downward trend since 2008 and have gradually stabilized after 2018.

To obtain better prediction accuracy, we select an appropriate predictor for each IMF based on the complexity and correlation. The components that RE values of IMF are greater than that of the original time series are regarded as the high-complexity ones, and the rest shows low-complexity. According to the MIC results of low-complexity components, those with a value lower than 0.1 are classified as low-correlation, and the rest are high-correlation components. The classification results in Table 4 indicate that the first few components represent high-complexity characteristics, while the last few possess high-correlation with raw time series in most cases.

For high-complexity components, we first combine them with preselected influence factors time series to form a matrix, and then use LASSO to extract effective features from all the columns of matrix and fuse them into a time series, meanwhile the input structure is determined. As shown in Table 5, the adjustment coefficients of LASSO are chosen according to the grid search method. Then, BiGRU is applied to forecast each high-complexity component based on the same network hyperparameters, which are set as follows: a number of recurrent layers = 3, hidden units = 50, batch sizes = 30 and epochs = 200. Since the training parameters of BiGRU network have certain randomness, the final prediction results are obtained by averaging the values of 10-time runs. For low-complexity and low-correlation components, the optimal parameters of forecasting model SVR are determined by the grid search and cross validation methods (see Table 6). For low-complexity and high-correlation components, the optimal forms of ARIMA are listed in Table 7 based on AIC and BIC.

D. EXPERIMENTAL RESULTS ANALYSIS
To verify the effectiveness and robustness of the designed approach, a systematic analysis including various benchmark models and metric methods is implemented from the point of error and statistics test.

1) ERROR CRITERION ANALYSIS
The forecasting error-based evaluation can directly reflect the differences between predicted values and observed values. Table 8, Table 9 and Table 10 show the preliminary comparison results of one-step-ahead, three-step-ahead and six-step-ahead forecasting for all considered models, respectively, where the bold represents the best results. Meanwhile, Figure 6, Figure 7 and Figure 8 display visually the forecasting results of all models on the three testing datasets by scatter plots and line plots of forecasting error and forecasting value, respectively. Some meaningful conclusions and analyses can be summarized as follows.

First, from the perspective of single prediction model, deep learning has obvious advantages in complex time series. In the experiment of USA-SH, the BiGRU shows a significant advantage over the other single models (i.e. ARIMA, SARIMA, MLP, SVR and GRU). For one-step-ahead forecasting of USA-SH, MAPE, NRMSE, TC, DA, and $R^2$ of the BiGRU are 6.5939%, 9.3602%, 0.0393, 0.9730 and 0.8656, respectively, which are the lowest values in the single models. And Figure 6 shows that the line plot of forecasting error of BiGRU is closer to zero than other single models. In the experiment of Korea-SH and Japan-SH, the deep learning model (i.e. BiGRU) shows the same advantages in one-step-ahead and multi-step-ahead forecasting. Specifically, compared with the three-step-ahead forecasting of SARIMA, the MAPE of BiGRU is improved by 5.60% (Korea-SH) and 27.28% (Japan-SH), respectively, and compared with that of MLP, the NRMSE of BiGRU is reduced by 1.16% (Korea-SH) and 4.14% (Japan-SH), respectively. The scatter plot of the forecasting values and true value is also closer to a straight line in three-step-ahead forecasting in Figure 7 and Figure 8.

Therefore, it is reasonable to conclude that BiGRU is more suitable for training nonlinear and highly complex datasets.
TABLE 8. The evaluation results of forecasting performance of each model for the tourist flow from USA to Shanghai.

| Model       | MAPE | NRMSE | TC | DA   | R²    |
|-------------|------|-------|----|------|-------|
| EMD-ARIMA   | 13.0290 | 0.748 | 0.378 | 0.531 | 16.5938 |
| EEMD-ARIMA  | 12.4708 | 0.0929 | 0.669 | 0.336 | 12.7035 |
| SVR         | 8.2626 | 0.1543 | 0.459 | 0.791 | 9.1359 |
| GRU         | 8.9634 | 0.2910 | 0.459 | 0.791 | 8.2164 |
| BIGRU       | 6.5397 | 0.3602 | 0.393 | 0.730 | 6.7269 |

Second, the decomposition-ensemble approaches combining different models are superior to the corresponding single model. For example, in USA-SH Korea-Sh and Japan-Sh, the decomposition-ensemble models (i.e. EMD-MLP, EEMD-MLP, CEEMDAN-MLP and ICEEMDAN-MLP) can obtain more accurate forecasting than single MLP predictor, the MAPE values of them increase by about 10% on average. Similar experimental results can be obtained by comparing other single models. Notably, our experiments show that some decomposition-ensemble approaches with ARIMA (such as EMD-ARIMA, EEMD-ARIMA, CEEMDAN-ARIMA and ICEEMDAN-ARIMA) have poorer prediction results than some single models (such as SARIMA, SVR and GRU) in USA-Sh and Japan-Sh. R² values of these decomposition-ensemble approaches are all below 65%. The interesting facts indicate that the unreasonable choice of model will lead to a decrease of prediction performance. When the decomposition-ensemble approach combining BiGRU is used to train and forecast each IMF, the obtained results will be significantly better than those of other decomposition-ensemble approaches with simple forecasting model (such as ICEEMDAN-ARIMA, ICEEMDAN-MLP, ICEEMDAN-SVR and ICEEMDAN-GRU). In the experiments USA-Sh Korea-Sh and Japan-Sh, the MAPE values of ICEEMDAN-BiGRU are all below 2.90% in the one-step-ahead forecasting, and its R² are all above 80% in the three-step-ahead forecasting. From Figure 6, Figure 7 and Figure 8, the line plots also display that the prediction errors of decomposition-ensemble approaches are significantly reduced compared with the single model. Hence, the decomposition-ensemble approach is helpful for the extraction of complex data features, which can significantly reduce the difficulty of prediction through modal decomposition, so as to improve the overall prediction performance.
and CEEMDAN-BSA even reached 100, which can also be found in Figure 8. In general, with the increase of the prediction step, the accuracy of all models decreases to a certain extent. But on the whole, selecting adaptive predictor according to different IMF features is helpful to slow down the decline of accuracy. Based on this reason, the MAPE values of ICEEMDAN-BSA with respect to three-step-ahead forecasting can remain roughly between 3.5% and 5.5%, and the $R^2$ values of that are above 0.85. In addition, for six-step-ahead forecasting, the DA values of ICEEMDAN-BSA are also above 75 and the NRMSE values are all below 12.2. Hence, matching adaptive prediction models by using the IMF classification method can enhance the forecasting performance of hybrid approach significantly.

Forth, in the empirical analysis, we find that it is difficult to improve further the prediction accuracy for low-complexity components, but there is still much room for improving high-complexity components. Here, the influence factors (such as exchange rate, consumer price index (CPI) and personal or household consumption expenditure or income) are added to high-frequency components to compensate for the negative effects caused by short-term fluctuations. In fact, the forecasting accuracy (MAPE) and fitting effect ($R^2$) of proposed approach (ICEEMDAN-LXBSA) are the best in either one-step-ahead or multi-step-ahead forecasting. The MAPE of ICEEMDAN-LXBSA is at a low value (1.8481%, 1.7987% and 1.7545% from USA, Korea and Japan to Shanghai in one-step-ahead forecasting, respectively), and $R^2$ is at the higher value (0.9854, 0.9737 and 0.9831 from USA, Korea and Japan to Shanghai in one-step-ahead forecasting, respectively). In terms of direction accuracy, ICEEMDAN-LXBSA also achieves a great advantage compared with most of the benchmark models. Especially, the DA values of ICEEMDAN-LXBSA with respect to one, three and six-step ahead forecasting in the experiment from USA to Shanghai are 0.9730, 0.9730 and 1.0000, respectively. Similarly, the DA of ICEEMDAN-LXBSA the tourist flow from Korea to Shanghai also reaches the optimal value. Also, the line plots and scatter plots (i.e. Figure 6, Figure 7 and Figure 8) indicate that the forecasting values are closest to the true values in one and multi-step-ahead forecasting, they are closer to a true line and straight line, respectively. Therefore, it is an effective method to improve the prediction effect by adding influence factors to high-frequency components (IMF).
2) MULTI-FACTOR FUSION ANALYSIS

In order to display better the improved effect of multi-factors fusion on the prediction accuracy of high-frequency components, Table 11 shows a detailed comparison of the prediction results of the high frequency components with and without influence factors. For IMF1 of USA-SH and Japan-SH, the average NRMSE of LX-BiGRU improves by about 25% and 32% compared to BiGRU. In one-step-ahead forecasting for IMF1, IMF2 and IMF3 of Korea-SH, the NRMSE of LX-BiGRU that uses ICEEMDAN method increases by 50.57%, 22.49% and 54.50%, respectively. For Korea-SH and Japan-SH, only IMF4 and IMF2 obtained by CEEMDAN method are the high-complexity components. Its fitting effect is the best when the influence factors are added into IMF4 and IMF2, R2 of one-step-ahead forecasting reaches 0.95, and the multi-step-ahead forecasting also achieves a high level. Although a few high-complexity components with influence factors have the opposite effect (the table is marked with an asterisk), there is an overall significant positive effect on one-step-ahead forecasting (all R2 are around 90%) and a negative effect on a few six-step-ahead forecasting (such as IMF1 of CEEMDAN and EMD, IMF of CEEMDAN and EEMD).

In addition, it can also be seen that using ICEEMDAN gains greater forecasting accuracy compared with other decomposition methods. For one thing, ICEEMDAN can extract the short-term fluctuation components (i.e. noise), for another, the fitting effect of high-complexity components of ICEEMDAN is significantly enhanced by incorporating influence factors. In conclusion, no matter which decomposition method is adopted, introducing appropriate influence factors is helpful to improve the forecasting accuracy of high-complexity components.

3) STATISTIC TEST ANALYSIS

In order to prove the feasibility of the developed hybrid approach, DM and KSPA tests are implemented from a statistical perspective. The test results of DM and KSPA are listed in Table 12, Table 13 and Table 14 with one-step-ahead and multi-step-ahead forecasting, respectively. The following is a detailed analysis.

In DM statistical test, except for CEEMDAN-BSA of Korea-SH time series, it can be confirmed that the results of ICEEMDAN-LXBSA reject the null hypothesis at the p = 0.1 level compared to others, in other words, the proposed hybrid approach outperforms other benchmark models with the 90% confidence level. Especially, in most
TABLE 11. The evaluation results of forecasting performance of high-complexity components for three time series.

| Source | Component | Method        | 1-step | 3-step | 6-step |
|--------|-----------|---------------|--------|--------|--------|
|        |           |               | RMSE   | R²     | RMSE   | R²     | RMSE   | R²     |
| USA-SH | IMF₁      | ICEMDAN-LXBS | 6.0794  | 0.9412 | 5.4167  | 0.9661 | 5.0093  | 0.9796 | 4.7983  | 0.9843 | 4.5368  | 0.9902 |
|        |           | ICEMDAN      | 7.3992  | 0.9139 | 5.7093  | 0.9437 | 5.2017  | 0.9503 | 4.8731  | 0.9604 | 4.6826  | 0.9662 |
|        |           | CEEMDAN      | 8.6317  | 0.8848 | 5.9462  | 0.9345 | 5.4473  | 0.9476 | 5.1087  | 0.9557 | 4.8985  | 0.9627 |
|        |           | EEMD         | 9.6515  | 0.8304 | 6.3467  | 0.9106 | 5.8664  | 0.9419 | 5.4383  | 0.9525 | 5.1307  | 0.9594 |
|        |           | EM           | 5.5753  | 0.9496 | 3.9546  | 0.9394 | 3.5924  | 0.9464 | 3.3083  | 0.9542 | 3.0678  | 0.9613 |
|        |           | EMD          | 12.5702 | 0.6401 | 6.3069  | 0.9094 | 5.8748  | 0.9399 | 5.4396  | 0.9462 | 5.1048  | 0.9527 |
|        |           | TED           | 10.9633  | 0.7354 | 6.6672  | 0.9021 | 6.3349  | 0.9260 | 5.9702  | 0.9350 | 5.5526  | 0.9462 |
|        |           | EMD          | 9.1672  | 0.8082 | 8.7313  | 0.9152 | 8.1742  | 0.9252 | 7.7852  | 0.9316 | 7.4462  | 0.9376 |

**Notes:**
- L. X. B. S. A represent models employing LASSO method, influence factor, BiGRU, SVM and ARIMA, respectively.
- Similar to DM test, in one-step-ahead and multi-step-ahead forecasting, KSPA test results also indicate that the error distribution of ICEMDAN-LXBSA has a positive significant difference than that of other models at the p = 0.1 confidence.

TABLE 12. The DM and KSPA test results of testing data set for USA-SH time series.

| Model | 1-step | 3-step | 6-step |
|-------|--------|--------|--------|
|       | DM     | KSPA   | DM     | KSPA   | DM     | KSPA   |
| CEEMDAN-LXBSA | -1.5702 | -0.0328 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
|       | -1.0500 | -0.0328 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| EEMD-LXBSA | -3.0930 | -0.0622 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| ICEMDAN-LXBSA | -3.6333 | -0.0512 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| CEEMDAN-BiGRU | -1.6607 | -0.0250 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
|       | -1.8427 | -0.0250 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| EEMD-BiGRU | -2.3586 | -0.0318 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| ICEMDAN-BiGRU | -2.7124 | -0.0318 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| ICEMDAN-SVR | -3.5874 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
|       | -3.6712 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| EEMD-SVR | -3.9795 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| ICEMDAN-MLP | -3.8791 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
|       | -2.8788 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| ICEMDAN-ARIMA | -4.1327 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
|       | -4.5388 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| EEMD-ARIMA | -4.9273 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| ICEMDAN-ARIMA | -2.7555 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
|       | -4.2834 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |
| EMD-ARIMA | -4.6588 | -0.0526 | -0.0000 | 0.0000 | -0.0000 | 0.0000 |

**Notes:**
- L, X, B, S. A represent models employing LASSO method, influence factor, BiGRU, SVM and ARIMA, respectively.
level. In other words, the forecasting error of the proposed approach is closer to the normal distribution, showing the characteristics of white noise, while there may still be unextracted effective information of the prediction error in other models.

Generally, whether from DM or KSPA results, ICEEMDAN-LXBSA shows the characteristics of high precision and sufficient information extraction (i.e. the prediction error is closer to white noise). And from the statistical point of view, the proposed approach has amended, to a certain extent, some problems of the commonly used decomposition ensemble framework.

E. DISCUSSION AND COMPARISON WITH BENCHMARK MODELS

In order to understand better the comprehensive advantages of the hybrid ensemble learning approach constructed in this study, this section discusses in detail the deep reasons for the impact of each module of our approach and benchmark model on the forecasting performance.

First, for the comparison of a single model, SARIMA model achieves higher accuracy than ARIMA model in three time series, and the improvement effect is the most obvious in the USA time series. The reason for it is that the tourist flow from USA is more cyclical than the other two sources. Meanwhile, the performance of SARIMA model is better than that of the AI-based model (SVR, GRU, MLP) in some cases. This opinion is similar to that of Claveria et al. [39], whose research indicated that an AI-based model is more likely to cause information loss than linear models in the filtering process, even though it is excellently equipped to handle nonlinear pattern. In addition, SVR and GRU models outperform MLP model in most cases in the stability and forecasting performance. BiGRU model has higher prediction accuracy than GRU model in one-step-ahead and multi-step-ahead forecasting because of the application of bidirectional transmission mechanism of hidden layers. This mechanism can not only enable the model to have stronger learning ability, but also increase the time cost. In general, BiGRU model performs better than other models in all cases. The reason for this result is that deep learning technology can not only capture the hidden regular components in complex time series processing but also better adapt to its nonlinear components.

Second, using decomposition technique reduces the data complexity and prediction difficulty. According to the results of three time series, the performance of single model is worse than that of the decomposition ensemble model that uses
corresponding single model as an individual predictor. However, the deep learning method, BiGRU, gains stronger processing capacity than some decomposition ensemble models (such as EMD-ARIMA, EEMD-ARIMA, and EMD-MLP) in different prediction horizons of three time series. The previous studies manifest that the defect of EMD method leads to the decomposed information confusion and development of EEMD, CEEMDAN and ICEEMDAN methods suppresses this problem to a varying degree by adding white noise to the original data to change the distribution of extremum points. In the process of multiple integration averages, the added white noise cancels each other out, so as to reveal the real data component gradually. In all decomposition ensemble models, the ICEEMDAN-based models outperform the models based on other decomposition methods (EMD, EEMD, CEEMDAN) in most cases. For the model using the same predictor to forecast each decomposed component, in most cases, the hybrid approach based on different decomposition methods and BiGRU displays the best performance in prediction level and direction accuracy.

Third, in order to select adaptively a reasonable predictor to forecast each IMF, we propose an IMFs classification method that considers both complexity and correlation to classify each component. In addition, to further improve forecasting effect, the influence factors are introduced and combined with high-complexity components. However, in some three-step-ahead and six-step-ahead prediction cases, the forecasting results of BiGRU without considering influence factors are closer to the actual results. The reason for this result may be that the influence factors do not supplement the useful information for these high-complexity components, so playing an opposite role, or the feature selection method does not effectively filter the redundant variables. But, on the whole, the proposed IMFs classification method is reasonable and can logically select a suitable predictor to improve the overall prediction performance.

Finally, compared with all benchmark models, the performance of the proposed approach (ICEEMDAN-LXBSA) is the most competitive one in different prediction horizons. The main reasons are as follows: (1) ICEEMDAN method, as a data decomposition technique, can effectively extract information with different frequencies and amplitudes in non-linear and complex time series of tourist flow for tourism demand forecasting. (2) The IMFs classification method is completely divided into three categories based on the complexity and correlation of each component. (3) The deep learning model (BiGRU) is suitable for the most complex IMF with its powerful model learning capabilities, and the

### TABLE 14. The DM and KSPA test results of testing data set for Japan-SH time series.

| Model          | 1-step   | 3-step | 6-step  |
|----------------|----------|--------|---------|
|                | DM p-value | KSPA p-value | DM p-value | KSPA p-value | DM p-value | KSPA p-value |
| CEEMDAN-LXBSA  | -1.3801  | 0.0838 | 0.3158 | 0.0008 | -1.3111 | 0.0929 | 0.3684 | 0.0041 |
| EEMD-LXBSA     | -2.5860  | 0.0049 | 0.5789 | 0.0000 | -3.4340 | 0.0003 | 0.6579 | 0.0000 |
| EMD-LXBSA      | -2.8809  | 0.0020 | 0.3421 | 0.0087 | -3.1070 | 0.0000 | 0.4474 | 0.0000 |
| CEEMDAN-BSA    | -1.4587  | 0.0769 | 0.4211 | 0.0008 | -2.8944 | 0.0019 | 0.3158 | 0.0175 |
| EEMD-BSA       | -1.8434  | 0.0395 | 0.3158 | 0.0175 | -2.7986 | 0.0026 | 0.3158 | 0.0175 |
| EMD-BSA        | -2.3789  | 0.0087 | 0.3684 | 0.0041 | -3.6657 | 0.0001 | 0.5263 | 0.0000 |
| ICEEMDAN-BiGRU | -3.4960  | 0.0002 | 0.5900 | 0.0000 | -3.9446 | 0.0000 | 0.5263 | 0.0000 |
| EEMD-BiGRU     | -3.1202  | 0.0009 | 0.3158 | 0.0175 | -3.7123 | 0.0001 | 0.4737 | 0.0001 |
| EMD-BiGRU      | -2.9140  | 0.0018 | 0.4211 | 0.0008 | -3.7476 | 0.0001 | 0.4737 | 0.0001 |
| ICEEMDAN-SVR   | -3.3343  | 0.0004 | 0.5000 | 0.0000 | -5.2129 | 0.0000 | 0.6579 | 0.0000 |
| CEEMDAN-SVR    | -4.4146  | 0.0000 | 0.5719 | 0.0000 | -3.8042 | 0.0001 | 0.6316 | 0.0000 |
| EEMD-SVR       | -3.8854  | 0.0001 | 0.6316 | 0.0000 | -3.9829 | 0.0001 | 0.6316 | 0.0000 |
| EMD-SVR        | -3.8838  | 0.0001 | 0.5719 | 0.0000 | -4.2283 | 0.0000 | 0.7368 | 0.0000 |
| ICEEMDAN-MLP   | -4.4340  | 0.0000 | 0.7895 | 0.0000 | -4.8665 | 0.0000 | 0.7895 | 0.0000 |
| CEEMDAN-MLP    | -3.0794  | 0.0010 | 0.5719 | 0.0000 | -3.6421 | 0.0001 | 0.6035 | 0.0000 |
| EEMD-MLP       | -3.0240  | 0.0012 | 0.6053 | 0.0000 | -3.4547 | 0.0003 | 0.5789 | 0.0000 |
| EMD-MLP        | -3.0632  | 0.0011 | 0.6316 | 0.0000 | -3.6326 | 0.0003 | 0.6579 | 0.0000 |
| ICEEMDAN-ARIMA | -3.8283  | 0.0001 | 0.7895 | 0.0000 | -3.4312 | 0.0000 | 0.7368 | 0.0000 |
| CEEMDAN-ARIMA  | -3.8953  | 0.0000 | 0.7895 | 0.0000 | -3.7224 | 0.0001 | 0.7105 | 0.0000 |
| EEMD-ARIMA     | -4.3166  | 0.0000 | 0.7632 | 0.0000 | -3.9174 | 0.0000 | 0.7632 | 0.0000 |
| EMD-ARIMA      | -4.2812  | 0.0000 | 0.7368 | 0.0000 | -4.4997 | 0.0000 | 0.6579 | 0.0000 |
| BIGRU          | -3.9213  | 0.0000 | 0.7105 | 0.0000 | -4.9945 | 0.0000 | 0.7632 | 0.0000 |
| GRU            | -2.9129  | 0.0018 | 0.5719 | 0.0000 | -3.0135 | 0.0013 | 0.5789 | 0.0000 |
| SVR            | -4.2686  | 0.0000 | 0.6579 | 0.0000 | -4.2878 | 0.0000 | 0.7105 | 0.0000 |
| MLP            | -3.7407  | 0.0000 | 0.8158 | 0.0000 | -3.7387 | 0.0001 | 0.7105 | 0.0000 |
| SARIMA         | -3.8688  | 0.0001 | 0.6579 | 0.0000 | -3.6357 | 0.0000 | 0.7368 | 0.0000 |
| ARIMA          | -3.0068  | 0.0013 | 0.7895 | 0.0000 | -4.6486 | 0.0000 | 0.7368 | 0.0000 |

Notes: L, X, B, S, A represent models employing LASSO method, influence factor, BiGRU, SVR and ARIMA, respectively.
FIGURE 6. The comparison of prediction results of all models for tourist flow from USA.

FIGURE 7. The comparison of prediction results of all models for tourist flow from Korea.

FIGURE 8. The comparison of prediction results of all models for tourist flow from Japan.

regular and linear components are predicted by the ARIMA model, while the remaining components are fitted by the SVR model with strong generalization ability. (4) Considering that high-complexity components contain random factors and are affected by the short-term market, some reasonable influence factors are selected as supplementary information, and on this basis, the feature selection method (LASSO) is used to eliminate redundant information, thereby improving forecasting accuracy. Therefore, the above reasons and the experimental results prove convincingly the superiority and rationality of our approach.

IV. CONCLUSION

Providing an accurate forecast of tourism demand is helpful for tourism market managers to provide effective tourism services and for investors to specify reasonable investment plans and programs. However, due to the defect of forecasting technology and the randomness and irregularity of time series caused by multiple factors, the accurate prediction of tourism demand becomes a challenging task. To ameliorate these shortcomings, a multi-factor hybrid ensemble learning approach is proposed to forecast the monthly overnight tourist flow from three major origins (USA, Korea and Japan) to Shanghai. The main work includes the following aspects: data decomposition, intrinsic mode functions classification, multi-factors fusion, individual prediction and ensemble results. The empirical results demonstrate that the proposed approach outperforms all benchmark models in terms of different prediction horizons.

The contributions of the study can be summarized as follows. First, unlike previous studies that only use historical data, this study develops a decomposition-ensemble approach with multi factor fusion strategy, thereby improving the
prediction ability and the explainability. To be specific, the IMFs with high complexity usually present the high frequency and high volatility. These IMF components are often affected by short-term economic factors and are difficult to be effectively extracted, resulting in a decline of the forecasting accuracy to a certain extent. Therefore, we fuse some influence factors (i.e., exchange rate, consumer price index (CPI) and personal or household consumption expenditure or income) into the high-complexity components in an attempt to capture the dynamic of time series well. Second, the advantages of decomposition-ensemble approach have been demonstrated in many studies including the field of tourism demand forecasting, but there are only a few ways to conduct further feature analysis on IMFs. This study proposes innovatively an IMFs classification method based on a comprehensive strategy combing complexity (range entropy) and correlation (maximal information coefficient) to investigate the internal characteristics among all components effectively. Furthermore, we select appropriate prediction models for IMF in each category according to its features rather than the same model for prediction, thus improving the adaptability of prediction model. Third, according to the characteristics analysis of IMFD and predictor, we assign the deep Learning technique (BiGRU), machine learning method (SVR) and time series model (ARIMA) to predict the corresponding IMF. BiGRU is used to forecast the high complex IMFs combing influence factors to obtain a better training effect, SVR is applied to non-linear and non-smooth IMF, and ARIMA is employed to the low complexity IMF with obvious trend characteristics. Based on this decomposition and ensemble idea, the advantages of each technical module can be fully utilized, thus improving the overall prediction performance. Thus, a novel hybrid approach (ICEEMDAN-LXBSA) integrating the ICEEMDAN, IMFs classification and multi-factors fusion is proposed to forecast monthly tourism demand. Compared with the existing models, ICEEMDAN-LXBSA has higher precision and stability in different horizons. So, it can server as an effective reference tool for decision-makers to make corresponding plans.

This study reveals some practical implications worthy of further study in the field of tourism demand forecasting. First, in the theoretical framework, the reasonable influence factors can effectively improve the forecasting accuracy of high-complexity components. According to the empirical results, we can see that the factors, which affect directly or indirectly the tourism industry, also deserve further attention in tourism demand forecasting. Second, from a technical point of view, the proposed classification method based on feature selection of each decomposition component can provide a basis for selecting a suitable predictor. The reliability of the strategy in terms of accuracy and stability is verified in this study. Third, for practitioners and managers in the tourism industry, precise forecast results can provide references for the effective allocation and utilization of resources and policy formulation. In addition, grasping the development of tourist arrivals can effectively transform the business plan of tourism-related enterprises, such as hotels, travel agencies, and catering services.

Although the performance of the presented approach is excellent, there are still some limitations. First, the theoretical framework of this paper is only verified by the tourism demand data from three major origins (USA, Japan, and Korea) to Shanghai, and the amount of data is limited. The proposed model may have certain limitations if the data is missing or huge. Further research needs to be tested in other cities and the amount of data expanded as well. Second, the selected influence factors are mainly some macroeconomic indicators, while other active data sources should be found and introduced into the proposed approach. In addition, since the data in this study are available and complete, how to integrate other influence factor data into the model needs further consideration. Third, predicting each IMF will result in a large amount of computational cost, and the reconstruction strategies and clustering methods are effective means to solve this problem. Therefore, to achieve superior accuracy and lower computational consumption, these limitations are possible directions for future research.

REFERENCES

[1] U. Gunter and I. Önder, “Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data,” Tourism Manage., vol. 46, pp. 123–135, Feb. 2015.
[2] G. Xie, Y. Qian, and S. Wang, “A decomposition-ensemble approach for tourism forecasting,” Ann. Tourism Res., vol. 81, Mar. 2020, Art. no. 102891.
[3] C. Goh and R. Law, “The methodological progress of tourism demand forecasting: A review of related literature,” J. Travel Tourism Marketing, vol. 28, no. 3, pp. 296–317, Apr. 2011.
[4] B. Bokelmann and S. Lessmann, “Spurious patterns in Google trends data—An analysis of the effects on tourism demand forecasting in Germany,” Tourism Manage., vol. 75, pp. 1–12, Dec. 2019.
[5] D. C. Wu, H. Song, and S. Shen, “New developments in tourism and hotel demand modeling and forecasting,” Int. J. Contemp. Hospitality Manage., vol. 29, no. 1, pp. 507–529, 2017.
[6] H. Li, M. Hu, and G. Li, “Forecasting tourism demand with multisource big data,” Ann. Tourism Res., vol. 83, Jul. 2020, Art. no. 102912.
[7] H. Song and G. Li, “Tourism demand modelling and forecasting—A review of recent research,” Tourism Manage., vol. 29, no. 2, pp. 203–220, Apr. 2008.
[8] E. X. Jiao and J. L. Chen, “Tourism forecasting: A review of methodological developments over the last decade,” Tourism Econ., vol. 25, no. 3, pp. 469–492, May 2019.
[9] S.-X. Lv, L. Peng, and L. Wang, “Stacked autoencoder with echo-state regression for tourism demand forecasting using search query data,” Appl. Soft Comput., vol. 73, pp. 119–133, Dec. 2018.
[10] S. Li, T. Chen, L. Wang, and C. Ming, “Effective tourist volume forecasting supported by PCA and improved BPN model using Baidu index,” Tourism Manage., vol. 68, pp. 116–126, Oct. 2018.
[11] X. Jiao, J. L. Chen, and G. Li, “Forecasting tourism demand: Developing a general nesting spatiotemporal model,” Ann. Tourism Res., vol. 90, Sep. 2021, Art. no. 103277.
[12] C. Li, P. Ge, Z. Liu, and W. Zheng, “Forecasting tourist arrivals using denoising and potential factors,” Ann. Tourism Res., vol. 83, Jul. 2020, Art. no. 102943.
[13] E. S. Silva, H. Hassani, S. Heravi, and X. Huang, “Forecasting tourism demand with denoised neural networks,” Ann. Tourism Res., vol. 74, pp. 134–154, Jan. 2019.
[14] C.-F. Chen, M.-C. Lai, and C.-C. Yeh, “Forecasting tourism demand based on empirical mode decomposition and neural network,” Knowl.-Based Syst., vol. 26, pp. 281–287, Feb. 2012.
[15] R. Kummong and S. Supratid, “Thailand tourism forecasting based on a hybrid of discrete wavelet decomposition and NARX neural network.” Ind. Manage. Data Syst., vol. 116, no. 6, pp. 1242–1258, Jul. 2016.

[16] G. Zhang, J. Wu, B. Pan, J. Li, M. Ma, M. Zhang, and J. Wang, “Improving daily occupancy forecasting accuracy for hotels based on EEMD-ARIMA model.” Tourism Econ., vol. 23, no. 7, pp. 1496–1514, Nov. 2017.

[17] X. Li and R. Law, “Forecasting tourism demand with decomposed search cycles.” J. Travel Res., vol. 59, no. 1, pp. 52–68, Feb. 2020.

[18] H. Y. Song, T. R. Qiu, and J. Park, “A review of research on tourism demand forecasting: Launching the Annals of Tourism Research curated collection on tourism demand forecasting.” Ann. Tourism Res., vol. 75, pp. 338–362, Mar. 2019.

[19] J. Chi, “The impact of third-country exchange rate risk on international air travel flows: The case of Korean outbound tourism demand.” Transp. Policy, vol. 89, pp. 66–78, Apr. 2020.

[20] C. K. W. Chow, W. H. K. Tsui, and H. Wu, “Airport subsidies and domestic inbound tourism in China.” Ann. Tourism Res., vol. 90, Sep. 2021, Art. no. 103275.

[21] J. Zhang, “The effects of tourism on income inequality: A meta-analysis of econometrics studies.” J. Hospitality Tourism Manage., vol. 48, pp. 312–321, Sep. 2021.

[22] J.-W. Bi, Y. Liu, and H. Li, “Daily tourism volume forecasting for tourist attractions.” Ann. Tourism Res., vol. 83, Jul. 2020, Art. no. 102923.

[23] A. Kulshrestha, V. Krishnaswamy, and M. Sharma, “Bayesian BILSTM approach for tourism demand forecasting.” Ann. Tourism Res., vol. 83, Jul. 2020, Art. no. 102925.

[24] Y. Zhang, G. Li, B. Muskat, R. Law, and Y. Yang, “Group pooling for deep tourism demand forecasting.” Ann. Tourism Res., vol. 82, May 2020, Art. no. 102899.

[25] X. Li, X. Ma, F. Xiao, C. Xiao, F. Wang, and S. Zhang, “Time-series production forecasting method based on the integration of bidirectional gated recurrent unit (Bi-GRU) network and sparrow search algorithm (SSA).” J. Petroleum Sci. Eng., vol. 208, Jan. 2022, Art. no. 109309.

[26] W. Jun, L. Yuyan, T. Lingyu, and G. Peng, “Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting.” Chaos, Solitons Fractals, vol. 108, pp. 136–147, Mar. 2018.

[27] K.-Y. Chen and C.-H. Wang, “Support vector regression with genetic algorithms in forecasting tourism demand.” Tourism Manage., vol. 28, no. 1, pp. 215–226, Feb. 2007.

[28] W.-C. Hong, Y. Dong, L.-Y. Chen, and S.-Y. Wei, “SVR with hybrid chaotic genetic algorithms for tourism demand forecasting.” Appl. Soft Comput., vol. 11, no. 2, pp. 1881–1890, Mar. 2011.

[29] A. Khatibi, A. P. C. Silva, J. M. Almeida, and M. A. Gonçalves, “Fine-grained tourism prediction: Impact of social and environmental features.” Inf. Process. Manage., vol. 58, Feb. 2020, Art. no. 101670.

[30] M. A. Colominas, G. Schlotthauer, and M. E. Torres, “Improved complete ensemble EMD: A suitable tool for biomedical signal processing.” Biomed. Signal Process. Control, vol. 14, pp. 19–29, Nov. 2014.

[31] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis.” Proc. Roy. Soc. London A, Math., Phys. Eng. Sci., vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[32] Z. Wu and N. E. Huang, “Ensemble empirical mode decomposition: A noise-assisted data analysis method.” Adv. Adapt. Data Anal., vol. 1, no. 1, pp. 1–41, Jan. 2009.

[33] O. Amir, M. Mostefa, P. Mangor, and J. Graeme, “Range entropy: A bridge between signal complexity and self-similarity.” Entropy, vol. 20, no. 12, pp. 1–22, 2018.

[34] R. Tibshirani, “Regression shrinkage and selection via the lasso.” J. Roy. Statist. Soc., B Methodological, vol. 58, no. 1, pp. 267–288, 1996.

[35] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. Mcvean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, “Detecting novel associations in large data sets.” Science, vol. 334, pp. 1518–1524, Dec. 2011.

[36] V. Vapnik, “The nature of statistical learning theory.” Technometrics, vol. 38, no. 4, p. 409, 1995.

[37] F. X. Diebold and R. S. Mariano, “Comparing predictive accuracy.” J. Bus. Econ. Statist., vol. 20, no. 1, pp. 134–144, Jan. 2002.

[38] H. Hassani and E. Silva, “A Kolmogorov-Smirnov based test for comparing the predictive accuracy of two sets of forecasts.” Econometrics, vol. 3, no. 3, pp. 590–609, Aug. 2015.

[39] O. Claveria and S. Torra, “Forecasting tourism demand to catalonia: Neural networks vs. time series models.” Econ. Model., vol. 36, pp. 220–228, Jan. 2014.

HONGWEI WANG was born in China, in 1978. He is currently the Vice Dean of the School of Traffic and Transportation, Lanzhou Jiaotong University. His current research interests include tourism management and planning, intelligent transportation systems, regional comprehensive traffic planning, and transportation planning and management.

WENZHENG LIU was born in Gansu, China, in 1996. He received the M.S. degree from the School of Traffic and Transportation, Lanzhou Jiaotong University, in 2021. His research interests include machine learning, data mining, and tourism demand forecasting.