Annals of Tourism Research 90 (2021) 103271

Contents lists available at ScienceDirect

Annals of Tourism Research

journal homepage: https://www.journals.elsevier.com/annals-of-tourism-research

Multi-attraction, hourly tourism demand forecasting

Weimin Zheng [a], Liyao Huang [a,*], Zhibin Lin [b]

[a] School of Management, Xiamen University, 422 South Siming Road, 361005 Xiamen, China
[b] Durham University Business School, Mill Hill Lane, Durham DH1 3LB, United Kingdom

ARTICLE INFO

Article history:
Received 15 March 2021
Received in revised form 21 June 2021
Accepted 22 June 2021
Available online xxxx

Editor in Chief: Haiyan Song

Keywords:
Tourism demand forecasting
Spatial-temporal effect
Correlated time series
Long-short-term-memory
Attention mechanism

ABSTRACT

Forecasting tourism demand for multiple tourist attractions on an hourly basis provides important insights for effective and efficient management, such as staffing and resource optimization. However, existing forecasting models are not well equipped to hand the hourly data, which is dynamic and nonlinear. This study develops an improved, artificial intelligent-based model, known as Correlated Time Series oriented Long Short-Term Memory with Attention Mechanism, to solve this problem. The validity of the model is verified through a forecasting exercise for 77 attractions in Beijing, China. The results show that our model significantly outperforms the baseline models. The study advances the tourism demand forecasting literature and offers practical implications for resource optimization while enhancing staff and customer satisfaction.

Introduction

Accurate demand forecasting enables tourism organizations to properly arrange resources in advance to serve the demand (Jiao et al., 2020; Song & Li, 2008). Numerous forecasting models have been suggested and tested (Bi et al., 2020), including econometric models, time series models, and artificial intelligence (AI) models (Jiao & Chen, 2019). Despite the rapid progress in tourism demand forecasting research in recent years, many issues remain (Huang et al., 2019).

First, most existing models aim at relatively long time spans, such as monthly and annual forecasting (Alvarez Diaz & Mateu-Sbert, 2011; Bi et al., 2020; Divino & McAleer, 2010). However, for daily operations, finer time granularity and higher frequency tourism demand forecasting are required. For example, based on the hourly tourism demand forecasting results, tourist attraction managers can formulate effective real-time crowd management strategies, and optimize staff and resource arrangements (Song & Li, 2008; Wu et al., 2017). Hourly data is more dynamic and non-linear than other granular data (e.g., monthly, quarterly, or annual tourist arrivals) (Pereira & Nobre, 2016; Yang et al., 2014), making the forecasting more challenging. To date, little research has been devoted to hourly demand forecasting.

Second, most existing studies focus on the prediction of tourism demand for a single entity, e.g., a tourism destination (Assaf et al., 2019; Li & Law, 2020; Song et al., 2019), a tourist attraction (Bi et al., 2020; Li et al., 2020), or a hotel (Assaf & Tsionas, 2019; Yang et al., 2014), rather than multiple ones. Tourist attractions in a destination are geographically connected and tourists prefer multi-attraction visits; as such, supply interactions occur between attractions (Jiao & Chen, 2019; Long et al., 2019; Yang & Wong, 2012). Therefore, incorporating spatial information could potentially improve the accuracy of forecasting (Long et al., 2019; Stewart & Vogt, 1997; Yang & Zhang, 2019). Empirical research on this issue has not appeared until recently (Jiao et al., 2020). The few studies that consider the spatial effect are still based on the traditional time series model (Jiao et al., 2020; Yang &

* Corresponding author.
E-mail addresses: zhengweimin@xmu.edu.cn, (W. Zheng), huangly@stu.xmu.edu.cn. (L. Huang).

https://doi.org/10.1016/j.annals.2021.103271
0160-7383/

Zhang, 2019) or the econometric model (Emili et al., 2020; Long et al., 2019). However, these traditional models are unsuitable for handling nonlinear patterns and exogenous variables (Jiao & Chen, 2019).

This study aims to narrow the gaps in the demand forecasting literature by developing an improved AI-based model, known as the Correlated Time Series oriented Long Short-Term Memory with Attention Mechanism (CTS-LSTM-AM), to solve the research problem. The model compensates for the lack of spatial dependence characteristics between sequences extracted by LSTM (Wan et al., 2020) and could perfectly fit into tourism demand forecasting (Law et al., 2019). The validity of the model is verified through its application to forecasting the hourly visitor numbers for 77 attractions in Beijing, China.

The main contribution of the present study is the development of an improved model that incorporates spatial effects. Forecasting the demand of multiple tourist attractions with spatial effect from the hourly granularity is challenging because the hourly data is dynamic and nonlinear (Höpken, Eberle, Fuchs, & Lexhagen, 2021), and intra-sequence time and inter-sequence space dependencies should be captured for forecasting (Wan et al., 2020). This study improves the architecture of LSTM and expands the application of AI-based models in demand forecasting. The proposed model expands demand forecasting literature by focusing on hourly granularity, and such an approach is particularly helpful for managers' decision-making.

Literature review

Tourism demand forecasting

Tourism demand forecasting is a research hotspot with highly practical implications (Song et al., 2012), and various models have been proposed to advance forecasting accuracy in the past few decades. The most commonly used models are econometric and time series models (Jiao & Chen, 2019). Based on historical data, time series models can figure out the trends, and then predict the future (Song & Li, 2008). Econometric models can forecast demand by identifying tourist arrival relevant factors (Cao et al., 2017; Pan & Yang, 2017; Song & Wong, 2003). These models require advanced diagnosis and fixed time lag to realize the prediction. However, the time dependence may change, which suggests that the above models might not make full use of the time dependence of time series (Bi et al., 2020).

Recently, AI-based models have gained popularity, and a variety of models have been designed, including artificial neural networks (ANN), fuzzy time series, grey theory, support vector machines, rough sets approach, and hybrid models (Jiao et al., 2020). These traditional AI-based models have several weaknesses. First, with the increase of data frequency, handling the data becomes very challenging for these models, due to the non-stationarity, seasonality, and complexity of the data (Yang et al., 2014). Second, these traditional models cannot easily learn the tourist flow time series' long-term dependency (Bi et al., 2020). Third, they cannot automatically extract the data features and often encounter problems of overfitting or local optima (Zhang, Li, Muskat, & Law, 2021; Zhang, Li, Shi, & Law, 2020). Comparatively, deep learning techniques have the capability to extract discriminative features without the need for much human effort and domain knowledge (Pouyanfar et al., 2019). Research based on deep learning methods is becoming popular, among which LSTM is a typical one (Bi et al., 2020; Law et al., 2019). For instance, Law et al. (2019) put forward a deep network architecture to forecast tourist arrivals in Macau, demonstrating the superiority of this approach over traditional ones. Bi et al. (2020) incorporate multivariate time series data into the LSTM network, further improving forecasting performance.

However, the use of LSTM in tourism demand forecasting still presents several challenges, e.g., the deficiency of information processing (Kulshrestha et al., 2020), model overfitting, and high complexity (Zhang, Li, Muskat, & Law, 2021; Zhang, Li, Muskat, Law, & Yang, 2020). Thus, several scholars have attempted to improve the forecasting performance by improving LSTM to address the challenge. Kulshrestha et al. (2020) proposed a variant of LSTM, namely, Bayesian Bidirectional LSTM (BBiLSTM), to predict quarterly tourist arrivals. In their model, either backward or forward information can be utilized, and the hyperparameters can be optimized. Zhang et al. have also done considerable constructive work: they incorporated the group-pooling method into LSTM structure to predict the tourist arrival volumes (Zhang, Li, Muskat, Law, & Yang, 2020); and introduced a decomposed deep learning approach that combines trend and seasonal decomposition through the Loess and duo attention layer structure (Zhang, Li, Muskat, & Law, 2021).

Spatial effects in tourism demand

Spatial effects, including spatial spillover and spatial heterogeneity, have been well examined in previous studies (Balli et al., 2015; Li et al., 2016; Yang & Fik, 2014). Spatial dependence and regional interactions are important areas of research in tourism geography and economics (Fingleton & López-Bazo, 2006), particularly in regional tourism growth (Ma et al., 2015; Zhang et al., 2011) and destination tourism flows (Yang & Fik, 2014; Yang & Wong, 2012).

The spillover effect is mainly a result of multi-destination tourism (Yang, Fik, & Zhang, 2017). As an economic externality, spatial spillovers may occur through the sharing of common infrastructure and resources, collaboration or competition among destinations, or various events that happened in the area (Chhetri et al., 2013). Spatial proximity and attraction compatibility help drive the spillover effects in tourist flow from one attraction or destination to another (Weidenfeld, Butler, & Williams, 2010), because tourists generally visit multiple attractions or destinations in a single trip (Santos et al., 2011). Empirical evidence has confirmed the spillover effects. For example, Balli et al. (2015) reveal the spillovers of tourism demand from three main hubs to other regions in the country. Gooroochurn and Hanley (2005) found the spiller overs of tourism demand between the two main regions

in the Irish island are significant. Cao et al. (2017) showed evidence to support the tourism demand interdependence across countries. More recently, Assaf et al. (2019) reported the spillovers of international tourism demand in nine Southeast Asian countries.

Due to each location's unique characteristics, spatial heterogeneity is associated with spatial differentiation that is due to each location's unique characteristics. It reflects the differences between regions (Zhang et al., 2011). These differences can be reflected in the distributions and means of a set of variables and their variance and covariance. For instance, the various geographic regions in China have large heterogeneity in terms of history, culture, and economic growth (Lin et al., 2018). Using panel data, Li et al. (2016) revealed the tourism income inequality between different regions in China. Examining the environmental characteristics and historical evolution of places in Italy, Considering the social, economic and cultural, economic diversity of Spain, Cotos-Yáñez et al. (2018)) revealed the geographic heterogeneity in senior tourists' travel patterns.

Previous researches on identifying spatial spillover and spatial heterogeneity have provided important insights for tourism development (Capone & Domenech, 2008; Khadaroo & Seetanah, 2008). However, research on incorporating spatial effects into forecasting models is at an early stage. Only four recent studies have taken into account the spatial effects. Specifically, Yang and Zhang (2019) used dynamic spatial panel models and STARMA. Long et al. (2019) adopted the pooled ordinary least squares model by incorporating spatial and temporal effects. Emili et al. (2020) used a dynamic panel model through the difference Generalized Method of Moments estimation approach. Jiao et al. (2020) further extended the basic ARIMA model and applied spatial lag in response variables and moving average terms.

Despite the advances made by the recent studies (Emili et al., 2020; Jiao et al., 2020; Long et al., 2019; Yang & Zhang, 2019), the extant literature of tourism demand forecasting considering spatial effect has several limitations. First, the spatial-temporal models adopted in these studies rely on time series (Jiao et al., 2020; Yang & Zhang, 2019) or econometric models (Long et al., 2019), which cannot easily adequately capture nonlinear patterns and exogenous variables (Jiao & Chen, 2019). Second, these studies focus on long-term forecasting of relatively large areas, whereas very few models forecast tourism demand at the attraction level at a short time scale, such as hourly. Thus, this study attempts to advance the field by proposing a new model that can capture temporal and spatial dependencies and simultaneously predict the demand of multiple tourist attractions within a destination from hourly granularity.

Proposed model

Architecture of standard LSTM

As an improved version of traditional neural networks, recurrent neural networks (RNN) have good performance in dealing with timing prediction problems due to their ability to capture the time correlation between events that are far apart in the sequence (Li & Law, 2020). However, RNN has inherent problems, such as vanishing gradients or exploding ones, when capturing the long-term dependence of time series (Wan et al., 2020; Xu et al., 2018).

LSTM networks, first proposed by Hochreiter and Schmidhuber (1997), are capable of learning long-term dependencies, overcoming RNNs' inherent problems. LSTM networks remember information for long durations and run well to solve many different practical problems (Bi et al., 2020). Unlike RNNs, but the LSTM networks have four layers of chain structure: a memory cell and three basic structures named “gates” (forget, input, and output gates) to realize the protection and control of information (shown in Fig. 1). The gate allows selective passage of information, largely through a point-by-point multiplication operation. First, the forget gate (σ_f) determines the removal of specific cell state information. The information source of the forget gate is the hidden layer vector at time *t-1* ($\boldsymbol{h_{t-1}}$) and the time series information at time *t* ($\boldsymbol{x_t}$) (Xu et al., 2018). Second, the input gate (σ_i) decides the addition of new information to the cell state, which involves an input gate layer and a tanh level. The *tanh* level is used as an excitation function to determine the increase or decrease of information (Wan et al., 2020). According to the cell state, the output gate (σ_o) controls the output information. Finally, we get the hidden layer vector ($\boldsymbol{h_t}$) and the cell state of the current time ($\boldsymbol{C_t}$). For more details of LSTM, please refer to Law et al. (2019) and Bi et al. (2020).

Architecture of CTS-LSTM-AM

The tourist attractions in a destination are geographically connected. They may demonstrate strong relevance of tourism demand given the competitive and cooperative relations between attractions (Jiao & Chen, 2019; Long et al., 2019; Yang & Wong, 2012). The time series of tourism demand in attractions are influenced and dependent on each other, which can be denoted as correlated time series. Therefore, to improve the forecasting performance, attention should be paid to the correlated time series of tourists' attractions in the destination rather than a single time series of an attraction. However, in correlated time series, the conventional LSTM is unable to capture the spatial-temporal features fully (Wan et al., 2020). In addition, LSTM cannot learn the importance of different input elements in the same time series.

To deal with these challenges, inspired by the study of Wan et al. (2020) and Law et al. (2019), this study proposed a novel AI-based model named CTS-LSTM-AM. This model adds an attention mechanism on the original CTS-LSTM proposed by Wan et al. (2020) to distinguish the differences in the impact of different input elements in the same time series on the prediction performance, thus making it better adapted to the scenario of tourism demand forecasting. In other words, in addition to the ability of CTS-LSTM that can simultaneously capture inter-sequence space dependencies and intra-sequence dependencies in time series, our model is able to disregard irrelevant information and as a result, it is highly interpretable. Fig. 2 illustrates the architecture

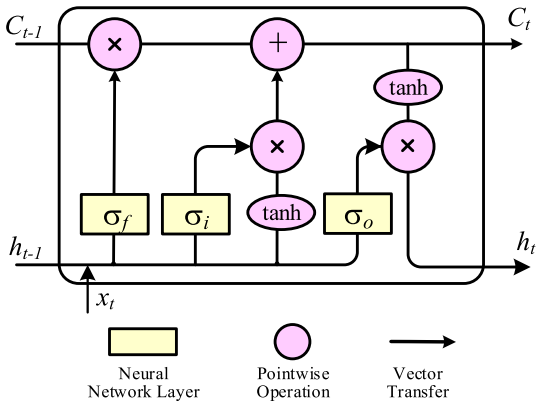

Fig. 1. Network structure of LSTM.

of our model, with three paramount modules: attention mechanism, spatio-temporal cells (ST-cell), and spatio-temporal fusion (ST-fusion).

We define the input series at time t as $x_t \in \mathbb{R}^{\mu \times N}$, where μ indicates the number of embedding representation dimensions, and N denotes the correlated time series in terms of their number. Then, the input matrix x_t passes the attention mechanism, which is to assign different weights to the time steps within the intra-sequence. Subsequently, the ST-cell module is modified based on the standard LSTM so that it can encapsulate the inter-sequence space and intra-sequence time dependencies in the correlated time series. ST-fusion is located at the top of the ST-cell module, and its main function is to fuse the information captured by the ST-cell. Finally, the predicted value of matrix $\boldsymbol{x}$ was obtained, denoted as $\hat{x}_t$. Backpropagation and Adam algorithm are used to train our network, to reduce the mean square error between the actual matrix x_t and the predicted matrix $\hat{x}_t$. Thus, a loss level is introduced to evaluate the forecasting performance.

Attention mechanism

As a feature engineering method, the attention mechanism can be integrated into the model and works along with models that are LSTM-based. The standard LSTM cannot pay special attention to the important features of the sequence, which can be compensated by the attention mechanism. This is very important in tourism demand forecasting, because the attention mechanism enables the model to capture the entire flow dynamics in the input sequence and pay attention to factors related to tourism demand to improve the interpretation of the model (Law et al., 2019). Specifically, in the attention mechanism, the context vector $\boldsymbol{v}$ from the given time series ($\boldsymbol{X} = (x_1, x_2, \ldots, x_i)$) is extracted. $\boldsymbol{v_t}$ is the weighted sum of each column x_i in $\boldsymbol{X}$, representing information related to the present time step. And $\boldsymbol{v_t}$ is then incorporated with the current state x_t to generate predictions (Shih, Sun, &

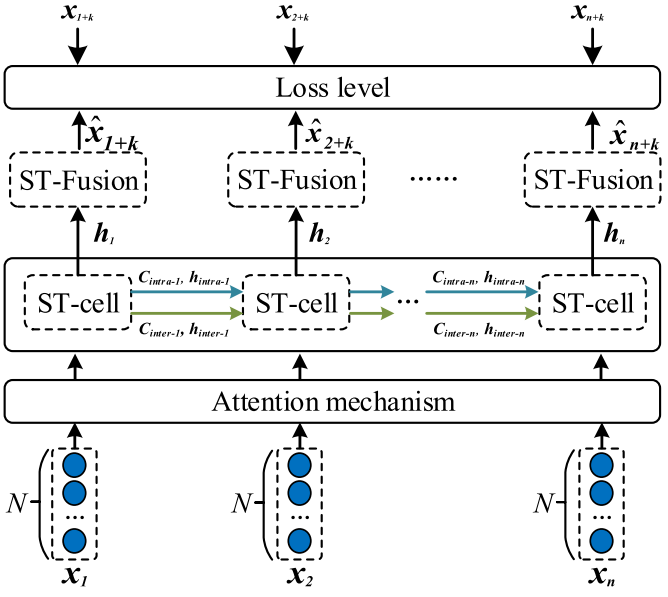

Fig. 2. The architecture of CTS-LSTM-AM.

Lee, 2019). We can formulate this process based on Eq. (1), where $\boldsymbol{a_i}$ is the attention weight, which can automatically capture the correlation between x_i and the predicted target $\hat{x}_i$. For more details of the attention mechanism, please refer to Law et al. (2019).

$$v_t = \sum_{i=1}^{t} a_i x_i \tag{1}$$

Spatio-temporal cell

The spatio-temporal cells are equipped with two channels to obtain the spatial-temporal dependencies. As shown on the left-hand side of Fig. 3, the intra-sequence cell state captures time dependence, consistent with the conventional LSTM. Similarly, we define three gates $\boldsymbol{I_{intra\text{-}t}}$, $\boldsymbol{F_{intra\text{-}t}}$, $\boldsymbol{O_{intra\text{-}t}}$ to determine the amount of information of each embedded feature contained in each time series to be retained in the cell state in each time step. As shown in Eqs. (2)–(7), ST-cell basically retains the architecture of LSTM. Its biggest difference from LSTM is that ST-cell builds a separate representation for each sequence. The advantage of this is that the update of each cell depends on the characteristics of each sequence, and the information is extracted to the maximum.

$$I_{intra\_t} = sigmoid(W_i[X_{intra\_t}, h_{intra\_t-1}] + b_i) \tag{2}$$

$$F_{intra\_t} = sigmoid\left(W_f[x_{intra\_t}, h_{intra\_t-1}] + b_f\right) \tag{3}$$

$$O_{intra\_t} = sigmoid(\mathrm{w_o}[\mathrm{x_{intra\_t}}, \mathrm{h_{intra\_t-1}}] + \mathrm{b_o}) \tag{4}$$

$$\tilde{C}_{\mathrm{intra\_t}} = relu(\mathrm{W_c}[\mathrm{x_{intra\_t}}, \mathrm{h_{intra\_t-1}}] + \mathrm{b_c}) \tag{5}$$

$$\mathrm{C_{intra\_t}} = \mathrm{I_{intra\_t}} {}^{\circ} \tilde{\mathrm{C}}_{\mathrm{intra\_t}} + \mathrm{F_{intra\_t}} {}^{\circ} \mathrm{C_{intra\_t-1}} \tag{6}$$

$$\mathrm{h_{intra\_t}} = \mathrm{O_{intra\_t}} {}^{\circ} relu(\mathrm{C_{intra\_t}}) \tag{7}$$

In the right-hand side of Fig. 3, the channel models the inter-sequence spatial effect. The largest difference from the basic LSTM is that the characterization of each series is reconstructed by fusing the information of all series through the spatial matrix $\boldsymbol{S}$ before the sequence enters the ST-cell. Here, $S \in R^{N\times N}$ reflects the paired influence between related series. A fixed matrix can be set to reflect a domain's knowledge a priori. For example, the geospatial correlated time series $\boldsymbol{S}$ may represent an inter-attraction similarity or proximity matrix in our case. Specifically, we use the inverse physical distances between attractions to develop an $\boldsymbol{S}$ matrix. This step enables the representation of each series to include information about the other series, thereby capturing the interactions between related sequences, that is, the spatial dependencies between sequences. As with the intra-sequence temporal channel, we define three gates $\boldsymbol{I_{intra\_t}}$, $\boldsymbol{F_{intra\_t}}$, $\boldsymbol{O_{intra\_t}}$ to determine the amount of information for each embedded feature within each time series to be retained in the cell state in each time step, using sigmoid as the activation function.

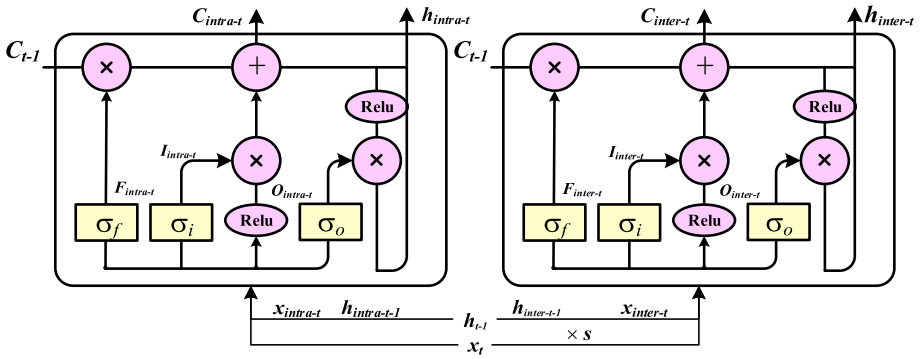

Fig. 3. Architecture of ST-cell.

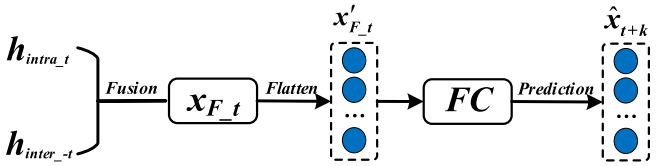

Fig. 4. Process of ST-fusion.

Spatio-temporal fusion

ST-fusion is located at the top of the ST-cells, and its function is to fuse the information obtained in ST-cells. Two tasks are contained in the fusion module, as shown in Fig. 4. First, the information of intra-sequence temporal channel (h_{intra_t}) and inter-sequence spatial channel (h_{inter_t}) achieved in ST-cells are assigned different weights. Then, they are fused to obtain the fusion matrix x_{F_t} according to the weights, as shown in Eq. (8). The weights involved in this process are learnable parameters. Second, the fusion matrix x_{F_t} is flattened to achieve a new matrix x_{F_t}', which serves as the input of the fully connected (FC) layer. Through the FC layer, the spatio-temporal features of the sequences can be better extracted, and the predicted value $\hat{x}_{t+k}$ can be calculated according to Eq. (9). In this equation, ω represents the weight of x_{F_t}', b is the bias, **ReLU** is the activation function, and "∘" denotes element-wise multiplication.

$$x_{F\_t} = w_1 \circ h_{intra\_t} + w_2 \circ h_{inter\_t} \tag{8}$$

$$\hat{x}_{t+k} = ReLu\left(\omega \cdot x'_{F\_t} + b\right) \tag{9}$$

Empirical test

Empirical research was conducted based on tourist attractions in Beijing, the capital of China. Its numerous world-famous tourist attractions attract a large number of international tourists every year (Zhang et al., 2019; Zhang, Yang, Zhang, & Zhang, 2020). A map of Beijing with the distribution of 77 tourist attractions is shown in Fig. 5.

Data collection and preprocessing

Hourly tourist arrival volumes of tourist attractions in Beijing are available from the Beijing Tourism Network (http://www.visitbeijing.com.cn/), which provides real-time visitor flow and congestion information of 233 × 2A and above-rated attractions in the city (5A is the highest rating, based on the tourism attraction classification system in China). The data is updated every 15 min. We collected the data from October 1st, 2020 to October 31st, 2020. Among these 233 attractions, 81 of them were 4A or 5A, 115 are 3A, and the remaining 37 are 2A attractions. Four highly-rated attractions (4A and above) were not open to the public during this period. Therefore, we focused on the 77 attractions with 4A or 5A ratings, and their distributions are shown in Fig. 5.

The raw data of 15-minute granularity was transformed into the data of hourly granularity using the average method. Fig. 6 shows the samples of six of these attractions (Summer Palace, The Old Summer Palace, Olympic Forest Park, Prince Gong Mansion, Forbidden City, and Temple of Heaven). The figure indicates that these time series have different degrees of correlations. We only used the data from 9:00 to 18:00 every day as the dataset for this study, based on the opening hours of most attractions. We used Python to predict the time series. By normalizing the zero-mean value of the data, we divided them into training and test datasets. The 180 samples from October 1st to October 20th were collected to train the model, while another 99 samples from October 21st to October 31st were used to test the model.

Performance evaluations

Several key parameters such as time step, batch size, and learning rate should be set in advance for training the LSTM networks. We used historical data to forecast the hourly demand from 9:00 to 18:00 each day following a window rolling approach. We set the timestep as 9, based on the period of the most common opening hours of all the attractions. For the selection of batch size [16, 32, 64, 128, 256, 512] and learning rate [0.1, 0.01, 0.005, 0.001, 0.0005, 0.0001], we determined through an exhaustive grid method and calculated the RMSE value of each combination (as shown in Fig. 7). The results were as follows: the lighter the color, the better the prediction effect of the parameter combination. The parameter combination we chose is batch size 64 and learning rate 0.005.

To achieve a better prediction of tourism demand, we focused on the different weights of time steps and the spatial correlation between attractions. We added the attention mechanism to implement allocate different weights of time steps. Spatial correlation was obtained by assigning spatial matrix $\boldsymbol{S}$ to sequences, where $\boldsymbol{S}$ is composed of the reciprocal of the geospatial distance between sequences (Wan et al., 2020). Many traditional LSTM models encounter problems of overfitting, to deal with this problem, our

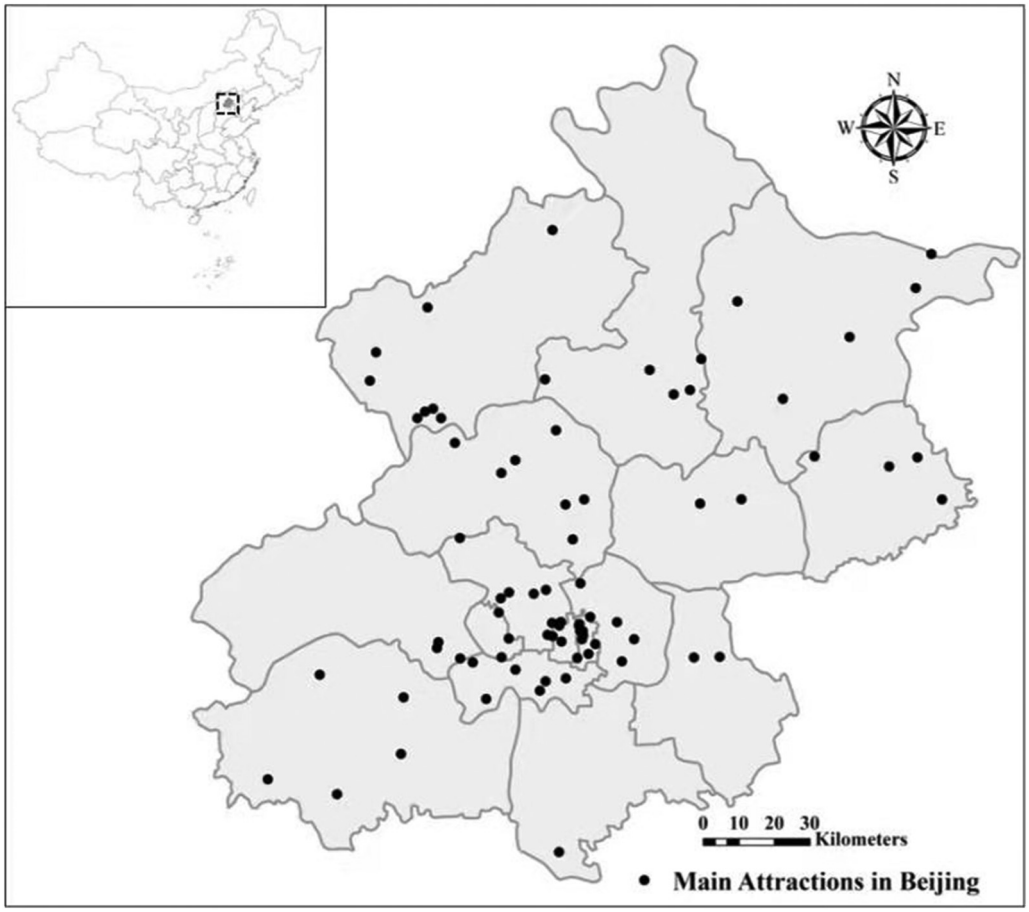

Fig. 5. Map of Beijing City.

model adds a dropout regularization in the recurring layer. Dropout regularization refers to the random deletion of input elements in each update to reduce the possibility of overfitting (Xu et al., 2018). However, the dropout may lead to the randomness of model training, which makes the results of the model fluctuate (Bi et al., 2020). Therefore, this study obtains the model results by averaging those over five runs, following the practices used in prior studies (e.g. Bi et al., 2020).

To investigate the model's performance, we used SARIMAX (Tsui & Balli, 2017), ANN (Law, 2000), LSTM (Bi et al., 2020), LSTM-AM (Law et al., 2019), and CTS-LSTM (Wan et al., 2020) as the baseline models. SARIMAX (p, d, q) $(P, D, Q)_s$ is a variant of the autoregressive integrated moving average (ARIMA) model. Three parameters are involved in SARIMAX: p is the autoregression order, q represents the moving average's order, and d denotes the number of differences made when time becomes stationary; (p, d, q) indicates the model's non-seasonal component, and (P, D, Q) indicates the seasonal component. In this study, a fully automated procedure provided in R, called auto.arima, was adopted to select the parameters of SARIMAX. We used an automatic traversal of the parameter combinations to obtain the model with the minimum AIC value. The ANN model was trained with the back-propagation algorithm to predict $\widehat{y}_{t+9}$. *Relu* was selected as the excitation function, and a fully connected layer is built in the ANN model. The LSTM model has been widely used in previous studies. Thus, we used this model as the baseline. To maintain consistency, we set the parameters of LSTM to be the same as those of CTS-LSTM-AM, that is, learning rate $= 0.005$, batch size $= 64$. LSTM-AM is an improved LSTM by adding AM, which enables the model to increase the explanatory power by focusing on the important features of the time series. The setting was consistent with that of LSTM. The CTS-LSTM model, proposed by Wan et al. (2020), is capable of capturing the spatial correlation of sequences. Hence, it is also used as a benchmark model. As with CTS-LSTM-AM, the batch size and learning rate of CTS-LSTM were 64 and 0.005, respectively.

In this study, the rolling window method was served for the prediction to better imitate reality (Law et al., 2019). Three commonly used criteria were used to assess to model performance evaluation, namely, mean absolute deviation (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE), as described in Eqs. (10)–(12): where f_i is the actual tourism demand data, $\hat{f}_i$ is the tourism demand to be predicted, and n is the number of test samples. The original normalized data were restored to calculate the above indicators. A smaller value indicates a better performance. Even if the operating hours of most

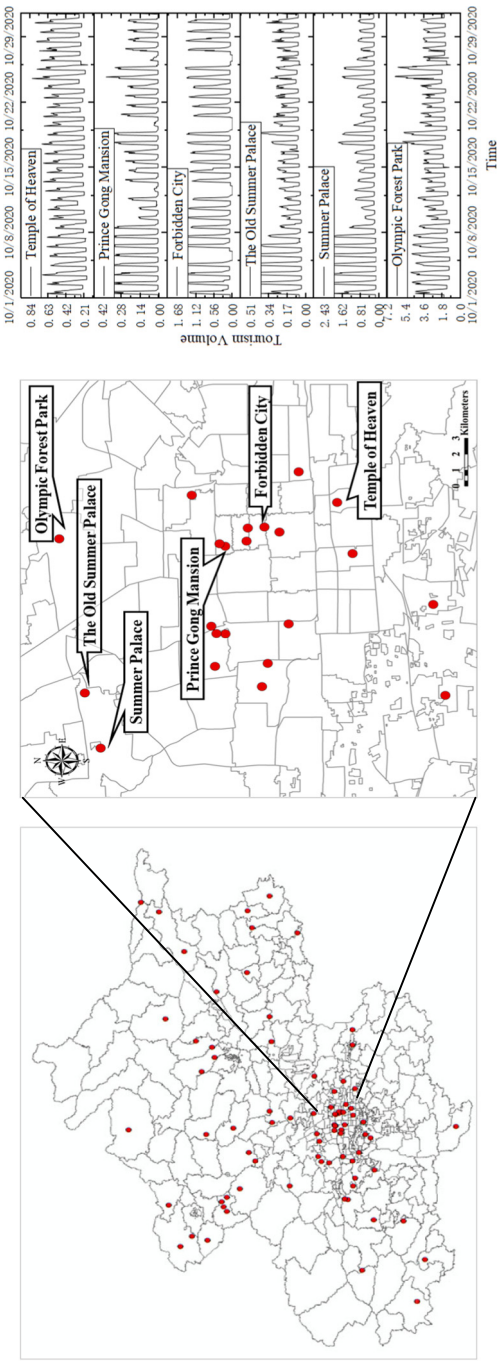

Fig. 6. Tourist arrival volumes of different tourist attractions.

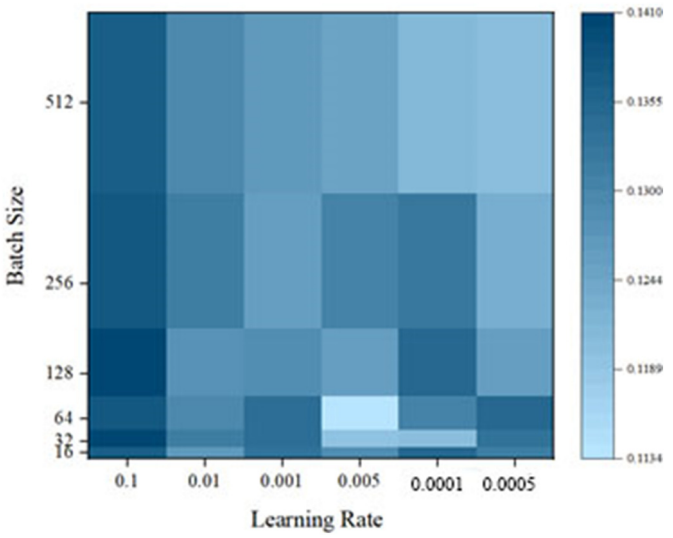

Fig. 7. The values of RMSE concerning different parameter combinations.

attractions were selected for forecasting, there were still situations where the demand for some attractions is 0. Therefore, in order to accurately calculate the MAPE value, we added 1 for the demand of all attractions before forecasting.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | f_i - \hat{f}_i | \tag{10}$$

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n} \left| f_i - \hat{f}_i \right|^2\right]^{\frac{1}{2}} \tag{11}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{f_i - \hat{f}_i}{f_i} \right| \tag{12}$$

Results

Model performance test

Using the obtained parameters, the CTS-LSTM-AM model was first trained using a training sample set. The hourly tourist volume from October 21st to October 31st, 2020 was then forecasted. In accordance with the optimal parameters, the 99 samples were forecasted using our model and five baseline models (SARIMAX, ANN, LSTM, LSTM-AM, and CTS-LSTM), respectively. The discrepancy between the predicted and actual tourism volume of the 77 attractions obtained by the six models is presented in Fig. 8, where the color and the degree of deviation from the 0-horizontal plane reflect the discrepancy between the predicted and actual volume. The MAE/RMSE/MAPE values for each attraction obtained by these six models are shown in Fig. 9.

The average MAE/RMSE/MAPE values of each model are presented in Table 1 and the MAE/RMSE/MAPE values of each attraction are presented in Appendix 1. The results show that the CTS-LSTM-AM model achieved minimal errors compared with the baseline models. The precision of MAPE drops from 0.388 to 0.186, which is a significant achievement.

We used Diebold-Mariano (DM) test to assess whether there were statistically significant differences between the new model and the baseline models. The DM method (Diebold & Mariano, 2002) is widely used in the studies of tourism demand forecasting (Bangwayo-Skeete & Skeete, 2015; Kulshrestha et al., 2020). Its null hypothesis is that two models have the same predictive power. Thus, a pairwise comparison between our model and the baseline model was performed. The results were all negative (see Appendix 2), indicating that the CTS-LSTM-AM achieved better prediction performance than all the baseline models. Specifically, our model outperforms ANN in all (100%) attractions, SARIMAX in 99% of the attractions (i.e., 76 out of 77), LSTM in 97% of the attractions (i.e., 75 out of 77), LSTM-AM in 95% of the attractions (i.e., 73 out of 77), and CTS-LSTM in 71% of the attractions (i.e., 55 out of 77).

Model discussion

There is a spatial effect in tourism demand due to the interaction of supply with neighboring areas (Long et al., 2019; Yang & Zhang, 2019). Therefore, our model considers the spatial effect by introducing correlated time series and attention mechanisms.

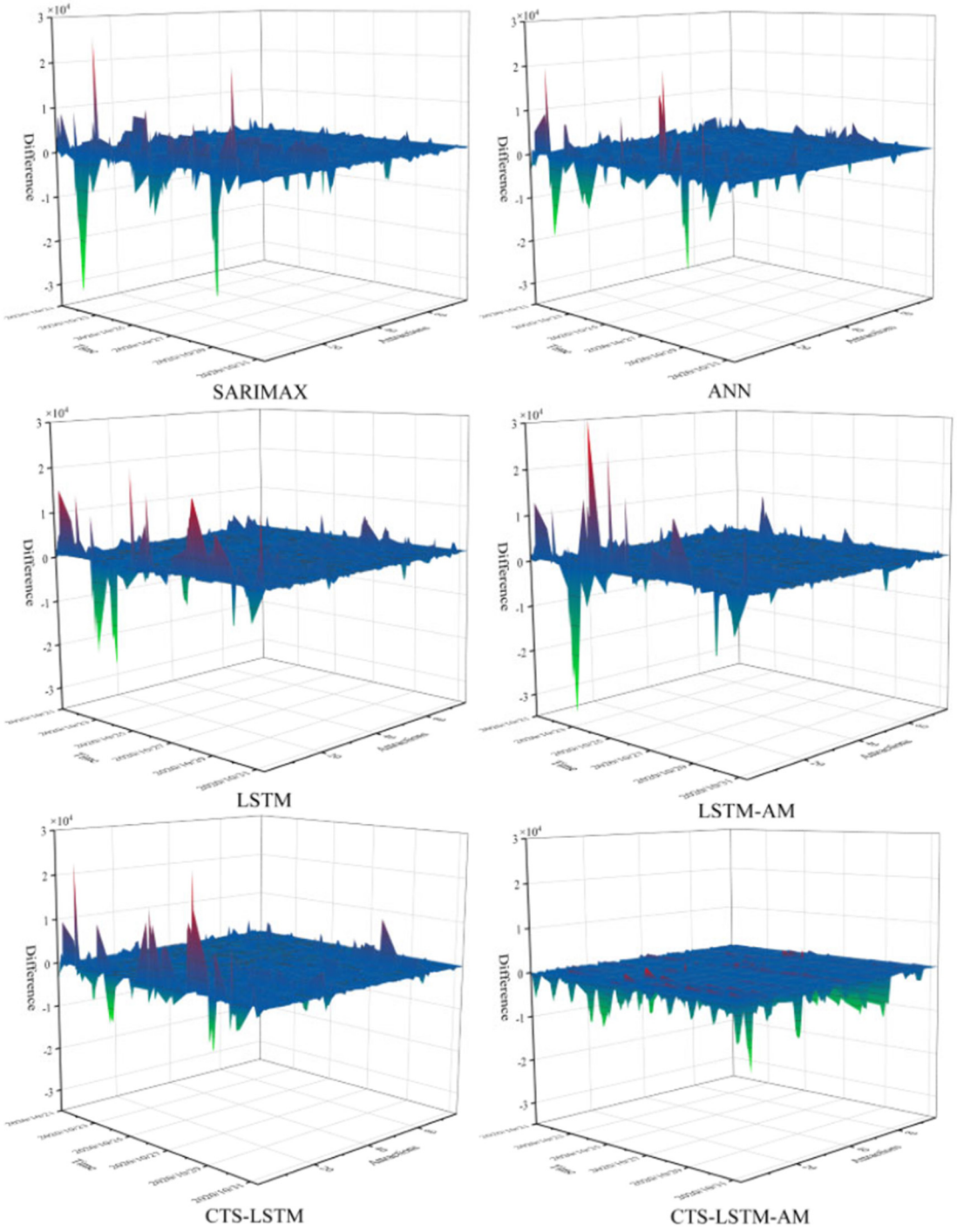

Fig. 8. Difference comparison between models.

To verify whether the consideration of spatial effect can improve forecasting performance, a baseline model was introduced for comparison. Compared with our model, this baseline model did not consider the demand series of the 76 other tourist attractions when forecasting the demand of an attraction. The prediction results of the two models are shown in Fig. 10, which indicates the discrepancy between the predicted and actual tourism volume of the 77 attractions. The mean values of MAE, RMSE and MAPE of the baseline model are 0.104, 0.068 and 0.452. The forecast accuracy of MAPE is 0.452, which is significantly larger than that of our model (0.186). Similarly, we further utilized the DM test to compare the two models' predictive accuracy (see the seventh column of Table A2-1 in Appendix 2). The results suggest that our proposed model is statistically better than the baseline

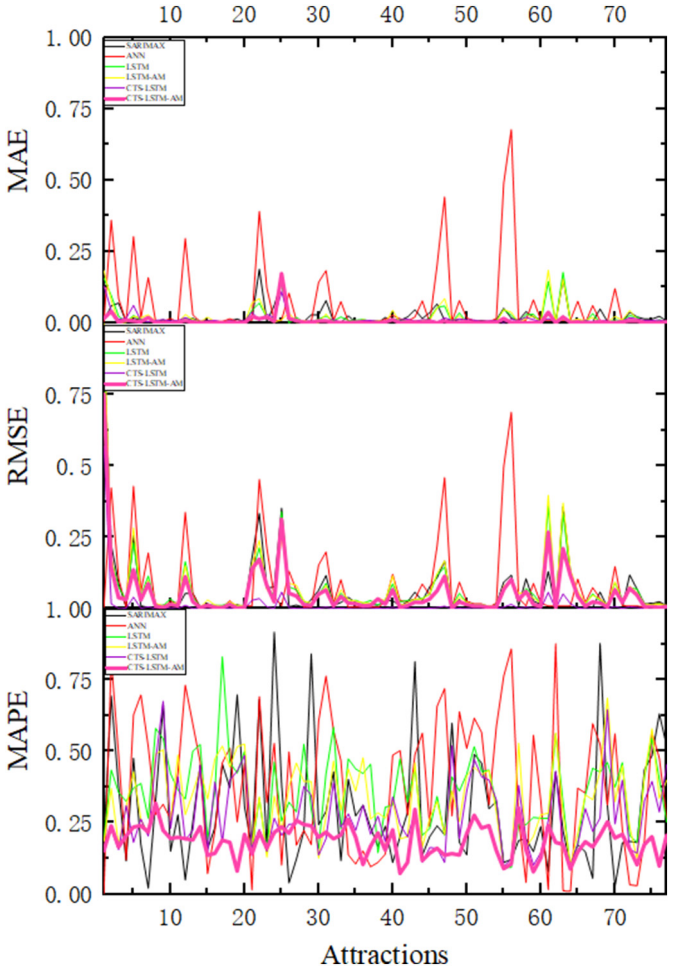

Fig. 9. MAE, RMSE, and MAPE comparison between models.

model in 99% of the attractions (i.e., 76 out of 77), which confirms that considering spatial effects in the model helps to improve the forecasting performance.

To further verify the model's robustness, we made detailed observations on the prediction results under different prediction time intervals. We selected the time intervals of 1 h, 3 h, and 6 h for comparative experiments. The MAPE values of the six models at different time intervals are shown in Fig. 11. The different line colors represent different models. The results demonstrate that our method outperforms all baseline models at different forecasting time intervals, confirming that our model has good robustness. Moreover, for all the six models, when the time interval was set as 9 h, its forecasting performance was better than the other three time intervals, confirming that setting the time interval to 9 h in the above analysis was appropriate.

Conclusion

This study addresses the challenges of forecasting hourly tourist arrivals in multiple tourist attractions. Our proposed model uses the spatial effect via correlated time series and attention mechanism to the forecasting model to simultaneously forecast the tourist volume for multiple attractions in a destination. The empirical test shows that our model outperforms the baseline models. Previous studies often only target the demand forecast of a single area, such as a single attraction or a single tourist

Table 1
Average of MAE, RMSE and MAPE for each model.

	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
MAE	0.283	0.314	0.122	0.100	0.058	0.053
RMSE	0.301	0.334	0.153	0.110	0.075	0.071
MAPE	0.534	0.607	0.364	0.353	0.275	0.186

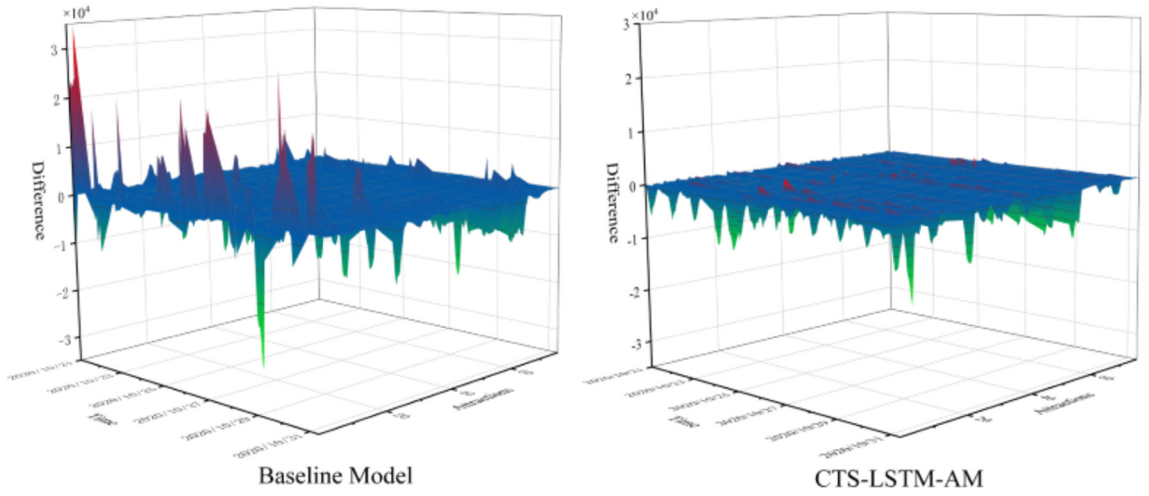

Fig. 10. Forecasting performance comparison.

destination. Our study considers the time series of a tourist attraction and that of other attractions in the destination; thus, the model simultaneously forecasts the demand of multiple attractions. Moreover, the extant studies mainly focus on long-term and mid-term forecasts. Short-term forecasts were relatively few. The major challenges are a) short-term data are more difficult to obtain and b) short-term data are more complex, with greater nonlinearity and randomness. Our proposed model effectively handles hourly data.

This study has several contributions. First, we advance the demand forecasting literature by proposing a model that captures both intra-sequence time and inter-sequence space dependencies, which have not been adequately addressed in prior studies. Second, we improve the architecture of LSTM and expand the application of the AI-based model, which can be widely applied with the capability to accommodate additional predicting variables. Two major improvements have been made: a) We consider correlated time series in LSTM to handle the historical demand data of multiple tourist attractions in a destination. b) We introduce an attention mechanism that can evaluate the importance of different elements of a time series so different weights are determined for different elements. Third, we further advance the demand forecasting field by presenting a model that is well

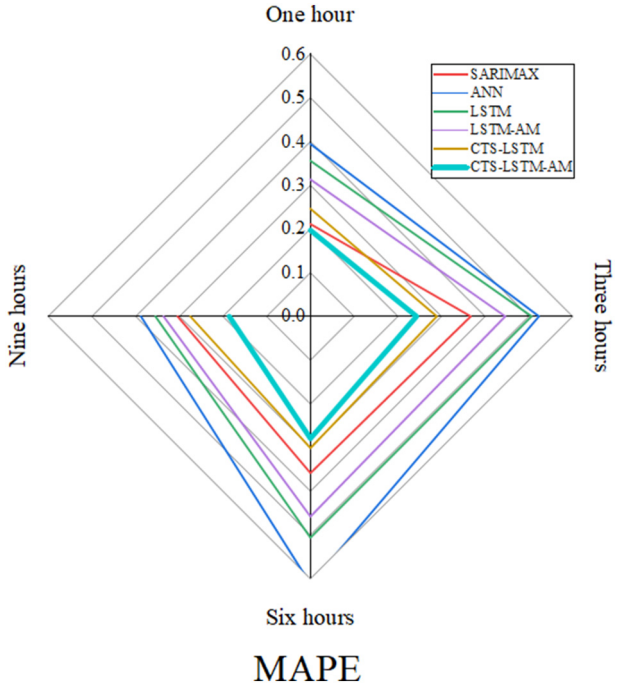

Fig. 11. Forecasting performance comparison for different time intervals.

equipped to handle the dynamic and non-linear characteristics of hourly data, thus providing finer granularity in the forecasting performance.

Our research also has practical implications. First, the model proposed in this study helps destination managers to perform micro-level forecasts based on high-frequency data, and to improve their resource planning and daily operation efficiency, thereby achieving higher levels of staff and tourist satisfaction. Second, the empirical results support that considering spatial effects can positively influence tourism demand forecasts. This suggests that destination managers should consider the spatial relevance between attractions in their planning and marketing. For example, attractions that are closely relevant and complementary with each other can be clustered together for joint promotion, while attractions that compete with each other with similar appeals can be repositioned.

One of the limitations of the study is our model's reliance on historical data. Future studies should consider several other variables that are not captured in our data, for instance, the data of search index (Yang et al., 2014) and weather (Bi et al., 2020). The data used in the study is based on short time intervals, which suits our research aims, however, tourism demand is highly influenced by seasonality (Xie et al., 2020), future studies may take seasonality into account by using longer interval tourist data to improve forecasting performance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research received grants from the National Natural Science Foundation of China (No. 71971179) and Natural Science Foundation of Fujian Province, China (No. 2020J01033).

Appendix 1. The MAE/RMSE/MAPE values of each attraction

Table A1-1
MAE values of each attraction.

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
1	1.240	3.541	1.125	0.843	0.530	0.452
2	0.004	0.001	0.012	0.020	0.003	0.004
3	0.109	0.084	0.053	0.049	0.028	0.025
4	0.925	0.859	0.268	0.199	0.088	0.091
5	0.004	0.001	0.016	0.024	0.003	0.004
6	0.200	0.111	0.051	0.045	0.031	0.028
7	0.060	0.070	0.055	0.058	0.030	0.022
8	0.004	0.040	0.019	0.015	0.003	0.004
9	0.266	0.613	0.156	0.105	0.080	0.082
10	0.021	0.001	0.028	0.035	0.012	0.012
11	0.070	0.172	0.067	0.078	0.047	0.042
12	0.361	0.274	0.138	0.113	0.074	0.064
13	0.406	0.094	0.115	0.128	0.050	0.043
14	0.036	0.035	0.061	0.056	0.015	0.015
15	0.216	0.171	0.079	0.031	0.025	0.024
16	0.115	0.310	0.131	0.107	0.080	0.072
17	0.310	0.162	0.088	0.082	0.063	0.067
18	0.382	0.165	0.162	0.110	0.054	0.055
19	2.181	2.475	0.816	0.730	0.470	0.447
20	0.005	0.001	0.019	0.026	0.007	0.008
21	0.328	0.338	0.157	0.111	0.077	0.078
22	0.440	0.488	0.189	0.228	0.108	0.094
23	0.319	0.063	0.117	0.082	0.061	0.052
24	0.005	0.039	0.014	0.016	0.004	0.003
25	0.921	0.965	0.250	0.201	0.232	0.224
26	0.122	0.578	0.207	0.123	0.075	0.085
27	0.048	0.091	0.038	0.040	0.022	0.022
28	0.993	0.836	0.267	0.187	0.090	0.092
29	0.016	0.042	0.034	0.034	0.012	0.011
30	0.046	0.083	0.050	0.044	0.048	0.041
31	0.128	0.001	0.048	0.053	0.024	0.022

(continued on next page)

Table A1-1 *(continued)*

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
32	0.365	0.577	0.149	0.075	0.067	0.062
33	0.096	0.107	0.051	0.062	0.029	0.022
34	0.004	0.001	0.011	0.020	0.003	0.003
35	0.960	0.764	0.211	0.129	0.108	0.091
36	0.550	0.102	0.133	0.088	0.062	0.056
37	0.021	0.001	0.021	0.025	0.009	0.010
38	1.002	0.884	0.267	0.207	0.093	0.088
39	0.032	0.001	0.024	0.028	0.010	0.008
40	0.004	0.001	0.012	0.013	0.003	0.004
41	0.004	0.036	0.016	0.019	0.003	0.004
42	0.033	0.031	0.028	0.020	0.007	0.008
43	0.007	0.059	0.018	0.012	0.005	0.005
44	0.020	0.045	0.043	0.047	0.016	0.014
45	0.109	0.117	0.040	0.049	0.016	0.014
46	0.100	0.108	0.077	0.078	0.045	0.031
47	0.151	0.186	0.128	0.113	0.074	0.059
48	1.023	0.345	0.259	0.223	0.165	0.149
49	0.062	0.122	0.032	0.028	0.019	0.019
50	0.357	0.918	0.274	0.216	0.090	0.086
51	0.112	0.379	0.156	0.113	0.077	0.080
52	0.004	0.026	0.016	0.018	0.003	0.004
53	0.026	0.103	0.024	0.030	0.009	0.009
54	0.563	0.651	0.171	0.090	0.085	0.079
55	0.495	0.211	0.158	0.100	0.055	0.056
56	0.528	0.318	0.206	0.136	0.077	0.086
57	0.005	0.001	0.016	0.019	0.003	0.004
58	0.127	0.171	0.062	0.067	0.029	0.035
59	0.015	0.080	0.038	0.030	0.011	0.011
60	0.008	0.001	0.020	0.023	0.006	0.005
61	0.031	0.087	0.040	0.034	0.022	0.023
62	0.452	0.567	0.155	0.114	0.077	0.078
63	0.918	0.775	0.239	0.302	0.156	0.121
64	0.889	0.648	0.267	0.192	0.089	0.085
65	0.198	0.126	0.061	0.072	0.033	0.036
66	0.019	0.034	0.020	0.020	0.007	0.006
67	0.057	0.001	0.049	0.040	0.016	0.014
68	0.189	0.269	0.168	0.159	0.075	0.066
69	0.564	0.601	0.149	0.114	0.055	0.054
70	0.109	0.080	0.090	0.115	0.040	0.029
71	0.010	0.079	0.021	0.024	0.009	0.009
72	0.253	0.121	0.085	0.073	0.046	0.044
73	0.264	0.294	0.093	0.067	0.050	0.049
74	0.028	0.001	0.031	0.023	0.007	0.008
75	0.379	0.561	0.166	0.094	0.055	0.057
76	0.011	0.001	0.019	0.016	0.003	0.003
77	0.329	0.899	0.252	0.189	0.092	0.078

Table A1-2
RMSE values of each attraction.

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
1	1.599	3.695	1.458	1.169	0.728	0.658
2	0.000	0.001	0.016	0.027	0.004	0.005
3	0.124	0.108	0.065	0.060	0.038	0.034
4	0.935	0.869	0.317	0.239	0.119	0.129
5	0.000	0.001	0.019	0.030	0.004	0.004
6	0.204	0.123	0.065	0.055	0.038	0.035
7	0.069	0.103	0.075	0.084	0.042	0.028
8	0.000	0.055	0.025	0.021	0.004	0.004
9	0.293	0.623	0.198	0.153	0.108	0.113
10	0.024	0.001	0.035	0.042	0.014	0.014
11	0.089	0.188	0.082	0.105	0.059	0.054
12	0.396	0.328	0.184	0.151	0.095	0.082
13	0.408	0.120	0.147	0.147	0.061	0.054
14	0.044	0.045	0.079	0.073	0.020	0.022
15	0.236	0.187	0.113	0.040	0.031	0.030
16	0.133	0.353	0.161	0.132	0.098	0.087
17	0.318	0.196	0.110	0.100	0.078	0.080

Table A1-2 *(continued)*

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
18	0.445	0.203	0.195	0.138	0.065	0.065
19	2.313	2.559	1.005	0.899	0.615	0.572
20	0.004	0.001	0.025	0.033	0.009	0.009
21	0.352	0.367	0.196	0.158	0.106	0.106
22	0.497	0.541	0.238	0.325	0.154	0.135
23	0.329	0.087	0.144	0.110	0.079	0.071
24	0.000	0.048	0.020	0.020	0.005	0.004
25	0.976	1.030	0.326	0.262	0.281	0.276
26	0.156	0.595	0.255	0.151	0.095	0.107
27	0.059	0.110	0.046	0.050	0.028	0.028
28	1.001	0.856	0.321	0.217	0.120	0.126
29	0.019	0.051	0.042	0.041	0.015	0.014
30	0.068	0.103	0.059	0.065	0.059	0.052
31	0.131	0.001	0.067	0.063	0.028	0.026
32	0.378	0.584	0.198	0.101	0.084	0.078
33	0.118	0.124	0.071	0.089	0.041	0.029
34	0.000	0.001	0.014	0.024	0.004	0.004
35	0.966	0.774	0.277	0.239	0.131	0.112
36	0.558	0.129	0.170	0.115	0.075	0.073
37	0.024	0.001	0.029	0.029	0.012	0.013
38	1.010	0.895	0.319	0.244	0.124	0.126
39	0.038	0.001	0.030	0.039	0.014	0.011
40	0.000	0.001	0.015	0.017	0.004	0.004
41	0.000	0.046	0.020	0.025	0.004	0.005
42	0.039	0.040	0.037	0.025	0.010	0.010
43	0.000	0.076	0.028	0.015	0.006	0.006
44	0.025	0.057	0.068	0.056	0.020	0.018
45	0.118	0.127	0.050	0.069	0.024	0.020
46	0.130	0.130	0.102	0.117	0.059	0.039
47	0.171	0.229	0.159	0.160	0.098	0.082
48	1.050	0.385	0.311	0.280	0.206	0.192
49	0.067	0.147	0.038	0.038	0.026	0.026
50	0.382	0.931	0.319	0.249	0.119	0.120
51	0.150	0.429	0.201	0.158	0.105	0.108
52	0.000	0.034	0.025	0.021	0.004	0.005
53	0.028	0.119	0.029	0.035	0.011	0.011
54	0.573	0.664	0.230	0.172	0.108	0.099
55	0.501	0.241	0.189	0.125	0.067	0.067
56	0.541	0.353	0.253	0.165	0.099	0.105
57	0.000	0.001	0.023	0.023	0.004	0.005
58	0.142	0.178	0.082	0.090	0.041	0.046
59	0.017	0.090	0.052	0.034	0.013	0.013
60	0.008	0.001	0.028	0.028	0.008	0.007
61	0.038	0.098	0.047	0.045	0.026	0.028
62	0.469	0.576	0.191	0.165	0.106	0.110
63	0.963	0.831	0.327	0.403	0.208	0.161
64	0.898	0.661	0.315	0.228	0.119	0.122
65	0.203	0.147	0.079	0.094	0.041	0.045
66	0.020	0.045	0.028	0.027	0.009	0.007
67	0.071	0.001	0.066	0.055	0.022	0.019
68	0.250	0.307	0.195	0.190	0.100	0.085
69	0.569	0.606	0.181	0.142	0.067	0.065
70	0.119	0.098	0.112	0.163	0.057	0.040
71	0.008	0.119	0.026	0.029	0.011	0.011
72	0.260	0.140	0.121	0.108	0.058	0.056
73	0.273	0.301	0.109	0.082	0.067	0.067
74	0.031	0.001	0.054	0.031	0.009	0.010
75	0.390	0.568	0.202	0.120	0.067	0.068
76	0.000	0.001	0.036	0.019	0.004	0.005
77	0.350	0.908	0.303	0.231	0.122	0.116

Table A1-3
MAPE values of each attraction (%).

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
1	0.310	0.952	0.303	0.244	0.160	0.152
2	0.154	0.040	0.464	0.795	0.226	0.239

(continued on next page)

Table A1-3 *(continued)*

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
3	0.866	0.498	0.352	0.307	0.155	0.166
4	0.897	0.833	0.262	0.19	0.311	0.204
5	0.138	0.031	0.488	0.75	0.181	0.235
6	0.869	0.471	0.222	0.21	0.263	0.240
7	0.512	0.461	0.408	0.492	0.203	0.219
8	0.138	1.224	0.568	0.459	0.469	0.318
9	0.372	0.899	0.232	0.164	0.676	0.224
10	0.462	0.019	0.617	0.565	0.231	0.195
11	0.214	0.542	0.234	0.287	0.411	0.199
12	0.920	0.652	0.367	0.357	0.189	0.195
13	0.881	0.199	0.250	0.286	0.223	0.190
14	0.541	0.531	1.128	1.078	0.452	0.236
15	0.903	0.667	0.320	0.125	0.147	0.136
16	0.232	0.579	0.285	0.249	0.393	0.144
17	0.930	0.452	0.265	0.249	0.182	0.189
18	0.630	0.251	0.258	0.169	0.400	0.182
19	0.740	0.867	0.321	0.309	0.431	0.082
20	0.092	0.019	0.366	0.477	0.486	0.209
21	0.530	0.477	0.233	0.178	0.162	0.141
22	0.810	0.915	0.407	0.488	0.227	0.217
23	0.757	0.175	0.283	0.201	0.150	0.162
24	0.186	1.494	0.544	0.616	0.266	0.216
25	0.867	0.901	0.283	0.246	0.206	0.236
26	0.148	0.740	0.266	0.154	0.245	0.213
27	0.478	0.817	0.333	0.358	0.247	0.259
28	0.964	0.807	0.262	0.181	0.377	0.244
29	0.176	0.466	0.367	0.352	0.346	0.240
30	0.206	0.455	0.234	0.233	0.137	0.199
31	0.850	0.007	0.333	0.349	0.194	0.217
32	0.593	0.899	0.233	0.121	0.391	0.194
33	0.620	0.900	0.413	0.539	0.234	0.207
34	0.138	0.031	0.323	0.604	0.281	0.262
35	0.963	0.765	0.214	0.134	0.223	0.200
36	0.936	0.175	0.237	0.15	0.311	0.104
37	0.438	0.024	0.512	0.55	0.210	0.168
38	0.973	0.859	0.263	0.198	0.273	0.214
39	0.543	0.016	0.383	0.471	0.184	0.155
40	0.138	0.031	0.348	0.385	0.340	0.225
41	0.138	1.105	0.480	0.556	0.243	0.073
42	0.492	0.489	0.463	0.345	0.200	0.112
43	0.205	1.748	0.561	0.351	0.286	0.297
44	0.155	0.352	0.352	0.366	0.112	0.118
45	0.728	0.783	0.278	0.36	0.163	0.144
46	0.315	0.436	0.288	0.289	0.160	0.161
47	0.288	0.289	0.222	0.2	0.111	0.137
48	0.963	0.313	0.260	0.241	0.519	0.142
49	0.468	1.071	0.245	0.242	0.194	0.137
50	0.337	0.888	0.269	0.207	0.362	0.212
51	0.158	0.528	0.234	0.182	0.472	0.277
52	0.138	0.786	0.493	0.542	0.411	0.231
53	0.396	1.753	0.398	0.489	0.399	0.240
54	0.746	0.863	0.229	0.122	0.329	0.154
55	0.786	0.326	0.251	0.153	0.098	0.090
56	0.673	0.398	0.262	0.17	0.122	0.106
57	0.173	0.031	0.493	0.569	0.381	0.302
58	0.478	0.689	0.246	0.301	0.191	0.156
59	0.118	0.592	0.286	0.224	0.101	0.079
60	0.189	0.026	0.564	0.58	0.146	0.127
61	0.204	0.588	0.272	0.235	0.258	0.240
62	0.650	0.834	0.230	0.177	0.430	0.182
63	0.990	0.820	0.284	0.361	0.173	0.174
64	0.862	0.624	0.263	0.183	0.099	0.089
65	0.924	0.601	0.268	0.341	0.177	0.148
66	0.643	1.440	0.873	0.855	0.300	0.184
67	0.957	0.013	0.580	0.438	0.219	0.166
68	0.496	1.438	0.642	0.621	0.267	0.206
69	0.897	0.955	0.238	0.173	0.646	0.253
70	0.697	0.454	0.480	0.701	0.246	0.197
71	0.179	1.321	0.352	0.386	0.399	0.210
72	0.816	0.431	0.299	0.267	0.157	0.159

Table A1-3 *(continued)*

Attractions	SARIMAX	ANN	LSTM	LSTM-AM	CTS-LSTM	CTS-LSTM-AM
73	0.854	0.855	0.296	0.22	0.144	0.104
74	0.516	0.019	0.613	0.439	0.327	0.174
75	0.624	0.894	0.268	0.142	0.395	0.201
76	0.475	0.040	0.789	0.682	0.288	0.099
77	0.311	0.874	0.248	0.181	0.454	0.207

Appendix 2. The Diebold–Mariano test results for each attraction

Table A2-1
Values of the Diebold–Mariano test based on MAE (h = 9).

Attractions	CTS-LSTM-AM vs SARIMAX	CTS-LSTM-AM vs ANN	CTS-LSTM-AM vs LSTM	CTS-LSTM-AM vs LSTM-AM	CTS-LSTM-AM vs CTS-LSTM	CTS-LSTM-AM vs baseline
1	−18.045***	−15.622***	−4.860***	−4.725***	−0.763	−3.185***
2	−29.078***	−13.055***	−6.870***	−6.372***	−0.571	−4.794***
3	−8.381***	−13.717***	−8.322***	−8.814***	−2.200**	−6.386***
4	−11.854***	−9.956***	−7.408***	−6.548***	−3.126***	−9.008***
5	−18.899***	−7.956***	−6.391***	−6.578***	−2.028**	−8.080***
6	−16.219***	−11.126***	−4.582***	−5.605***	−0.182	−5.780***
7	−9.375***	−8.672***	−8.424***	−6.625***	−4.888***	−7.516***
8	−31.683***	−11.802***	−7.461***	−8.960***	−21.009***	−15.206***
9	−21.706***	−7.702***	−8.596***	−5.120***	−3.233***	−4.684***
10	−15.876***	−8.360***	−6.949***	−5.326***	−4.799***	−10.823***
11	−7.516***	−16.665***	−6.694***	−7.005***	0.038	−5.236***
12	−36.811***	−6.754***	−8.224***	−6.770***	−2.372**	−5.784***
13	−16.784***	−36.523***	−10.109***	−8.463***	−21.997***	−10.140***
14	−24.741***	−19.066***	−3.769***	−5.332***	−0.994	−3.355***
15	−15.270***	−21.615***	−5.689***	−8.299***	−12.998***	−13.855***
16	−16.289***	−26.144***	−8.755***	−5.326***	−25.430***	−23.890***
17	−15.224***	−10.227***	−8.494***	−6.640***	−5.914***	−7.240***
18	−17.428***	−10.204***	−13.345***	−9.985***	−47.705***	−20.679***
19	−40.179***	−6.358***	−5.888***	−5.554***	−10.379***	−8.945***
20	−17.879***	−38.456***	−3.081***	−5.212***	−2.872***	−4.575***
21	−28.971***	−13.040***	−5.905***	−5.943***	−0.877	−5.442***
22	−19.704***	−36.348***	2.978***	−0.663	−0.121	−0.945
23	−21.192***	−7.146***	−8.656***	−7.916***	−5.179***	−5.115***
24	−21.276***	−21.047***	−2.771**	−0.020	−1.024	−19.493***
25	−9.671***	−10.919***	−3.944***	−4.979***	−2.899***	−8.617***
26	−5.855***	−12.555***	−8.799***	−6.901***	−2.778**	−7.548***
27	−17.255***	−8.958***	−10.222***	−7.904***	−4.237***	−8.043***
28	−25.998***	−14.938***	−7.461***	−3.888***	−3.687***	−9.628***
29	−17.123***	−17.070***	−0.967	−1.328	−2.342**	−3.979***
30	−23.377***	−17.472***	−4.895***	−4.814***	0.096	−21.532***
31	−17.168***	−8.223***	−7.588***	−9.073***	−7.359***	−7.639***
32	−12.627***	−9.966***	−5.158***	−5.276***	−0.547	−12.894***
33	−15.348***	−16.324***	−5.341***	−5.801***	−1.643	−7.780***
34	−17.123***	−13.541***	−6.182***	−5.596***	−0.697	−9.927***
35	−4.105***	−16.565***	−12.272***	−9.819***	−8.336***	−11.769***
36	−3.001***	−13.310***	−9.324***	−3.433***	−2.899***	−5.288***
37	−8.480***	−12.622***	−10.556***	−8.250***	−2.489**	−7.039***
38	−14.381***	−16.807***	−7.294***	−5.651***	−2.822**	−10.172***
39	−6.000***	−6.855***	−7.181***	−7.126***	−1.706*	−12.028***
40	−25.711***	−15.360***	−13.099***	−4.673***	−16.730***	−12.817***
41	−20.121***	−6.548***	−5.813***	−4.543***	−8.792***	−11.922***
42	−12.129***	−8.072***	−4.783***	−5.196***	−0.535	−10.055***
43	−0.245	−19.749***	−4.068***	−3.531***	−0.184	−5.696***
44	−22.042***	−3.398***	−4.421***	−4.117***	−2.102**	−9.216***
45	−32.324***	−21.336***	−5.486***	−5.321***	−0.253	−11.848***
46	−33.981***	−25.430***	−2.948***	−4.157***	−2.221**	−10.725***
47	−33.065***	−11.379***	−7.881***	−5.752***	−19.161***	−9.255***
48	−35.531***	−21.368***	−9.448***	−4.414***	−4.313***	−20.377***
49	−12.090***	−6.123***	−3.585***	−5.941***	−5.298***	−11.676***
50	−21.539***	−4.642***	−7.687***	−4.021***	−4.944***	−6.710***
51	−30.311***	−9.318***	−8.491***	−6.827***	−4.941***	−6.356***
52	−15.035***	−18.868***	−11.468***	−6.333***	−6.151***	−10.701***
53	−14.801***	−9.846***	−18.299***	−5.486***	−8.785***	−8.494***

(continued on next page)

Table A2-1 *(continued)*

Attractions	CTS-LSTM-AM vs SARIMAX	CTS-LSTM-AM vs ANN	CTS-LSTM-AM vs LSTM	CTS-LSTM-AM vs LSTM-AM	CTS-LSTM-AM vs CTS-LSTM	CTS-LSTM-AM vs baseline
54	−51.269***	−42.278***	−2.177**	−1.195	−1.390	−44.715***
55	−56.993***	−45.571***	−1.408	−2.977***	−1.280	−28.266***
56	−5.339***	−9.441***	−6.781***	−5.377***	−1.203	−10.134***
57	−17.850***	−24.780***	−5.447***	−4.460***	−3.441***	−9.591***
58	−74.852***	−26.202***	−11.307***	−6.494***	−3.815***	−68.800***
59	−17.306***	−11.082***	−7.757***	−4.979***	−2.069**	−4.436***
60	−6.477***	−17.043***	−4.528***	−6.331***	−1.266	−18.875***
61	−35.204***	−16.205***	−8.755***	−6.827***	−18.327***	−5.626***
62	−27.316***	−25.072***	−5.712***	−5.975***	−0.137	−4.568***
63	−53.492***	−52.730***	−2.896***	−3.834***	−1.074	−40.461***
64	−38.905***	−10.617***	−2.995***	−2.251***	−1.796*	−33.606***
65	−32.177***	−8.196***	−5.642***	−4.403***	−4.861***	−10.521***
66	−20.775***	−15.087***	−7.980***	−6.827***	−3.607***	−12.500***
67	−18.901***	−7.935***	−8.731***	−6.697***	−3.781***	−7.645***
68	−17.197***	−10.474***	−8.076***	−6.843***	−10.713***	−8.369***
69	−14.165***	−11.902***	−7.253***	−6.072***	−1.903*	−12.477***
70	−49.943***	−11.184***	−8.755***	−7.437***	−4.738***	−5.587***
71	−28.851***	−30.941***	−4.826***	−5.307***	−0.396	−6.292***
72	−43.707***	−41.498***	−2.770**	−3.302***	−4.134***	−5.011***
73	−12.625***	−16.122***	−5.566***	−7.814***	−7.285***	−13.394***
74	−19.917***	−8.568***	−8.445***	−5.482***	−6.752***	−8.734***
75	−59.162***	−15.477***	−8.727***	−8.031***	−13.465***	−11.353***
76	−17.376***	−20.676***	−8.755***	−6.234***	−13.311***	−12.987***
77	−9.375***	−8.672***	−8.424***	−6.625***	−4.888***	−6.356***yy

*** Significant at 1%.
** Significant at 5%.
* Significant at 10%.

References

Alvarez Diaz, M., & Mateu-Sbert, J. (2011). Forecasting daily air arrivals in Mallorca Island using nearest neighbour methods. *Tourism Economics*, *17*, 191–208.
Assaf, A. G., Li, G., Song, H., & Tsionas, M. G. (2019). Modeling and forecasting regional tourism demand using the Bayesian global vector autoregressive (BGVAR) model. *Journal of Travel Research*, *58*, 383–397.
Assaf, A. G., & Tsionas, M. G. (2019). Forecasting occupancy rate with Bayesian compression methods. *Annals of Tourism Research*, *75*, 439–449.
Balli, F., Curry, J., & Balli, H. O. (2015). Inter-regional spillover effects in New Zealand international tourism demand. *Tourism Geographies*, *17*, 262–278.
Bangwayo-Skeete, P., & Skeete, R. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, *46*, 454–464.
Bi, J. -W., Liu, Y., & Li, H. (2020). Daily tourism volume forecasting for tourist attractions. *Annals of Tourism Research*, *83*, Article 102923.
Cao, Z., Li, G., & Song, H. (2017). Modelling the interdependence of tourism demand: The global vector autoregressive approach. *Annals of Tourism Research*, *67*, 1–13.
Capone, F., & Domenech, R. (2008). Sources of growth and competitiveness of local tourist production systems: An application to Italy (1991-2001). *The Annals of Regional Science*, *42*, 209–224.
Chhetri, A., Arrowsmith, C., Chhetri, P., & Corcoran, J. (2013). Mapping spatial tourism and hospitality employment clusters: An application of spatial autocorrelation. *Tourism Analysis*, *18*, 559–573.
Cotos-Yáñez, T., Alén, E., Domínguez, T., & Losada, N. (2018). Spatial heterogeneity in Spain for senior travel behavior. *Tourism Management*, *70*, 444–452.
Diebold, F., & Mariano, R. (2002). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *20*, 134–144.
Divino, J. A., & McAleer, M. (2010). Modelling and forecasting daily international mass tourism to Peru. *Tourism Management*, *31*, 846–854.
Emili, S., Gardini, A., & Foscolo, E. (2020). High spatial and temporal detail in timely prediction of tourism demand. *International Journal of Tourism Research*, *22*, 451–463.
Fingleton, B., & López-Bazo, E. (2006). Empirical growth models with spatial effects. *Papers in Regional Science*, *85*, 177–198.
Gooroochurn, N., & Hanley, A. (2005). Spillover effects in long-haul visitors between two regions. *Regional Studies*, *39*, 727–738.
Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780.
Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2021). Improving tourist arrival prediction: A big data and artificial neural network approach. *Journal of Travel Research*, *60*(5), 998–1017.
Huang, S., Wang, D., Wu, X., & Tang, A. (2019). Dsanet: Dual self-attention network for multivariate time series forecasting. *28th ACM international conference* (pp. 2129–2132).
Jiao, E. X., & Chen, J. L. (2019). Tourism forecasting: A review of methodological developments over the last decade. *Tourism Economics*, *25*, 469–492.
Jiao, X., Li, G., & Chen, J. L. (2020). Forecasting international tourism demand: A local spatiotemporal model. *Annals of Tourism Research*, *83*, Article 102937.
Khadaroo, J., & Seetanah, B. (2008). The role of transport infrastructure in international tourism development: A gravity model approach. *Tourism Management*, *29*(5), 831–840.
Kulshrestha, A., Krishnaswamy, V., & Sharma, M. (2020). Bayesian BILSTM approach for tourism demand forecasting. *Annals of Tourism Research*, *83*, Article 102925.
Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management*, *21*, 331–340.
Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research*, *75*, 410–423.
Li, H., Chen, J. L., Li, G., & Goh, C. (2016). Tourism and regional income inequality: Evidence from China. *Annals of Tourism Research*, *58*, 81–99.
Li, H., Hu, M., & Li, G. (2020). Forecasting tourism demand with multisource big data. *Annals of Tourism Research*, *83*, Article 102912.
Li, X., & Law, R. (2020). Forecasting tourism demand with decomposed search cycles. *Journal of Travel Research*, *59*, 52–68.
Lin, V. S., Yang, Y., & Li, G. (2018). Where can tourism-led growth and economy-driven tourism growth occur? *Journal of Travel Research*, *58*, 760–773.
Long, W., Liu, C., & Song, H. Y. (2019). Pooling in tourism demand forecasting. *Journal of Travel Research*, *58*, 1161–1174.
Ma, T., Hong, T., & Zhang, H. (2015). Tourism spatial spillover effects and urban economic growth. *Journal of Business Research*, *68*, 74–80.
Pan, B., & Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *Journal of Travel Research*, *56*, 957–970.
Pereira, & Nobre, L. (2016). An introduction to helpful forecasting methods for hotel revenue management. *International Journal of Hospitality Management*, *58*, 13–23.

Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., ... Iyengar, S. S. (2019). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys*, *51*, 1–36.
Santos, G., Ramos, V., & Rey-Maquieira, J. (2011). A microeconomic model of multidestination tourism trips. *Tourism Economics*, *17*, 509–529.
Shih, S. -Y., Sun, F. -K., & Lee, H. -y. (2019). Temporal pattern attention for multivariate time series forecasting. *Machine Learning*, *108*, 1421–1441.
Song, H., Dwyer, L., Li, G., & Cao, Z. (2012). Tourism economics research: A review and assessment. *Annals of Tourism Research*, *39*, 1653–1682.
Song, H., & Li, G. (2008). Tourism demand modelling and forecasting-a review of recent research. *Tourism Management*, *29*, 203–220.
Song, H., Wen, L., & Liu, C. (2019). Density tourism demand forecasting revisited. *Annals of Tourism Research*, *75*, 379–392.
Song, H., & Wong, K. K. F. (2003). Tourism demand modeling: A time-varying parameter approach. *Journal of Travel Research*, *42*, 57–64.
Stewart, S. I., & Vogt, C. A. (1997). Multi-destination trip patterns. *Annals of Tourism Research*, *24*, 458–461.
Tsui, K., & Balli, F. (2017). International arrivals forecasting for Australian airports and the impact of tourism marketing expenditure. *Tourism Economics*, *23*(2), 403–428.
Wan, H., Guo, S., Yin, K., Liang, X., & Lin, Y. (2020). CTS-LSTM: LSTM-based neural networks for correlatedtime series prediction. *Knowledge-Based Systems*, *191*, Article 105239.
Weidenfeld, A., Butler, R., & Williams, A. (2010). Clustering and compatibility between tourism attractions. *International Journal of Tourism Research*, *12*(1), 1–16.
Wu, D., Song, H., & Shen, S. (2017). New developments in tourism and hotel demand modeling and forecasting. *International Journal of Contemporary Hospitality Management*, *29*, 507–529.
Xie, G., Qian, Y., & Wang, S. (2020). A decomposition-ensemble approach for tourism forecasting. *Annals of Tourism Research*, *81*, Article 102891.
Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C: Emerging Technologies*, *95*, 47–60.
Yang, Y., & Fik, T. (2014). Spatial effects in regional tourism growth. *Annals of Tourism Research*, *46*, 144–162.
Yang, Y., Fik, T., & Zhang, H. (2017). Designing a tourism spillover index based on multidestination travel: A twostage distance-based modeling approach. *Journal of Travel Research*, *56*, 317–333.
Yang, Y., Pan, B., & Song, H. (2014). Predicting hotel demand using destination marketing organization's web traffic data. *Journal of Travel Research*, *53*, 433–447.
Yang, Y., & Wong, K. K. F. (2012). A spatial econometric approach to model spillover effects in tourism flows. *Journal of Travel Research*, *51*, 768–778.
Yang, Y., & Zhang, H. (2019). Spatial-temporal forecasting of tourism demand. *Annals of Tourism Research*, *75*, 106–119.
Zhang, B., Li, N., Shi, F., & Law, R. (2020). A deep learning approach for daily tourist flow forecasting with consumer search data. *Asia Pacific Journal of Tourism Research*, *25*, 323–339.
Zhang, K., Chen, Y., & Li, C. L. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tourism Management*, *75*, 595–608.
Zhang, X. W., Yang, Y., Zhang, Y., & Zhang, Z. L. (2020). Designing tourist experiences amidst air pollution: A spatial analytical approach using social media. *Annals of Tourism Research*, *84*, Article 102999.
Zhang, Y., Li, G., Muskat, B., & Law, R. (2021). Tourism demand forecasting: A decomposed deep learning approach. *Journal of Travel Research*, *60*(5), 981–997.
Zhang, Y., Li, G., Muskat, B., Law, R., & Yang, Y. (2020). Group pooling for deep tourism demand forecasting. *Annals of Tourism Research*, *82*, Article 102899.
Zhang, Y., Xu, J. -H., & Zhuang, P. -J. (2011). The spatial relationship of tourist distribution in Chinese cities. *Tourism Geographies*, *13*, 75–90.

Weimin Zheng, PhD is an associate professor of tourism management at Xiamen University. His research interests focus on tourism demand forecast and smart tourism.
Liyao Huang, is a PhD student of tourism management at Xiamen University. Her research interests focus on tourism demand forecast and tourist mobility.
Zhibin Lin, PhD is an associate professor in marketing at Durham University Business School, UK. Research interests include transport, travel and tourism management.