Expert Systems With Applications 236 (2024) 121388

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Forecasting tourism demand with a novel robust decomposition and ensemble framework

Xin Li [a,*], Xu Zhang [a], Chengyuan Zhang [b], Shouyang Wang [c,d,e]

[a] School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China
[b] School of Economics and Management, Xidian University, Xi'an 710071, China
[c] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[d] School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China
[e] School of Entrepreneurship and Management, ShanghaiTech University, Shanghai 201210, China

ARTICLE INFO

Keywords:
Tourism demand forecasting
Decomposition
RobustSTL
Ensemble framework

ABSTRACT

Current research highlights the efficacy of decomposition and ensemble algorithms in enhancing forecasting accuracy; however, the investigation of robustness associated with decomposed components within these algorithms remains notably scarce in existing literature. To address this gap, we introduce a novel tourism demand forecasting framework, underpinned by a sophisticated decomposition algorithm. Our approach initially decomposes the original data into multiple sub-series and subsequently selects forecasting models based on their respective data attributes. We evaluated the proposed framework by forecasting monthly tourist arrivals in Hong Kong from six countries. The superiority of the novel decomposition method was further substantiated through comparisons with alternative decomposition techniques in both single-step and multi-step ahead forecasting contexts. The results indicate that the proposed forecasting framework consistently outperforms baseline models in terms of forecasting accuracy across all tourism demand forecasting scenarios, with the innovative decomposition method exhibiting exceptional performance.

1. Introduction

Tourism demand forecasting is crucial for the tourism and hospitality sectors, providing vital insights into emerging trends and evolving tourist preferences. This enables businesses to establish evidence-based planning and marketing strategies (Song & Li, 2008; Wu et al., 2023). The importance of forecasting accuracy in these industries cannot be overstated, yet it remains a daunting task (Silva et al., 2019; Song et al., 2019; Zhang et al., 2020). The complex and ever-changing nature of tourism demand data makes achieving reliable forecasts challenging. Generally, time series data in tourism display non-linear, non-stationary characteristics, often accompanied by seasonality and unpredictable changes (Zhang et al., 2022). Furthermore, unprecedented events such as pandemics have a profound effect on the demand for tourism, while simultaneously exposing both domestic and foreign tourism industries to substantial adverse shocks (Škare et al., 2021). To tackle the challenges posed by these data features, researchers and professionals in the tourism field have explored and developed various forecasting techniques, including time series methods (Ma et al., 2016), econometric approaches (Pan & Yang, 2017), artificial intelligence (AI) methodologies (Law et al., 2019), and deep learning strategies (Zheng et al., 2021). Although these methods have improved forecasting accuracy to a certain degree, numerous challenges and complexities remain.

The decomposition algorithm has been demonstrated to have significant advantages in tourism forecasting research, as it can dissect the original data into simpler components, thereby reducing modeling complexity and enhancing forecasting ability (Li & Law, 2020; Xie et al., 2020). By analyzing the causes of fluctuations and anomalies in each component, each sub-sequence can be predicted separately, and the results can be integrated to obtain more accurate forecasting results (Zhang et al., 2021). This strategy can overcome the non-linear and non-stationary characteristics of tourism demand data and can also cope with the impact of exceptional events. For instance, Zhang et al. (2020) applied the decomposition method to decompose the tourist arrival time series into global trend mode, fixed seasonal component, and local sensitivity mode. By decomposing the tourism arrival data into several straightforward sub-sequences, the forecasting accuracy was enhanced. Consequently, it is extensively utilized in the domain of tourism demand

* Corresponding author.
E-mail addresses: drxinli@ustb.edu.cn (X. Li), chyzhang@xidian.edu.cn (C. Zhang), sywang@amss.ac.cn (S. Wang).

https://doi.org/10.1016/j.eswa.2023.121388
Received 6 June 2023; Received in revised form 28 August 2023; Accepted 29 August 2023
Available online 1 September 2023
0957-4174/

forecasting (Li & Law, 2020; Zhang et al., 2022).

Within the existing body of research, numerous decomposition algorithms demonstrate potential for enhancing forecasting accuracy; however, two predominant limitations persist. Firstly, the performance of these methods is impeded by insufficient robustness in addressing seasonal fluctuations, shifts, and unexpected changes in trends and residuals. Secondly, the interpretability of decomposed components remains restricted. For instance, the widely applied ensemble empirical mode decomposition (EEMD) algorithm dissects the original data into intrinsic mode functions (IMFs), which, regrettably, do not embody inherent components of tourism time series data, resulting in a reduced explanatory capability (Li & Law, 2020).

To enhance the precision of tourism demand forecasting, some researchers have employed ensemble models that combine the strengths of multiple forecasting techniques (Zheng et al., 2021; Li et al., 2022). Li and Law (2020) utilized the ensemble empirical mode decomposition method to decompose the tourist arrival data from nine countries visiting Hong Kong. Subsequently, an autoregressive model was used for forecasting, resulting in a significant improvement in forecasting performance. Feng et al (2022) employed the singular spectrum analysis technique to decompose the tourist data from Mount Siguniang. Afterwards, the decomposed sub-sequences were predicted using the support vector regression model, resulting in the desired outcome. Although the aforementioned studies have been able to enhance prediction accuracy, they utilize a uniform model to forecast the decomposed components without considering the distinct characteristics of these components, such as linearity or non-linearity. Consequently, this constraint limits the further improvement of tourism demand forecasting accuracy.

This paper introduces an innovative decomposition and ensemble forecasting framework grounded in a robust decomposition algorithm. This methodology initially applies a robust decomposition algorithm to partition the original data into interpretable constituents, such as trend and seasonality. Following this, linear or non-linear models are chosen for modeling, contingent upon the attributes of the decomposed components. For instance, linear components are modeled utilizing a classical autoregressive integrated moving average (ARIMA) model, whereas non-linear components require the application of a long short-term memory network (LSTM) model. We endeavor to examine whether the implementation of a robust decomposition method and the utilization of hybrid forecasting models can contribute to an improvement in forecasting accuracy in both one-step and multi-step ahead forecasting scenarios. The efficacy of our proposed technique was evaluated using six monthly tourist arrival datasets pertaining to visitors to Hong Kong. Across the evaluated datasets, our methodology substantially improved forecasting accuracy in terms of root mean square error (RMSE) by an estimated 72.01% in both single-step and multi-step forecasting contexts.

The structure of the paper is as follows: Section 2 provides an extensive review of the pertinent literature, Section 3 describes the methods and forecasting framework adopted, Section 4 offers a detailed presentation of the empirical study, and the concluding remarks are presented in Section 5.

2. Literature review

This section initially offers a literature review of popular forecasting models employed in the field of tourism demand forecasting. Subsequently, we examine the decomposition techniques and ensemble approaches utilized in tourism forecasting, respectively.

2.1. An overview of tourism forecasting models

Three primary categories of models are employed in tourism demand forecasting: time series, econometric, and AI-based models (Song et al., 2019). Time series analysis has been extensively utilized for forecasting tourism demand based on historical tourism data in the literature. ARIMA and its variants are among the most prevalent models in the field of tourism demand forecasting, and in many instances, they yield accurate forecasting, establishing them as the most widely used and efficacious models (Li et al., 2021). For example, Baldigara and Mamula (2015) applied ARIMA to forecast German tourism demand in Croatia. Ma et al. (2016) used Seasonal ARIMA to forecast the demand for Chinese tourists arriving in Australia. Naive and exponential smoothing techniques and Prophet are also frequently employed for tourism demand forecasting (Lim & McAleer, 2001; Liu et al., 2022). However, the majority of ARIMA models assume a linear relationship between future and past time-step values, thus demonstrating strong linear fitting capabilities. Nonetheless, when the time series exhibits pronounced nonlinear characteristics, the resulting predictions may be less accurate.

Econometric models have also been broadly employed in tourism demand forecasting. These models consider various factors influencing tourism demand, such as income, weather, and prices, and construct models to predict future tourism demand. Prominent econometric models include autoregressive distributed lag model (ADLM), error correction model, vector autoregression (VAR), mixed data sampling (MIDAS) (Apergis et al., 2017; Liu et al., 2018; Song et al., 2019; Seabra et al., 2020; Hu et al., 2022; Wu et al., 2023). Sophisticated variants of VAR, including Bayesian VAR (BVAR), Global VAR (GVAR), and structural VAR model (Assaf et al., 2019; Hailemariam & Ivanovski, 2021), as well as spatial panel model (Yang & Zhang, 2019), have been explored. Additionally, time series bagging has been employed in tourism demand forecasting (Liu et al., 2023).

In recent years, AI-based approaches, including machine learning and deep learning methods, have gained increasing popularity in tourism demand forecasting and time series analysis, as they do not necessitate assumptions such as stationarity or distribution. Some commonly employed AI-based models encompass the back propagation neural network model (BPNN) (Kon and Turner, 2005), support vector regression model (SVR) (Chen et al., 2015), and extreme learning machine (ELM) and kernel extreme learning machine (KELM) (Sun et al., 2019), random vector functional link network (Zhang et al., 2022), and kernel random vector functional link network (Sun et al., 2022). Lately, deep learning models such as convolutional neural network (CNN) and LSTM have demonstrated remarkable performance in various visual and data processing tasks, attesting to the robust nonlinear fitting capabilities of these models (He et al., 2021). Among them, the LSTM is a notable example and has been extensively applied in time series and tourism demand forecasting (Yan et al., 2021; Li et al., 2022; Mao et al., 2023).

2.2. Decomposition methods in tourism forecasting

Decomposition methods have proven to be highly effective in reducing the complexity of time series data by partitioning them into distinct sub-series, allowing for the extraction of data features and, consequently, enhancing the precision of tourism forecasting (Li & Law, 2020). Prominent decomposition techniques encompass filtering, Fourier analysis, wavelet transforms, singular spectrum analysis (SSA), variational mode decomposition (VMD), empirical mode decomposition (EMD), and EEMD. A multitude of studies have corroborated that the implementation of decomposition methods considerably enhances the accuracy of time series forecasting (Li et al., 2018; Lin et al., 2018; Wu & Wu, 2019; Xie et al., 2020).

For example, Hassani et al. (2017) employed SSA to eliminate noise from the monthly international tourist arrivals in various European nations, thereby successfully ameliorating forecasting performance. Similarly, Li and Law (2020) utilized EEMD to decompose Google Trends search data into distinct components, effectively forecasting tourist arrivals in Hong Kong. Furthermore, Zhang et al. (2022) applied VMD to model nonlinear and nonstationary visitor arrival data, accurately forecasting the inbound tourism demand for the eight primary source markets of Hong Kong. These methods have improved the forecasting accuracy to some extent. However, these decomposition techniques may necessitate intricate inference procedures or involve ambiguous patterns. Distinct from these approaches, the seasonal trend decomposition method can partition the time series into trend, seasonal, and random noise components, enhance forecasting accuracy, and obviate the need for complex inference processes.

Seasonal trend decomposition is an effective seasonal adjustment technique extensively employed in time series analysis, which can enhance forecasting accuracy (Cleveland et al., 1990). The most classic and widely adopted decomposition method is Seasonal-Trend decomposition using Loess (STL). Zhang et al. (2021) applied STL to decompose inbound passenger flow from six source markets in Hong Kong and leveraged deep learning methodologies to augment the accuracy of short- and long-term tourism demand forecasting. Nevertheless, when trend changes and outliers are present in the time series, STL may not effectively extract seasonal components.

To address this issue, novel decomposition algorithms have been proposed, such as Seasonal-Trend decomposition based on regression (STR), which can concurrently extract trend, seasonal, and residual components to accommodate multiple seasons and seasonal changes without iteration (Cleveland et al., 1990; Dokumentov & Hyndman, 2015). Recently, an innovative seasonal trend decomposition method known as RobustSTL has been introduced. This method exhibits considerable robustness to noise and outliers and can yield superior decomposition outcomes (Wen et al., 2019). Table 1 highlights the widely used decomposition algorithms found in the literature for tourism demand forecasting, and provides a comparison between our study and relevant existing research.

2.3. Ensemble methods

Ensemble methods are widely used in various forecasting applications including tourism demand forecasting to improve the overall predictive performance (Timmermann, 2006). Ensemble methods often reduce errors and improve overall performance by leveraging the strengths of each individual model and minimizing the impact of their limitations (Rokach, 2010). By combining the decomposition and ensemble methods, the forecasting accuracy can be improved to some extent (Zhang et al., 2022).

While numerous studies have focused on decomposition and ensemble in tourism demand forecasting, the majority of them merely decompose the time series and subsequently combine it with a single forecasting model. For instance, Zhang et al. (2021) integrated STL decomposition with dual-attention deep learning models. Zhang et al. (2022) combined VMD with an extended version of a single hidden layer feedforward neural network (RVFL). Although these approaches have contributed to enhancing the short-term and long-term forecasting accuracy of time series, they have not taken into account the distinct data characteristics of the decomposed sub-sequences.

Only a handful of studies in the literature have taken into account the data characteristics of different sub-sequences. For instance, Zhang et al. (2021) proposed a novel decomposition algorithm, the noise-assisted multivariate empirical mode decomposition (NA-MEMD) method. This method consistently decomposes the time series into an IMFs and employs various forecasting techniques (e.g., linear regression, Seasonal ARIMA, BPNN, ELM, RVFL, and LSTM) to individually predict each IMF decomposition before aggregating the individual forecasting results to obtain the final outcome. This approach considers the distinct data characteristics of the decomposed sub-sequences, significantly improving forecasting accuracy. Recently, Bi et al. (2023) introduced an integrated deep learning model for decomposing complex tourism demand data into multiple components with simpler features using the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) method. They achieved more accurate forecasting of real-time tourism volume data for two attractions in Beijing (Ming Tombs and Temple of Heaven) by integrating CNN, LSTM, and AR models. Both of these studies have made substantial contributions toward enhancing the accuracy of tourism demand forecasting.

2.4. Research gap

In summary, decomposition methods can enhance the accuracy of tourism demand forecasting to a certain degree. However, time series decomposition still faces several unresolved challenges, including the inability to effectively handle seasonal changes, insufficient treatment of abrupt changes in trends and random components, and incompatibility

Table 1
Comparative analysis of characteristics and representative studies of common decomposition algorithms.

ID	Decomposition methods	Decomposed components	Forecasting methods	Outlier Robustness	Seasonality Shift	Long Period	Abrupt Trend Change	Time complexity	Representative papers
1	EMD	IMFs	GMDH; BP neural network;	√		√		$O(L^3)$	(Chen et al., 2012; Lin et al., 2018; Shabri, 2015)
2	EEMD	IMFs	ARX; Elman neural network; ARIMA	√	√	√	√	$O(NL^3)$	(Li & Law, 2020; Xie et al., 2020;Zhang et al., 2018)
3	Fourier	Multiple sine and cosine components	SARIMA; GreyModel		√			$O(LlogL)$	(Apergis et al., 2017; Hu, 2021)
4	SSA	multiple series	SSA-R; SVR					$O(L^2)$	(Hassani et al., 2017; Feng et al., 2022)
5	VMD	IMFs	RVFL; LSSVR	√	√	√	√	$O(KLlogL)$	(Zhang et al., 2022; Xing et al., 2022)
6	Wavelet	low-frequency components and high-frequency components	NARX; WPD	√				$O(LlogL)$	(Kummong & Supratid, 2016Wu & Wu, 2019)
7	HP	trend series and cycle series	Spectral analysis; STSM		√			$O(L)$	(Kožić,2014; Yao & Cao, 2020)
8	STL	Trend, seasonal, and residual	DADLM; LSTM			√	√	$O(KLW^2)$	(Y. Zhang et al., 2020; Adil et al., 2021)
9	RobustSTL	Trend, seasonal, and residual	ARIMA;LSTM	√	√	√	√	$O(KLW^2)$	This study

Note: EMD = empirical mode decomposition; EEMD = ensemble empirical mode decomposition; HP = Hodrick-Prescott filter; SSA = singular spectrum analysis; VMD = variational mode decomposition; STSM = structural time series model; STL = Seasonal and Trend decomposition using Loess; SVR = support vector regression; GMDH = Group Method Data Handling NARX = Nonlinear Auto-Regressive model with Exogenous Inputs; WPD = wavelet phase difference; DADLM = duo attention deep learning model; RVFL = random vector functional link network. And K represents the number of decomposition layers, L represents the length of the time series, and W represents the size of the local regression window. It should be noted that the actual time complexity in practical applications may vary due to specific implementation methods and parameter settings.

with data exhibiting long seasonal cycles. To address these issues, Wen et al. (2019) proposed the RobustSTL decomposition algorithm, which can manage sudden changes in trends and residuals to some extent while exhibiting robustness against noise and outliers in time series. Nevertheless, there are very few studies on the application of this algorithm in tourism demand forecasting.

Furthermore, although numerous studies have integrated decomposition algorithms with forecasting models, the majority have only combined a single model for ensemble forecasting, and only a few have focused on the data characteristics of the sub-series following time series decomposition. Consequently, to forecast tourism demand with greater accuracy, researchers need to amalgamate multiple forecasting models and decomposition methods, selecting the appropriate forecasting model based on the data characteristics of the sub-series.

It is crucial to acknowledge that every forecasting technique has its strengths and weaknesses, and no universal solution can consistently outperform others across all scenarios (Li et al., 2017; Song et al., 2019). As such, in tourism demand forecasting, the most suitable forecasting technique should be selected according to the specific situation. In this pursuit, we will employ three decomposition methods (STL, STR, and RobustSTL) and two forecasting models (ARIMA and LSTM) to examine whether selecting the appropriate model based on the data characteristics of the sub-series after decomposition can yield more accurate forecasts than direct forecasting following decomposition, and whether the RobustSTL decomposition algorithm is more robust and can offer more precise forecasts than other decomposition algorithms.

3. Methodology

In this section, we initially outline the three employed decomposition methods and two forecasting models. Following this, we introduce our proposed forecasting framework, which incorporates a robust decomposition method and hybrid forecasting strategy. Our proposed framework exhibits innovative aspects that are worth highlighting. Firstly, we incorporate robust decomposition techniques that have proven effective in partitioning time series data into well-defined and interpretable components. This approach enhances the interpretability of the results, surpassing existing methods like the EEMD approach employed by Li and Law (2020). Secondly, our framework allows for the adaptable application of forecasting methods that are tailored to the unique characteristics of the decomposed components. Linear components are forecasted using ARIMA, while nonlinear components are addressed using LSTM. Subsequently, the disparate segments are combined to generate a comprehensive and accurate final forecasting outcome.

3.1. STR decomposition

Decomposition techniques are widely utilized to identify and isolate the major components of time series, such as trends, seasonality, and cyclic patterns. By breaking down the data into its constituent components and forecasting each component separately, one can essentially filter out smaller, noise-like fluctuations and isolate the effects of stronger, more persistent elements. These decomposed components also tend to exhibit stronger determinism and thus are more easily inferable. As a result, forecasting individual components may potentially lead to more accurate forecasting than forecasting the series as a whole.

STR is a versatile method for handling multiple seasonal components and multiple linear covariates with various constant, flexible, and seasonal effects. It assumes the continuity of both trend and seasonal signals and enables fractional and flexible changes in the seasonal pattern over time. With this algorithm, it is possible to model complex seasonal effects and predictors that affect the data in complicated seasonal ways. We provide a comprehensive mathematical formulation of the STR below.

The observed time series Y_t is decomposed into several components, including the trend T_t, seasonal S_t, covariate effects $X_{t,j}$, and a residual component R_t:

$$Y_t = T_t + S_t + \sum_{j=1}^{J} \beta_j X_{t,j} + R_t, \tag{1}$$

where, T_t is estimated to capture the overall direction of the time series and can be modeled as a linear or nonlinear function of time:

$$T_t = f\left(T_{t-1}, \cdots, T_{t-p}\right) + \epsilon_t, \tag{2}$$

where f is a function that captures the trend, p is the order of autoregressive terms, and ϵ_t is the noise term.

The seasonal component S_t captures the repetitive patterns that occur at fixed intervals over time. STR can handle multiple seasonal patterns with different frequencies.

$$S_t = \sum_{k=1}^{K} A_k \cdot \cos(2\pi \cdot f_k \cdot t) + B_k \cdot \sin(2\pi \cdot f_k \cdot t), \tag{3}$$

where K is the number of seasonal patterns, A_k and B_k are coefficients for the sine and cosine terms, and f_k is the frequency of the k-th seasonal pattern.

Linear covariates $X_{t,j}$ can be included to account for external factors that influence the time series. The parameters β_j represent the effect of each covariate. The residual component R_t captures unexplained variability in the time series after accounting for trend, seasonal, and covariate effects.

$$R_t = Y_t - T_t - S_t - \sum_{j=1}^{J} \beta_j X_{t,j}. \tag{4}$$

The overall implementation process of STR is illustrated in Fig. 1.

The STR method is adept at accounting for the multiple and complex seasonal patterns in the input data. It can manage missing values, furnish confidence intervals, determine smoothing parameters, and accommodate the inclusion of time-varying and seasonal coefficients within regression models.

3.2. STL decomposition

STL is a general and robust method for decomposing time series, which decomposes a time series into three main components: trend, seasonal, and residual, and then model and predict each part separately. STL estimates trend and seasonality iteratively and uses locally estimated scatterplot smoothing to extract smooth estimates of the three components, making it suitable for capturing the inherent seasonal fluctuations and gradual trends present in tourism demand. The algorithm has been widely applied in many fields, but its ability to handle trend shifts and flexibility is limited in cases where the data has long periods and high noise. The STL algorithm operates through an iterative process that consists of two main parts: an inner loop for trend fitting and seasonal component calculation, and an outer loop for robustness weight adjustment. The mathematical expression is detailed below.

Consider a time series model with both trend and seasonality:

$$Y_t = T_t + S_t + R_t, t = 1, 2, \cdots N, \tag{5}$$

where Y_t is the observed value at time t, T_t is the trend of the time series, S_t is the seasonal signal with period t, and R_t is the residual signal. $T_t^{(k)}$ and $S_t^{(k)}$ represent the trend component and seasonal component, respectively, at the end of the k-1th iteration in the inner loop. Initially, $T_t^{(k)}$ is set to 0. The overall implementation process of STL is illustrated in Fig. 2.

It should be noted that the STL algorithm is iterative, and each iteration will decompose and stationary the residual sequence until the stopping criterion is met. In addition, in practical applications, some

Algorithm 1 STR Decomposition

procedure STRDECOMPOSE(Y, m, h)
  $N \leftarrow$ length of Y
  $t \leftarrow 1 : N$
  $H \leftarrow$ Fourier matrix of order m for t
  $X \leftarrow [\mathbf{1} \quad t]$
  $\beta \leftarrow$ coefficients of $X^T Y$ using QR decomposition
  $T \leftarrow X\beta$ ▷ Trend component
  $D \leftarrow Y - T$ ▷ Deseasonalized data
  $W \leftarrow$ diagonal matrix of weights $\omega_i = \begin{cases} 1 & \text{if } D_i \geq 0 \\ c & \text{if } D_i < 0 \end{cases}$ ▷ Weights for robustness
  $S \leftarrow H(WD)$ ▷ Seasonal component
  $R \leftarrow D - S$ ▷ Residual component
  return (T, S, R)
end procedure

Fig. 1. Description about the STR algorithm.

Algorithm 1 STL Decomposition

procedure STLDECOMPOSE$(Y, n(i), n(o), n(p), n(s), n(l), n(z))$
  $N \leftarrow$ length of Y
  $T_0 \leftarrow 0$ ▷ Initial trend component
  for $j \leftarrow 1$ to $n(o)$ **do** ▷ Outer loop
    $R_0 \leftarrow Y - T_{j-1}$ ▷ Initial remainder component
    $h \leftarrow 6 \times median(|R_0|)$ ▷ Robustness weight
    for $k \leftarrow 1$ to $n(i)$ **do** ▷ Inner loop
      $S_k \leftarrow$ CYCLESUBSERIESSMOOTH$(R_{k-1}, n(p), n(s))$
      $L_k \leftarrow$ LOWPASSFILTER$(S_k, n(p), n(l))$
      $T_k \leftarrow$ TRENDSMOOTH$(Y - S_k, n(z))$
      $R_k \leftarrow Y - T_k - S_k$ ▷ Update remainder component
      $\rho \leftarrow$ BISQUARE$(|R_k|/h)$ ▷ Robustness weight for each data point
      $w \leftarrow \rho \times$ NEIGHBORHOODWEIGHT$(\cdot)$ ▷ Weighted neighborhood weight
    end for
    $T_j \leftarrow T_k$
  end for
  return (T, S, R)
end procedure

Fig. 2. Description about the STL algorithm.

parameter adjustments and model evaluations are needed to obtain optimal forecasting results.

3.3. RobustSTL algorithm

RobustSTL, proposed by Wen et al. (2019), is a robust and versatile method for seasonal-trend decomposition. It can accurately extract seasonal features from data with long seasonal cycles and high noise levels, which makes it superior to existing algorithms. Specifically, it allows flexible and moving seasonal components that vary with time and can handle sudden changes in trend and remainder. RobustSTL is an algorithm designed to decompose the observed time series Y_t into three distinct components: trend, seasonal, and residual. Mathematically, this relationship is expressed as:

$$Y_t = T_t + S_t + R_t, t = 1, 2, \cdots N, \tag{6}$$

where Y_t is the observed value at time t, T_t is the trend of the time series, S_t is the seasonal signal with period t, and R_t is the residual signal.

The trend component T_t is extracted by solving a regression problem with the least absolute deviation loss and sparse regularization. It uses the least absolute deviation loss, coupled with sparse regularization, to identify the trend. After isolating the trend, the algorithm employs non-local seasonal filtering to extract the seasonal component S_t. This step recognizes repeating patterns across various time intervals, enabling RobustSTL to handle both moving and flexible seasonal patterns that may vary over time. The extraction process for both trend and seasonal components is performed iteratively. This iterative refinement ensures the accurate capture of these crucial elements of the time series.

Finally, the residual component R_t captures unexplained variability in the time series after accounting for trend and seasonal components.

In summary, RobustSTL is an efficient time series analysis method that can accurately extract seasonal features from data with long seasonal cycles and high noise levels, enabling more accurate forecasting. The mechanisms, such as outlier detection and handling, and weighted estimation, make RobustSTL better equipped to handle the complexities and uncertainties in tourism demand time series data. This robustness ultimately contributes to enhanced forecasting accuracy and a better understanding of underlying patterns in tourism demand data. The steps of the RobustSTL decomposition algorithm can be found in Fig. 3.

However, the algorithm still has some issues that need to be further studied and optimized, such as high computational complexity and difficulty in selecting appropriate parameters.

3.4. ARIMA model

ARIMA model is a widely used method for time series forecasting analysis. It consists of the autoregressive (AR) term with order p, the moving average (MA) term with order q, and the order of differencing d required to achieve stationarity of the time series. ARIMA is highly accurate for short-term forecasting and has been applied in various fields worldwide. However, the model cannot capture nonlinear data structures effectively, this is the reason that constrains further improvement

Algorithm 1 RobustSTL

procedure ROBUSTSTL(Y, n, p, q, λ)
  $S_0 \leftarrow$ median filter of length n applied to Y
  $Y_1 \leftarrow Y - S_0$ ▷ Detrend the data
  $i \leftarrow 0$
  while $i < p$ **do**
    $D_i \leftarrow q$-step Huber-weighted STL decomposition of Y_1
    $Y_2 \leftarrow Y_1 - D_i$ ▷ Deseasonalize the data
    $S_i \leftarrow$ median filter of length n applied to Y_2
    $Y_1 \leftarrow Y_2 - S_i$ ▷ Detrend the deseasonalized data
    $i \leftarrow i + 1$
  end while
  $D \leftarrow q$-step Huber-weighted STL decomposition of Y_1
  $W \leftarrow$ diagonal matrix of weights $\omega_i = \begin{cases} 1 & \text{if } |D| \leq \lambda \\ \lambda/|D| & \text{otherwise} \end{cases}$ ▷ Weights for robustness
  $S \leftarrow q$-step Huber-weighted STL decomposition of $Y_1 + WS_0$ ▷ Robust seasonal component
  $T \leftarrow S_0 + D + S$ ▷ Robust trend component
  $R \leftarrow Y - T$ ▷ Robust remainder component
  return (T, S, R)
end procedure

Fig. 3. Description about the RobustSTL algorithm.

of forecasting accuracy. The mathematical representation of the ARIMA model is as follows:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}, \tag{7}$$

where y_t represents the current value, μ represents the constant term, p and q are the orders, γ_i and θ_i are the autoregressive and moving average coefficients, respectively, and ϵ_t is the error term.

3.5. LSTM model

LSTM is a recurrent neural network (RNN) model that outperforms traditional RNNs in modeling long-term dependencies and processing long sequence data. LSTM uses three gate units, namely the input gate (i_t), forget gate (f_t), and output gate (o_t), to control the flow and retention of information, effectively addressing the gradient disappearance and explosion problems that occur in traditional RNN models. Furthermore, LSTM introduces memory cells that can efficiently store and extract historical information. The architecture of LSTM is depicted in Fig. 4.

The three gates in LSTM are determined by applying the sigmoid activation function to the weighted sum of current information (x_t) and previous state information (h_{t-1}). As shown in Fig. 4, the propagation of information in the storage unit mainly involves four steps. First, f_t determines which information from the previous cell state C should be discarded by h_{t-1}. If the value of f_t is close to 0, more information will be discarded. Second, i_t is used to determine how much new information needs to be retained in the current state c_t. To achieve this, a new candidate value is also added to the called state. Third, the current cell state c_t is obtained based on c_{t-1} and finally, o_t is used to control the output information. Therefore, the output h_t can be calculated. The LSTM model has strong non-linear fitting capabilities. The calculation formula is as follows:

$$\begin{cases} h_t = o_t * \tanh(c_t) \\ o_t = \sigma(w_o * [h_{t-1}, \alpha_t] + b_o) \\ c_t = f_t * c_{t-1} + i_t * \overline{c_t} \\ \overline{c_t} = \tanh(w_c * [h_{t-1}, x_t] + b_c) \\ i_t = \sigma(w_i * [h_{t-1}, x_t] + b_i) \\ f_t = \sigma(w_f * [h_{t-1}, x_t] + b_f) \end{cases}, t = 0, 1, 2, \cdots T, \tag{8}$$

where i_t, f_t, and o_t respectively denote the output of the input gate, forget gate, and output gate; c_t is the cell state;w_i, w_f, w_c, and w_o are the weights matrices; and b_o, b_c, b_i, and b_f are the biases vectors.

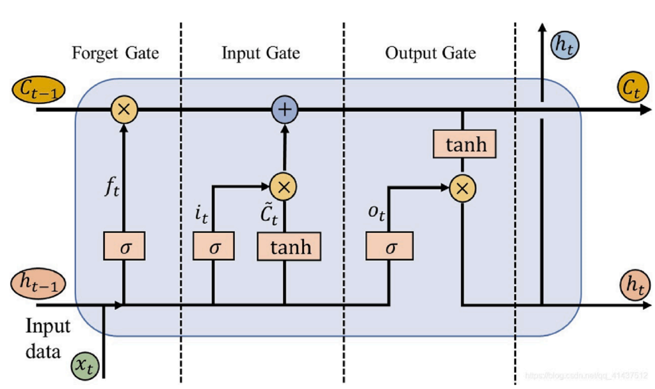

Fig. 4. LSTM cell structure diagram.

3.6. Proposed framework

Consequently, we propose a novel forecasting framework aimed at enhancing forecasting accuracy by employing a robust and reliable decomposition method, followed by constructing forecasting models in accordance with the data characteristics. Fig. 5 outlines the framework of the proposed methodology, encompassing five key steps: data collection and preprocessing, data decomposition, modeling, ensemble method, and forecasting evaluation. The specific details and underlying rationale for each step are outlined as follows.

Step 1: This step involves the collection and pre-processing of time series data on tourist arrivals. The goal of preprocessing is to enhance the data's quality, usability, and effectiveness for subsequent analysis and modeling. In this context, we employ a logarithmic transformation of the collected time series data. This transformation is primarily used to mitigate the impact of outliers, stabilizing the variance and rendering the data more suitable for analysis.

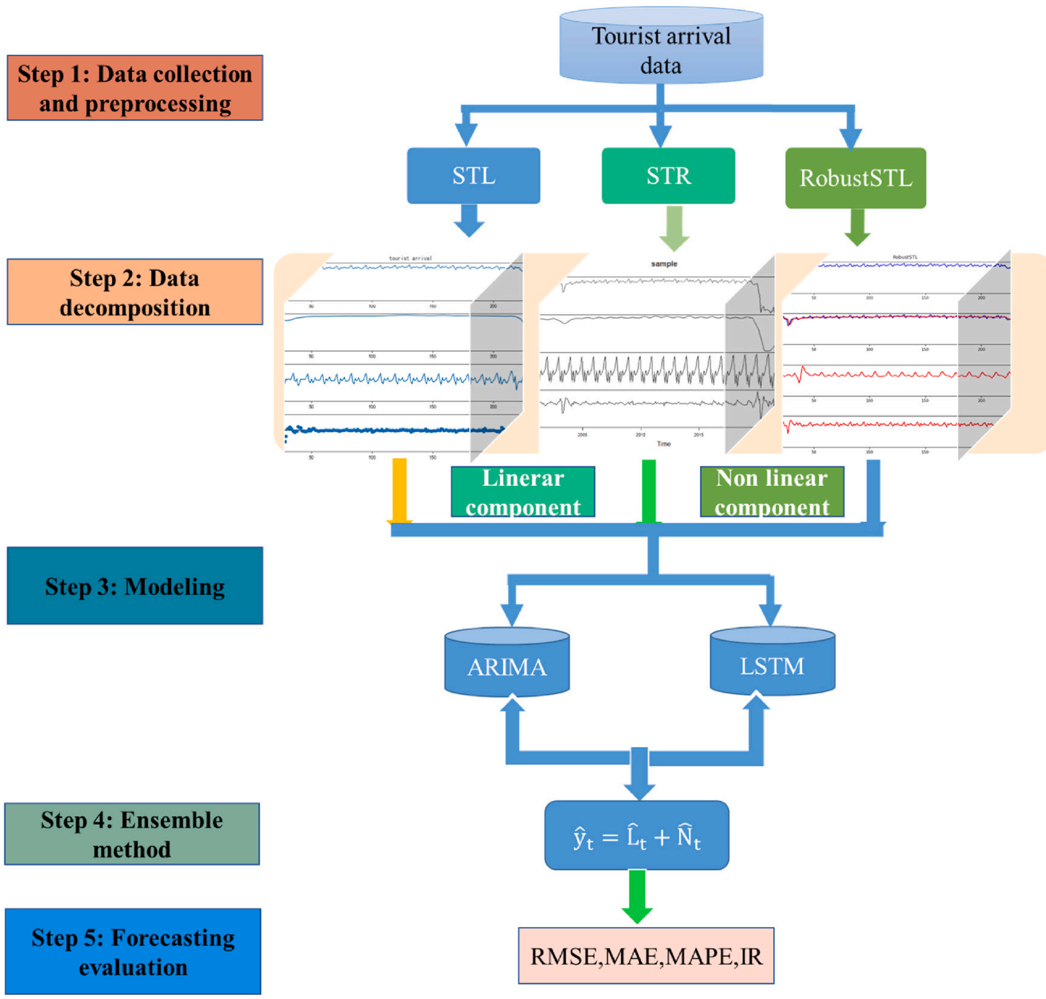

Fig. 5. Proposed forecasting framework.

Step 2: The data is then decomposed using the STL, STR, and RobustSTL decomposition algorithms, classifying the resulting subseries into linear and nonlinear components based on their data characteristics. These three methods were chosen for the following reasons:

First, these decomposition algorithms offer more accessible insights into seasonal, trend, and noise components than other complex decomposition techniques. This intuitive understanding is valuable in real-world applications, especially in decision making processes where clear comprehension of time series structures is vital.

Second, these methods tend to have lower computational costs and time complexity in comparison to some more intricate algorithms. This efficiency is particularly useful when handling large-scale tourism data, ensuring timely analysis and forecasting.

Third, the RobustSTL algorithm extends the capabilities of STL by incorporating outlier detection. In the context of tourism, unexpected events can create outliers that might distort the decomposition. RobustSTL's ability to identify and handle these anomalies ensures more accurate trend and seasonal component estimates, leading to more reliable forecasts.

In summary, by employing these three decomposition methods, the approach effectively addresses the complexities of tourism demand time series data. This careful selection allows for more precise modeling of patterns, fluctuations, and irregularities within the data, setting a solid foundation for subsequent forecasting models.

Step 3: To forecast the linear subseries, we employ the ARIMA model, renowned for its strong capability in linear fitting and wide application in time series forecasting. For the nonlinear subseries, the LSTM model is utilized, known for its excellent ability in handling complex, nonlinear patterns. The choice of ARIMA and LSTM is motivated by their complementary strengths and alignment with the specific characteristics of the time series data in tourism demand. Together, they offer a well-balanced approach that leverages the most suitable techniques for each component of the data.

Step 4: The forecasted results of each subseries are combined using an ensemble method, specifically additive integration, to derive the final forecasting outcome. By summing the individual forecasts, this method preserves the distinct characteristics of each subseries, ensuring that the overall forecast is a well-balanced representation of the underlying data.

Step 5: The integrated forecast results are evaluated using commonly adopted measures such as RMSE, mean absolute error (MAE), mean absolute percentage error (MAPE), and improvement rate (IR). The formulae for these measures are as follows:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(y_t - y_t^{'}\right)^2} \tag{9}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|y_t - y_t^{'}\right| \tag{10}$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|y_t - y_t^{'}\right|}{y_t} \tag{11}$$

$$IR_{i,j} = 1 - \frac{RMSE_i(MAE_i or MAPE_i)}{RMSE_j(MAE_j or MAPE_j)} \tag{12}$$

where $y_t^{'}$ and y_t capture predicted tourist arrivals and actual tourist arrival data. $RMSE_{i/j}$, $AE_{i/j}$, and $MAPE_{i/j}$ are the RMSEs, MAEs, and MAPEs of models i and j, respectively. $IR_{i,j}$, measures how model i improves forecasting accuracy compared with model j in terms of the reduction of forecasting errors of RMSEs, MAEs, and MAPEs.

Essentially, the forecasting can be summarized by the following equation, where $\widehat{L}_t$ is the predicted linear part and $\widehat{N}_t$ is the predicted nonlinear part:

$$\widehat{y}_t = \widehat{L}_t + \widehat{N}_t. \tag{13}$$

Through this approach, the ARIMA model's strengths in modeling linear data and the LSTM network's proficiency in modeling nonlinear components are optimally utilized. This methodology effectively investigates and leverages the intricate and complex features present in the time series of tourist arrivals, resulting in superior forecasting performance.

4. Empirical results

We conducted an empirical study using tourist arrival data to Hong Kong from six countries including Singapore, Korea, Japan, the Philippines, Canada, and United Kingdom to examine the performance of the proposed methods in tourism demand forecasting. Initially, we utilized three time series feature decomposition methods that are known for their interpretability to decompose the aforementioned data. Subsequently, we employed linear or nonlinear models to construct prediction models based on the characteristics of the decomposed components. We conducted predictions for 1-step, 3-step, 6-step, and 9-step ahead, and performed Diebold-Mariano (DM) tests to compare the predictive performance of the models. Finally, to emphasize the generalizability and broader applicability of our proposed forecasting framework, we conducted a supplementary empirical study focusing on outbound Chinese tourists traveling to Hawaii from January 2005 to January 2023. By comparing our approach with other benchmark models, we highlighted the robustness and flexibility of the proposed framework, affirming its relevance and practicality in various tourism demand forecasting scenarios.

4.1. Experimental design

4.1.1. Data description

We employ monthly data on visitor arrivals to Hong Kong from Singapore, Korea, Japan, the Philippines, Canada, and the United Kingdom to demonstrate the forecasting performance of the proposed framework. Hong Kong, a Special Administrative Region of China, has a thriving tourism industry and is a sought-after destination for international tourists, providing readily available tourism data that makes it a suitable subject for academic research. The rationale behind choosing these six source markets lies in their potential to be among the most crucial markets for Hong Kong's tourism industry. Representing the diversity of Southeast Asia, Europe, and North America, these countries enable us to acquire a broader understanding of travel patterns and elements influencing the tourism industry in these regions. Furthermore, the dependability and consistency of data originating from these countries might surpass that from other nations, bolstering our confidence in the data and guaranteeing that our findings are more precise and trustworthy.

Researchers can investigate various aspects of Hong Kong's tourism industry, such as its economic impact, cultural heritage, sustainability, and tourist behavior and preferences. Forecasting Hong Kong's tourism demand, given its unique characteristics and status, can provide valuable insights for tourism research and contribute to the growth of the industry.

The tourist arrival data used in this study were obtained from the Wind database (https://www.wind.com.cn) and cover the period from January 2001 to January 2023, comprising a total of 265 samples. Considering the significant impact of the COVID-19 pandemic in 2020, a comprehensive dataset consisting of 240 samples, spanning from January 2001 to December 2020, was utilized as the training set. This strategy aimed to enhance forecasting accuracy and equip the model with the ability to effectively learn from the unprecedented consequences of the pandemic, thereby contributing to more robust and reliable predictions. Additionally, a dataset of 25 samples from January 2021 to January 2023 was utilized as the testing set. To minimize the impact of outliers, all data were logarithmically transformed.

Fig. 6 displays the time series of tourist arrivals from the six selected countries, which exhibit a distinct seasonal periodicity, with the series from Korea and the Philippines displaying a certain trend feature. Notably, the SARS and COVID-19 epidemics had a significant impact on tourist arrivals from these countries, resulting in a substantial decline in both 2003 and 2020, with the latter year experiencing the most severe drop. The descriptive statistics results, presented in Table 2, indicate significant fluctuations in the time series data for three of the countries.

4.1.2. Model design

In this research, we employed three decomposition algorithms, namely STL, STR, and RobustSTL, along with two forecasting models, ARIMA and LSTM, culminating in four control groups. A depiction of the models is provided in Table 3.

The ARIMA model was determined using the R program, and the important parameters of the LSTM model were selected using the exhaustive grid search technique. Specifically, the learning rate was set to 0.01, batch size was set to 2, and the epoch was set to 500, with mean squared error serving as the loss function during training. The number of hidden layers in the LSTM network was set to 4, while other parameters were not discussed in detail.

It should be noted that both long-term and short-term forecasts are essential in tourism management decision-making. Thus, it is necessary to evaluate the performance of different models at different forecast horizons, particularly the h-step-ahead forecasting performance. In this study, h was set to 1, 3, 6, and 9 for comparison purposes.

4.2. Decomposition results

STL, STR, and RobustSTL are time series decomposition algorithms that are commonly used to partition time series data into three components: trend, seasonality, and residual. This class of approach allows for a more comprehensive understanding of the patterns and characteristics of the time series. Taking the Singapore dataset as an example, Fig. 7 demonstrates the decomposition outcomes of the three algorithms.

Fig. 7 indicates that RobustSTL behaves differently than STL and STR in the seasonality subsequence. In particular, the seasonality subsequence of RobustSTL demonstrates robustness by showcasing relatively stable oscillations during the initial and intermediate phases. Moreover, the extracted trend subsequence of RobustSTL demonstrates superior fit, agility in responding to sudden changes in trend, and robustness against outliers and non-stationary fluctuations.

In contrast, the trend subsequence of STL and STR exhibits inferior fit and limited responsiveness to changes. The trend subsequence of STL tends to lose its change pattern due to excessive smoothing, while the seasonality subsequence is significantly affected by offset and spike anomalies in the original sequence.

Furthermore, the STR algorithm does not assume a highly repetitive pattern for seasonality, making it more flexible than other algorithms. However, it is not effective in dealing with changes in trend. In contrast, the RobustSTL algorithm performs well in decomposing time series data into trend, seasonality, and residual parts, and demonstrates exceptional performance in handling outliers and changes in trend.

In summary, the RobustSTL algorithm offers several advantages in time series decomposition, and it can better handle a variety of time

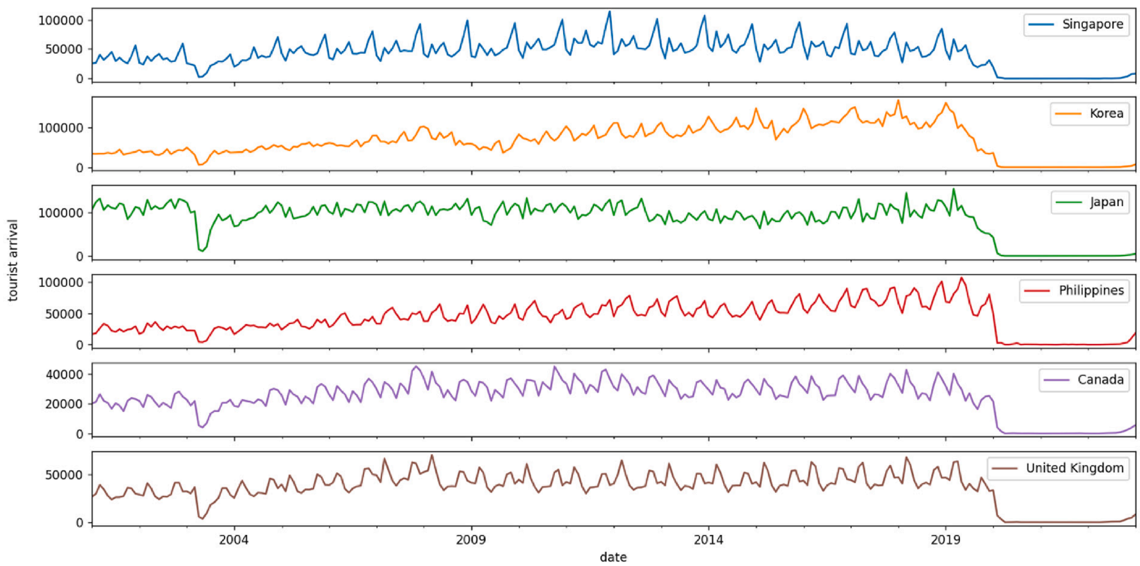

Fig. 6. Monthly tourist arrivals of Hong Kong from six countries.

Table 2
Descriptive analysis of Hong Kong tourist arrivals from six countries.

Source markets	mean	standard deviation	maximum	minimum	kurtosis	Skewness
Singapore	42961.54	24034.04	114295.00	7.00	0.10	−0.02
Korea	66109.65	39774.36	168152.00	4.00	−0.66	0.00
Japan	86385.45	38909.07	154586.00	3.00	0.66	−1.32
Philippines	42655.29	24417.86	107275.00	67.00	−0.49	−0.06
Canada	24569.72	11586.31	45148.00	11.00	0.12	−0.91
United Kingdom	35762.58	17086.46	70102.00	15.00	0.17	−0.81

Table 3
An introduction to constructed models.

Group	Models	Description
Group (1)	ARIMA and LSTM models	Baseline linear and nonlinear models
Group (2)	STL-ARIMA STR-ARIMA RobustSTL-ARIMA	Each subsequence obtained from decomposing the data using STL, STR, and RobustSTL algorithms combined with ARIMA for forecasting
Group (3)	STL-LSTM STR-LSTM RobustSTL-LSTM	Each subsequence obtained from decomposing the data using STL, STR, and RobustSTL algorithms combined with LSTM for forecasting
Group (4)	STL-ARIMA-LSTM STR-ARIMA-LSTM RobustSTL-ARIMA-LSTM	Linear components (trend) were obtained by decomposing the data using STL, STR, and RobustSTL algorithms combined with ARIMA for forecasting, and nonlinear components (seasonality and residuals) combined with LSTM for forecasting.

series data situations, thereby improving the accuracy and reliability of analysis results.

4.3. One-step forecasting evaluation

The importance of one-step forecasting in tourism management cannot be overstated. One-step forecasting serves as a valuable tool for tourism managers to gain rapid insight into future demand trends, enabling them to make timely adjustments in tourism product development, pricing, and marketing strategies. Additionally, one-step forecasting assists tourism practitioners in optimizing resource allocation, improving tourism service efficiency and quality, and ultimately boosting tourism revenue and the overall competitiveness of the tourism industry. The findings of the one-step forecasting are presented in Table 4.

Table 4 reveals that Group (2) exhibits superior forecasting accuracy in comparison to models of Group (1). Furthermore, Group (3) demonstrates a modest improvement in forecasting accuracy, except for STL-LSTM, which fails to handle abrupt trend shifts effectively. Notably, both STR-LSTM and RobustSTL-LSTM exhibit enhancements in forecasting performance, signifying that decomposition algorithms can improve predictive accuracy to a certain extent. Remarkably, Group (4) manifests significantly lower forecasting errors than the other three groups, thereby substantiating that the selection of different models based on the features of sub-sequences after time series decomposition is superior to using a single model or the same model for each decomposed sequence. Additionally, RobustSTL-ARIMA-LSTM achieves the most satisfactory forecasting results in all six datasets, underscoring the robustness of the RobustSTL algorithm to noise and outliers and its capacity to handle sudden trend changes effectively.

Fig. 8 illustrates that the MAPE values for one-step forecasting of RobustSTL-ARIMA-LSTM are consistently the best across all six datasets.

4.4. Multi-step forecasting evaluation

Multi-step forecasting is crucial in tourism management as it provides longer-term forecasting results compared to single-step forecasting, enabling tourism practitioners to make more informed and forward-looking decisions. For instance, multi-step forecasting aids in predicting the tourism demand trend for the next several months or years, facilitating the creation of comprehensive business plans and development strategies. Additionally, it assists tourism practitioners in responding better to market fluctuations and changes, mitigating operational risks. Through multi-step forecasting, tourism practitioners can gain a more comprehensive understanding of market demand and consumer trends, meet customer needs, enhance the quality and efficiency

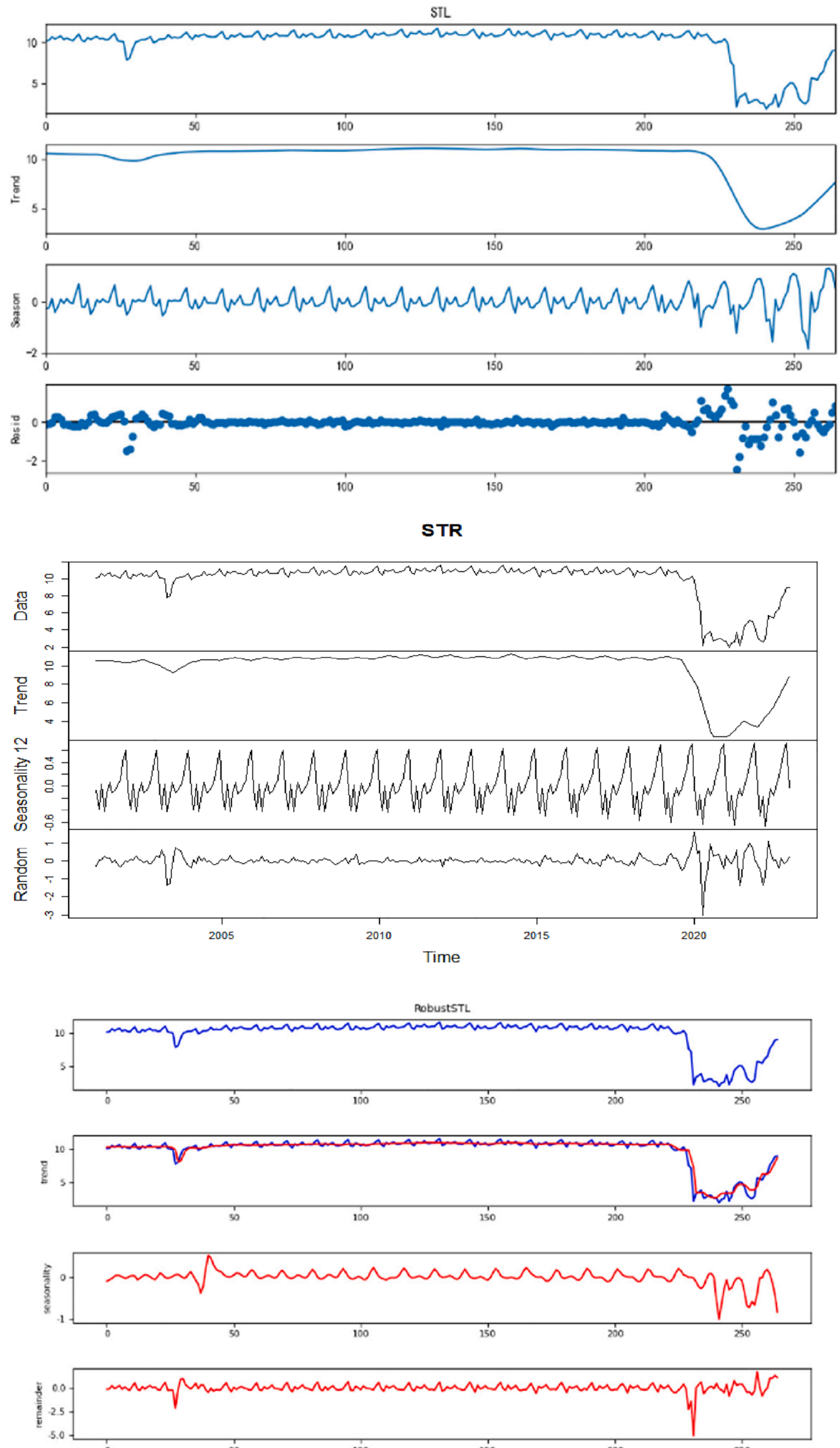

Fig. 7. Three decomposition results for the Singapore dataset.

of tourism services, and increase revenue and competitiveness. Tables 5–10 present the results of multi-step forecasting, including forecasts for 3, 6, and 9 steps ahead.

The results of multi-step forecasting indicate that the predictive accuracy of Group (2) and Group (3) models has improved in comparison to the ARIMA and LSTM models of Group (1). Additionally, the Group (4) model demonstrated superior predictive accuracy to other Group models, with the RobustSTL-ARIMA-LSTM model yielding the smallest forecast error and highest predictive accuracy within Group (4). These findings demonstrate that selecting different models for ensemble forecasting based on the data characteristics of different sub-sequences after decomposition can overall improve the predictive accuracy.

Specifically, when the forecast horizon reaches 9, the predictive performance of Group (2) models significantly declines, although they still outperform the benchmark ARIMA model. This suggests that ARIMA performs poorly in fitting nonlinear sub-sequences (seasonal and

Table 4
Evaluation of one-step forecasting results in six datasets.

Model		Singapore			Korea		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	0.51	0.267	0.048	0.441	0.245	0.034
	LSTM	2.503	1.134	0.292	1.743	0.909	0.133
Group (2)	STL-ARIMA	0.312	0.159	0.029	0.293	0.149	**0.019**
	STR-ARIMA	0.441	0.282	0.044	0.362	0.216	0.026
	RobustSTL-ARIMA	0.301	0.15	0.028	0.221	0.149	0.021
Group (3)	STL-LSTM	2.554	1.098	0.29	1.773	0.907	0.133
	STR-LSTM	2.473	1.089	0.283	1.732	0.901	0.131
	RobustSTL-LSTM	2.308	0.985	0.263	1.727	0.851	0.127
Group (4)	STL-ARIMA-LSTM	0.448	0.214	0.043	0.41	0.215	0.027
	STR-ARIMA-LSTM	0.281	0.155	0.026	0.419	0.217	0.027
	RobustSTL-ARIMA-LSTM	**0.187**	**0.123**	**0.022**	**0.189**	**0.152**	0.02
		Japan			Philippines		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	0.385	0.218	0.033	0.441	0.245	0.034
	LSTM	2.089	0.923	0.203	1.743	0.909	0.133
Group (2)	STL-ARIMA	0.259	0.119	0.021	0.293	0.149	**0.019**
	STR-ARIMA	0.272	0.166	0.024	0.362	0.216	0.026
	RobustSTL-ARIMA	0.217	0.129	0.021	0.221	0.149	0.021
Group (3)	STL-LSTM	2.015	0.862	0.193	1.773	0.907	0.133
	STR-LSTM	2.083	0.908	0.199	1.732	0.901	0.131
	RobustSTL-LSTM	2.061	0.888	0.197	1.727	0.851	0.127
Group (4)	STL-ARIMA-LSTM	0.389	0.173	0.03	0.41	0.215	0.027
	STR-ARIMA-LSTM	0.293	0.139	0.023	0.419	0.217	0.027
	RobustSTL-ARIMA-LSTM	**0.176**	**0.122**	**0.018**	**0.186**	**0.152**	0.02
		Canada			United Kingdom		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	0.385	0.218	0.033	0.424	0.223	0.041
	LSTM	2.089	0.923	0.203	2.201	0.952	0.208
Group (2)	STL-ARIMA	0.259	0.119	0.021	0.281	0.124	**0.021**
	STR-ARIMA	0.272	0.166	0.024	0.297	0.169	0.024
	RobustSTL-ARIMA	0.217	0.129	0.021	0.201	0.121	0.02
Group (3)	STL-LSTM	2.015	0.862	0.193	2.284	0.955	0.212
	STR-LSTM	2.083	0.908	0.199	2.212	0.933	0.206
	RobustSTL-LSTM	2.061	0.888	0.197	2.234	0.945	0.21
Group (4)	STL-ARIMA-LSTM	0.389	0.173	0.03	0.388	0.183	0.029
	STR-ARIMA-LSTM	0.293	0.139	0.023	0.283	0.153	0.024
	RobustSTL-ARIMA-LSTM	**0.178**	**0.122**	**0.018**	**0.172**	**0.153**	**0.021**

residual) after decomposition, which becomes more evident with an increase in the forecast horizon. Furthermore, the multi-step forecast error of Group (3) models is better than the benchmark LSTM; however, the performance of LSTM in fitting linear trend components after decomposition is poor, making it unsuitable for standalone forecasting, resulting in relatively larger errors.

The results of our analysis indicate that the Group (4) model exhibits significantly higher forecast accuracy compared to the other forecasting models considered in our study. This result further emphasizes the significance of selecting appropriate forecasting models that are suited to the characteristics of the sub-sequences obtained from the decomposition process. It underlines the importance of considering the unique features of each sub-sequence in order to achieve optimal forecasting results. Furthermore, our findings show that the RobustSTL-ARIMA-LSTM model consistently outperforms other models in terms of multi-step forecasting accuracy across all six datasets. This demonstrates the superiority of the RobustSTL algorithm in addressing sudden changes in trend, irregular seasonality components, and random fluctuations in residuals. These results are further supported by the evidence presented in Fig. 8. In conclusion, our multi-step forecasting results align with the findings of our single-step forecasting analysis, further underscoring the efficacy of the proposed method.

To demonstrate the efficacy of RobustSTL-ARIMA-LSTM in a clear and concise manner, we evaluate the improvement rate calculated based on RMSE values for the Singapore dataset in comparison to other models, indicated in Fig. 9 and Table 11.

As shown in Table 11 and Fig. 9, the improvement rate of RobustSTL-ARIMA-LSTM compared to other models is substantial, with a maximum improvement rate of 92.5% for single-step forecasting and 93.1% for multi-step forecasting. The average overall improvement rate of RMSE surpasses 75%, highlighting the exceptional performance of the RobustSTL algorithm and the soundness of selecting different models for predicting decomposed sub-sequences based on their respective data characteristics.

4.5. DM test

DM test is a commonly applied method for evaluating the comparative predictive accuracy of two time series models and determining which one has superior performance. To assess the significance of the superiority of the RobustSTL decomposition algorithm over other decomposition algorithms, a one-step-ahead forecasting DM test was conducted. The results of this test are presented in Table 12.

Based on the results presented in Table 12, it is apparent that the RobustSTL decomposition algorithm generally outperforms the STL and STR decomposition algorithms to varying extents, with the RobustSTL-ARIMA-LSTM model exhibiting the most substantial improvement. These results are statistically significant at the 10% significance level, providing robust evidence for the superiority of the RobustSTL decomposition algorithm in both one-step and multi-step forecasting scenarios.

4.6. Robustness analysis

To further scrutinize the model's robustness and examine its potential for generalization, we sought to apply it in forecasting tourism demand within different contexts. Hawaii was selected as the target destination, owing to the availability of pertinent tourism data (accessed via https://dbedt.hawaii.gov/visitor/tourismdata/). Specifically, we

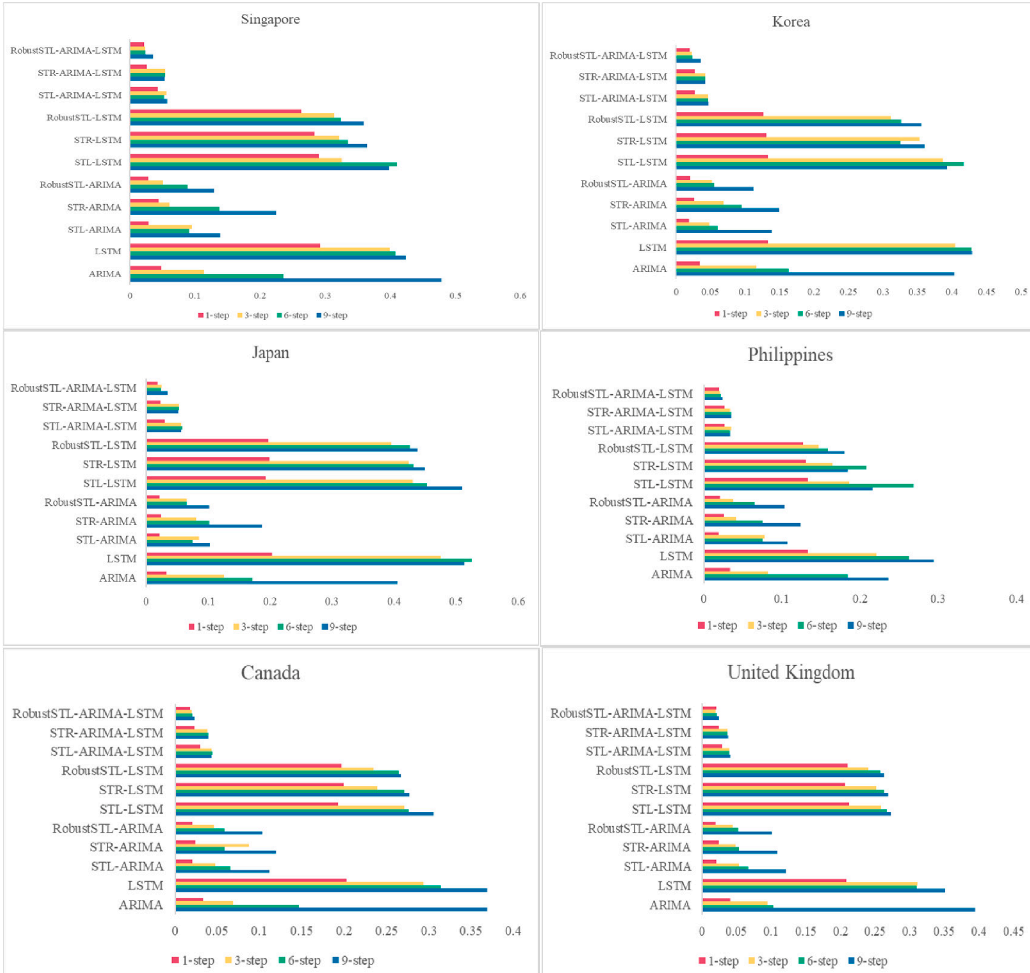

Fig. 8. The MAPE for one-step- and multi-step-ahead forecast results for all six datasets.

Table 5
Evaluation of multi-step forecasting results (Singapore dataset).

	Singapore	3-step			6-step			9-step		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	1.223	0.986	0.114	2.283	1.861	0.236	3.708	3.321	0.479
	LSTM	3.171	3.068	0.399	3.236	3.124	0.408	3.506	3.427	0.424
Group (2)	STL-ARIMA	0.87	0.841	0.095	0.904	0.762	0.091	1.333	1.067	0.139
	STR-ARIMA	0.529	0.487	0.06	1.027	0.879	0.137	1.353	1.113	0.224
	RobustSTL-ARIMA	0.505	0.442	0.051	0.832	0.709	0.089	1.161	1.045	0.129
Group (3)	STL-LSTM	2.522	1.662	0.325	3.11	2.807	0.41	3.013	2.575	0.398
	STR-LSTM	2.474	1.637	0.321	2.558	1.823	0.335	2.724	2.245	0.364
	RobustSTL-LSTM	2.406	1.606	0.314	2.504	1.811	0.324	2.682	2.024	0.359
Group (4)	STL-ARIMA-LSTM	0.582	0.362	0.056	0.509	0.359	0.052	0.602	0.372	0.057
	STR-ARIMA-LSTM	0.525	0.336	0.054	0.515	0.331	0.054	0.516	0.327	0.053
	RobustSTL-ARIMA-LSTM	**0.216**	**0.185**	**0.023**	**0.233**	**0.195**	**0.024**	**0.352**	**0.218**	**0.035**

engaged the dataset of outbound Chinese tourists to Hawaii spanning January 2005 to January 2023, incorporating 217 data points. An 8:2 ratio was applied for the training and testing split, and the time series plot corresponding to this dataset is illustrated in Fig. 10.

In our analysis, we extended beyond the standard benchmark models such as ARIMA and LSTM by including three AI techniques frequently utilized in relevant literature as benchmarks: gradient boosting regression trees (GBRT), temporal convolutional network (TCN), and DeepAR. By comparing our method against these established techniques, we aim to demonstrate its adaptability, robustness, and potential advantages. A comprehensive discussion of the forecasting evaluation for these models is included in our analysis. This sheds light on their comparative performance within the context of this study, as shown in Table 13.

Table 13 offers a detailed comparative analysis of the forecasting performance across various horizons (1-, 3-, 6-, and 9-step). The insights drawn from the evaluation can be summarized as follows:

First, the proposed model demonstrates remarkable proficiency, consistently outperforming the benchmark models across all forecasted horizons by manifesting the lowest RMSE, MAE, and MAPE values. These findings underline its robustness in accurately capturing the underlying temporal dynamics.

Second, traditional benchmarks such as the ARIMA model exhibiting

Table 6
Evaluation of multi-step forecasting results (Korea dataset).

Korea		3-step			6-step			9-step		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	1.11	0.98	0.116	1.388	1.151	0.163	2.979	2.633	0.403
	LSTM	3.315	3.221	0.404	3.636	3.566	0.428	3.642	3.571	0.429
Group (2)	STL-ARIMA	0.56	0.418	0.048	0.599	0.469	0.06	1.349	1.024	0.139
	STR-ARIMA	0.533	0.531	0.069	0.649	0.621	0.095	1.042	0.815	0.149
	RobustSTL-ARIMA	0.451	0.303	0.052	0.545	0.439	0.055	0.998	0.791	0.112
Group (3)	STL-LSTM	3.01	2.732	0.386	3.49	2.839	0.417	3.132	2.635	0.393
	STR-LSTM	2.756	2.197	0.3525	2.683	1.785	0.325	2.855	2.287	0.36
	RobustSTL-LSTM	2.509	1.737	0.311	2.679	1.724	0.326	2.765	2.313	0.355
Group (4)	STL-ARIMA-LSTM	0.519	0.271	0.046	0.518	0.273	0.046	0.513	0.268	0.047
	STR-ARIMA-LSTM	0.445	0.235	0.042	0.447	0.24	0.042	0.443	0.234	0.042
	RobustSTL-ARIMA-LSTM	**0.225**	**0.182**	**0.023**	**0.234**	**0.191**	**0.024**	**0.354**	**0.218**	**0.036**

Table 7
Evaluation of multi-step forecasting results (Japan dataset).

	Japan	3-step			6-step			9-step		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	1.075	0.977	0.126	1.616	1.249	0.172	2.756	2.472	0.405
	LSTM	3.623	3.467	0.475	4.792	4.702	0.525	4.184	4.066	0.513
Group (2)	STL-ARIMA	0.71	0.696	0.085	0.667	0.565	0.075	1.048	0.748	0.103
	STR-ARIMA	0.635	0.601	0.081	0.695	0.642	0.102	1.119	0.901	0.187
	RobustSTL-ARIMA	0.602	0.528	0.065	0.652	0.531	0.065	1.016	0.681	0.102
Group (3)	STL-LSTM	3.108	2.33	0.43	3.317	2.644	0.453	4.003	3.776	0.51
	STR-LSTM	3.036	2.261	0.423	3.158	2.313	0.431	3.281	2.669	0.449
	RobustSTL-LSTM	2.819	1.931	0.396	3.014	2.268	0.425	3.114	2.578	0.438
Group (4)	STL-ARIMA-LSTM	0.558	0.288	0.057	0.576	0.329	0.059	0.553	0.283	0.057
	STR-ARIMA-LSTM	0.479	0.265	0.053	0.477	0.262	0.053	0.469	0.252	0.052
	RobustSTL-ARIMA-LSTM	**0.228**	**0.183**	**0.025**	**0.239**	**0.189**	**0.024**	**0.345**	**0.101**	**0.035**

Table 8
Evaluation of multi-step forecasting results (Philippines dataset).

Philippines		3-step			6-step			9-step		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	1.071	0.744	0.082	1.914	1.52	0.184	2.171	1.795	0.236
	LSTM	2.118	2.025	0.221	2.596	2.547	0.263	3.026	2.949	0.294
Group (2)	STL-ARIMA	0.881	0.744	0.078	0.795	0.641	0.075	1.353	0.933	0.107
	STR-ARIMA	0.508	0.363	0.041	0.689	0.595	0.075	0.907	0.801	0.124
	RobustSTL-ARIMA	0.506	0.361	0.038	0.677	0.565	0.065	0.807	0.777	0.103
Group (3)	STL-LSTM	1.852	1.564	0.186	2.624	2.572	0.269	2.149	1.845	0.216
	STR-LSTM	1.794	1.304	0.165	2.058	1.792	0.208	1.971	1.478	0.184
	RobustSTL-LSTM	1.686	1.118	0.147	1.706	1.252	0.159	1.848	1.449	0.18
Group (4)	STL-ARIMA-LSTM	0.467	0.292	0.035	0.458	0.283	0.034	0.458	0.278	0.034
	STR-ARIMA-LSTM	0.48	0.291	0.034	0.482	0.297	0.035	0.484	0.293	0.035
	RobustSTL-ARIMA-LSTM	**0.198**	**0.127**	**0.021**	**0.211**	**0.133**	**0.022**	**0.231**	**0.161**	**0.024**

Table 9
Evaluation of multi-step forecasting results (Canada dataset).

	Canada	3-step			6-step			9-step		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	0.693	0.558	0.069	1.418	1.121	0.146	2.819	2.517	0.369
	LSTM	2.352	2.143	0.294	2.546	2.373	0.314	3.227	3.168	0.369
Group (2)	STL-ARIMA	0.436	0.405	0.048	0.595	0.517	0.065	1.022	0.827	0.112
	STR-ARIMA	0.687	0.672	0.087	0.479	0.423	0.059	0.977	0.741	0.119
	RobustSTL-ARIMA	0.431	0.401	0.046	0.477	0.421	0.059	0.987	0.732	0.103
Group (3)	STL-LSTM	2.25	1.772	0.271	2.286	1.844	0.276	2.518	2.177	0.306
	STR-LSTM	2.07	1.417	0.239	2.286	1.748	0.271	2.298	1.871	0.277
	RobustSTL-LSTM	2.021	1.394	0.235	2.152	1.738	0.264	2.237	1.68	0.267
Group (4)	STL-ARIMA-LSTM	0.481	0.284	0.043	0.486	0.293	0.044	0.475	0.271	0.043
	STR-ARIMA-LSTM	0.401	0.235	0.038	0.402	0.242	0.039	0.403	0.24	0.039
	RobustSTL-ARIMA-LSTM	**0.196**	**0.123**	**0.020**	**0.208**	**0.136**	**0.021**	**0.234**	**0.180**	**0.023**

Table 10
Evaluation of multi-step forecasting results (UK dataset).

United Kingdom		3-step			6-step			9-step		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Group (1):	ARIMA	0.992	0.79	0.094	0.954	0.79	0.103	3.073	2.775	0.394
	LSTM	2.583	2.427	0.311	2.572	2.347	0.31	3.042	2.963	0.351
Group (2)	STL-ARIMA	0.485	0.457	0.053	0.623	0.549	0.067	1.024	0.899	0.121
	STR-ARIMA	0.417	0.384	0.048	0.439	0.371	0.053	0.784	0.592	0.109
	RobustSTL-ARIMA	0.417	0.381	0.045	0.433	0.361	0.052	0.781	0.498	0.101
Group (3)	STL-LSTM	2.272	1.607	0.258	2.295	1.757	0.267	2.376	1.789	0.273
	STR-LSTM	2.211	1.543	0.251	2.321	1.668	0.263	2.321	1.775	0.269
	RobustSTL-LSTM	2.163	1.413	0.24	2.242	1.61	0.257	2.244	1.694	0.263
Group (4)	STL-ARIMA-LSTM	0.478	0.282	0.04	0.481	0.282	0.04	0.476	0.284	0.041
	STR-ARIMA-LSTM	0.417	0.258	0.037	0.417	0.26	0.037	0.416	0.261	0.038
	RobustSTL-ARIMA-LSTM	**0.202**	**0.121**	**0.020**	**0.208**	**0.128**	**0.022**	**0.236**	**0.168**	**0.024**

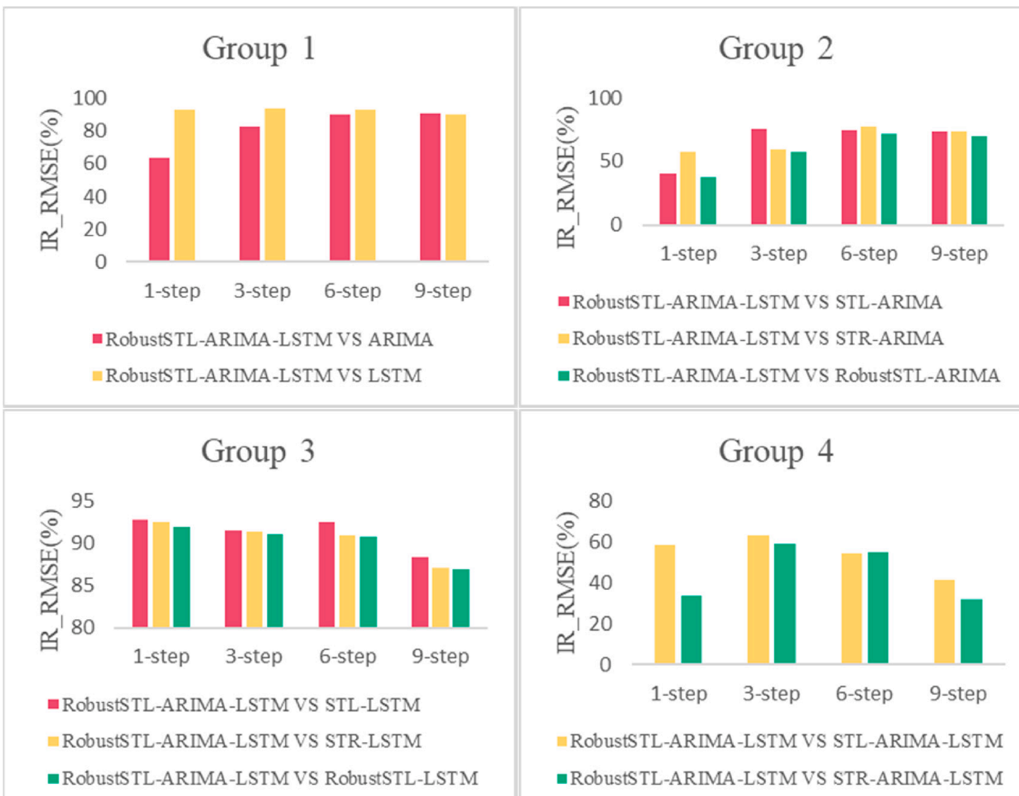

Fig. 9. IR based on RMSE of the proposed framework for the Singapore dataset.

Table 11
IR based on RMSE of the proposed framework for the Singapore dataset.

RobustSTL-ARIMA-LSTM *VS*	1-step	3-step	6-step	9-step
ARIMA	63.33	82.34	89.79	90.51
LSTM	92.53	93.19	92.80	89.96
STL-ARIMA	40.06	75.17	74.23	73.59
STR-ARIMA	57.60	59.17	77.3	73.98
RobustSTL-ARIMA	37.87	57.23	72.00	69.68
STL-LSTM	92.68	91.44	92.51	88.32
STR-LSTM	92.44	91.27	90.89	87.08
RobustSTL-LSTM	91.90	91.02	90.69	86.88
STL-ARIMA-LSTM	58.26	62.89	54.22	41.53
STR-ARIMA-LSTM	33.45	58.86	54.76	31.78

strong performance at the 1-step horizon, its accuracy diminishes for longer-term predictions, possibly due to its relative simplicity. Meanwhile, the LSTM model maintains steady performance across all horizons but falls behind the proposed model, which may indicate a slight overfitting on longer horizons.

Third, other AI models reveal average performance across all forecasting horizons. Challenges in capturing sufficient temporal information and dealing with long-term dependencies might contribute to their relatively lackluster performance.

In conclusion, the analysis accentuates the robustness and adaptability of the proposed RobustSTL-ARIMA-LSTM model in comparison to the established benchmark models. The performance of each model reflects its inherent strengths and limitations, providing valuable insights into their applicability and efficiency in the field of tourism demand forecasting.

5. Conclusion

This study endeavors to investigate and address the robustness phenomenon associated with decomposed components and

Table 12
The DM test results (proposed model compared to other models).

Method	Markets	*vs* STL-ARIMA-LSTM	*vs* STR-ARIMA-LSTM
RobustSTL-ARIMA-LSTM	Singapore	−3.395***	−2.279**
	Korea	−2.927***	−3.261**
	Japan	−3.952***	−2.101**
	Philippines	−4.252***	−3.845***
	Canada	−4.338***	−3.052***
	United Kingdom	−3.404**	−2.553***
Method	Markets	*vs* STL-ARIMA	*vs* STR-ARIMA
RobustSTL-ARIMA	Singapore	−2.692***	−3.194***
	Korea	−3.168***	−2.816***
	Japan	−3.386***	−2.332***
	Philippines	−2.011**	−2.903***
	Canada	−1.908***	−2.807***
	United Kingdom	−2.101**	−2.977***
Method	Markets	*vs* STL-LSTM	*vs* STR-LSTM
RobustSTL-LSTM	Singapore	−5.349***	−4.167***
	Korea	−4.527***	−5.543***
	Japan	−4.164***	−3.752***
	Philippines	−3.329***	−3.420***
	Canada	−1.535*	−4.544***
	United Kingdom	−2.972***	−1.127*

Note: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

subsequently proposes an innovative hybrid model, namely, the RobustSTL-ARIMA-LSTM, to augment the accuracy of tourism demand forecasting. An empirical investigation was carried out to forecast tourism demand from six countries to Hong Kong, encompassing both one-step and multi-step forecasting. The study's findings demonstrate that the employment of the novel RobustSTL method for time series data decomposition, in contrast to existing decomposition techniques, can result in enhanced predictive accuracy, as well as robustness, interpretability, and reduced complexity. Through the application of STL, STR, and RobustSTL algorithms for decomposing monthly tourist arrivals data from Singapore, Korea, Japan, the Philippines, Canada, and the United Kingdom to Hong Kong, a notable improvement in the forecasting performance of the model can be achieved. The improvement achieved by the RobustSTL-ARIMA-LSTM model, as compared to other models, is considerable, showcasing a maximal improvement rate of 92.5% for single-step forecasting and 93.1% for multi-step forecasting.

The proposed RobustSTL-ARIMA-LSTM model's adaptability within the tourism sector is further emphasized through a robustness analysis using an additional dataset. The model was applied to forecast the outbound Chinese tourists to Hawaii from January 2005 to January 2023. This successful application indicates that the model's robust combination of decomposition and forecasting techniques is not confined to a single tourism context. The favorable forecasting performance in the robustness analysis reinforces the models' potential as a valuable tool for diverse forecasting scenarios. While the empirical analyses in this study were concentrated on tourism demand forecasting, the adaptability and robustness of the RobustSTL-ARIMA-LSTM model suggest potential applicability beyond the confines of this specific sector. The combination of RobustSTL decomposition for extracting trend-seasonality, ARIMA for modeling linear patterns, and LSTM for capturing nonlinear dynamics is a flexible and sophisticated analytical framework that could be leveraged in various forecasting scenarios, such as the retail market where similar trends, seasonality, and nonlinear patterns may be present.

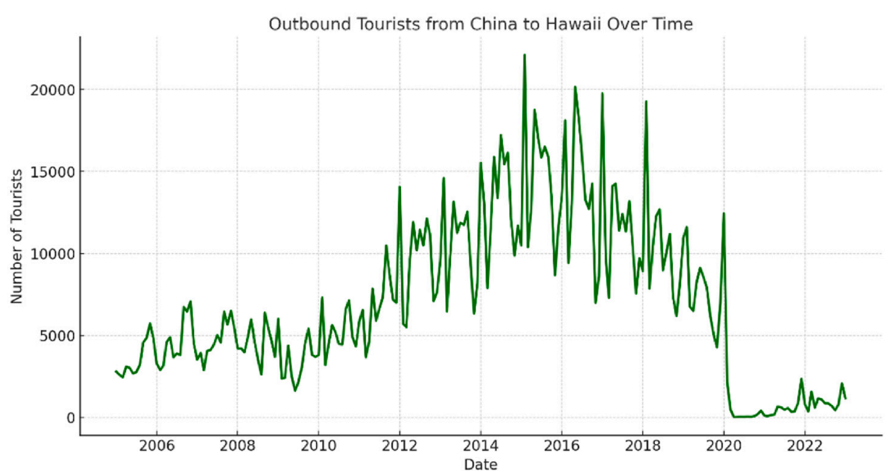

Fig. 10. Dataset of outbound tourists from China to Hawaii.

Table 13
Forecasting evaluation in robustness analysis.

Out-of-sample forecasting	Criteria	RobustSTL-ARIMA-LSTM	ARIMA	LSTM	GBRT	DeepAR	TCN
1-step	RMSE	**0.549**	0.745	1.773	4.421	4.218	4.133
	MAE	**0.262**	0.498	1.433	3.51	3.315	3.906
	MAPE	**0.042**	5.187	0.989	1.363	1.061	1.676
3-step	RMSE	**0.738**	0.978	1.277	4.649	4.438	4.103
	MAE	**0.439**	0.723	1.225	3.729	3.544	5.983
	MAPE	**0.074**	0.201	0.175	1.25	1.061	1.61
6-step	RMSE	**0.754**	1.005	1.571	5.011	4.744	7.07
	MAE	0.533	**0.521**	1.518	4.143	3.866	6.134
	MAPE	**0.079**	0.167	0.228	1.444	1.128	1.441
9-step	RMSE	**0.74**	1.115	1.783	5.287	5.027	6.951
	MAE	**0.468**	0.611	1.735	4.456	4.168	6.142
	MAPE	**0.076**	0.178	0.258	1.458	1.173	1.61

Following the decomposition of the time series, the importance of selecting ensemble algorithms based on the data characteristics is emphasized. The RobustSTL-ARIMA-LSTM hybrid model capitalizes on the strengths of the ARIMA model in linear fitting and the LSTM neural network in non-linear fitting, thereby augmenting the forecasting accuracy. Furthermore, the RobustSTL decomposition algorithm outperforms the STL and STR decomposition methods in the context of tourism demand forecasting. Regardless of whether it is one-step or multi-step forecasting, RobustSTL demonstrates superior predictive performance compared to STL and STR. The average overall decrease in RMSE surpasses 75%, emphasizing the exceptional performance of the RobustSTL algorithm and the sound rationale behind selecting distinct models for predicting decomposed sub-sequences based on their individual data characteristics.

Our study makes a significant contribution by establishing a sophisticated forecasting framework founded on a robust and reliable decomposition method. While numerous studies have implemented decomposition-ensemble approaches, the majority of existing research predominantly depends on methods such as EEMD, VMD, and their variants (Li & Law, 2020; Zhang et al., 2022; Zhang & Tian, 2022). These methods, although capable of attaining a certain degree of accuracy, often yield the decomposed components (e.g., IMFs) that lack intuitive interpretability, thereby constraining the practical application of decomposition-ensemble concepts in tourism forecasting. In contrast, the RobustSTL algorithm adopted in this study facilitates the extraction of trend-seasonality with enhanced robustness, making it more interpretable and suitable for various forecasting contexts. Consequently, our study employs a decomposition method that specifically considers the characteristics of the decomposed components, thus establishing a novel decomposition-ensemble forecasting framework. To the best of our knowledge, our investigation is the first to integrate a robust and interpretable decomposition algorithm with linear and non-linear models. For linear components, the ARIMA model is demonstrably more appropriate than complex deep learning models, while for non-linear components, deep learning models such as LSTM offer more precise predictions of tourism demand fluctuations.

This article is not without its limitations. First, although we conducted an in-depth evaluation of the performance of decomposition methods based on time series trend and seasonal characteristics (i.e., STL and STR) in tourism demand forecasting and applied a robust RobustSTL approach to this context, future research could further explore novel methods exhibiting robust decomposition features, favorable interpretability, low complexity, and adaptability to complex time series. Second, our study constructed a framework based on a combination of 11 forecasting models from four groups, using ARIMA and LSTM models as benchmarks and selecting prediction models according to the decomposed data characteristics. Future research could incorporate additional cutting-edge forecasting methods, such as transformer-based advanced deep learning techniques. Another limitation pertains to the sensitivity of the results to specific parameter settings of the RobustSTL-ARIMA-LSTM model. The careful tuning of these parameters is critical, as improper selections may lead to the suboptimal forecasting accuracy. Furthermore, the presence of outliers in the dataset, particularly unexpected spikes or drops, can affect the robustness of the decomposition process, thereby affecting the forecasting performance. It may also limit the model's generalizability to other domains without proper adjustments. Future research could address these limitations by exploring adaptive mechanisms for parameter tuning and implementing strategies to identify and mitigate the effects of outliers, thus enhancing the overall robustness and versatility of the forecasting methodology.

CRediT authorship contribution statement

Xin Li: Conceptualization, Methodology, Writing – review & editing. **Xu Zhang:** Methodology, Software, Visualization. **Chengyuan Zhang:** Investigation, Writing – review & editing. **Shouyang Wang:** Supervision, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the support of research fund from the National Natural Science Foundation of China (Nos. 72371025, 72025101, 72201201), Science and Technology Think Tank Young Talent Program from China Association for Science and Technology, China (No. 20220615ZZ07110103), and the Fundamental Research Funds for the Central Universities, China under Grant No. QTZX23085 and No. ZYTS23005.

References

Adil, M., Wu, J. Z., Chakrabortty, R. K., Alahmadi, A., Ansari, M. F., & Ryan, M. J. (2021). Attention-based STL-BiLSTM network to forecast tourist arrival. *Processes, 9*(10), 1759.

Apergis, N., Mervar, A., & Payne, J. E. (2017). Forecasting disaggregated tourist arrivals in Croatia: Evidence from seasonal univariate time series models. *Tourism Economics, 23*(1), 78–98.

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*(3), 383–397.

Baldigara, T., & Mamula, M. (2015). Modelling international tourism demand using seasonal ARIMA models. *Tourism and Hospitality Management, 21*(1), 19–31.

Bi, J. W., Han, T. Y., & Yao, Y. (2023). Fine-grained tourism demand forecasting: A decomposition ensemble deep learning model. *Tourism Economics*, Article 13548166231158705.

Chen, C. F., Lai, M. C., & Yeh, C. C. (2012). Forecasting tourism demand based on empirical mode decomposition and neural network. *Knowledge-Based Systems, 26*, 281–287.

Chen, R., Liang, C. Y., Hong, W. C., & Gu, D. X. (2015). Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm. *Applied Soft Computing, 26*, 435–443.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. *Journal of Official Statistics, 6*(1), 3–73.

Dokumentov, A., & Hyndman, R. J. (2015). STR: A seasonal-trend decomposition procedure based on regression. *Monash Econometrics and Business Statistics Working Papers, 13*(15), 2015–2203.

Feng, Y., Hao, J., Sun, X., & Li, J. (2022). Forecasting short-term tourism demand with a decomposition-ensemble strategy. *Procedia Computer Science, 199*, 879–884.

Hailemariam, A., & Ivanovski, K. (2021). The impact of geopolitical risk on tourism. *Current Issues in Tourism, 24*(22), 3134–3140.

Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., & Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research, 63*, 112–127.

He, K., Ji, L., Wu, C. W. D., & Tso, K. F. G. (2021). Using SARIMA–CNN–LSTM approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management, 49*, 25–33.

Hu, Y. C. (2021). Forecasting tourism demand using fractional grey prediction models with Fourier series. *Annals of Operations Research, 300*(2), 467–491.

Hu, M., Li, H., Song, H., Li, X., & Law, R. (2022). Tourism demand forecasting using tourist-generated online review data. *Tourism Management, 90*, Article 104490.

Kon, S. C., & Turner, L. W. (2005). Neural network forecasting of tourism demand. *Tourism Economics, 11*(3), 301–328.

Kožić, I. (2014). Detecting international tourism demand growth cycles. *Current Issues in Tourism, 17*(5), 397–403.

Kummong, R., & Supratid, S. (2016). Thailand tourism forecasting based on a hybrid of discrete wavelet decomposition and NARX neural network. *Industrial Management and Data Systems, 116*(6), 1242–1258.

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, X., Law, R., Xie, G., & Wang, S. (2021). Review of tourism forecasting research with internet data. *Tourism Management, 83*, Article 104245.

Li, C., Zheng, W., & Ge, P. (2022). Tourism demand forecasting with spatiotemporal features. *Annals of Tourism Research, 94*, Article 103384.

Li, S., Chen, T., Wang, L., & Ming, C. (2018). Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu Index. *Tourism Management, 68*, 116–126.

Li, X., & Law, R. (2020). Forecasting tourism demand with decomposed search cycles. *Journal of Travel Research, 59*(1), 52–68.

Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management, 59*, 57–66.

Lim, C., & McAleer, M. (2001). Forecasting tourist arrivals. *Annals of Tourism Research, 28*(4), 965–977.

Lin, S. J., Chen, J. Y., & Liao, Z. X. (2018). A EMD-BP integrated model to forecast tourist number and applied to Jiuzhaigou. *Journal of Intelligent & Fuzzy Systems, 34*(2), 1045–1052.

Liu, X., Liu, A., Chen, J. L., & Li, G. (2023). Impact of decomposition on time series bagging forecasting performance. *Tourism Management, 97*, Article 104725.

Liu, Y., Feng, G., Chin, K. S., Sun, S., & Wang, S. (2022). Daily tourism demand forecasting: The impact of complex seasonal patterns and holiday effects. *Current Issues in Tourism*, 1–20.

Liu, Y. Y., Tseng, F. M., & Tseng, Y. H. (2018). Big Data analytics for forecasting tourism destination arrivals with the applied vector Autoregression model. *Technological Forecasting and Social Change, 130*, 123–134.

Ma, E., Liu, Y., Li, J., & Chen, S. (2016). Anticipating Chinese tourists arrivals in Australia: A time series analysis. *Tourism Management Perspectives, 17*, 50–58.

Mao, W., Zhu, H., Wu, H., Lu, Y., & Wang, H. (2023). Forecasting and trading credit default swap indices using a deep learning model integrating Merton and LSTMs. *Expert Systems with Applications, 213*, Article 119012.

Pan, B., & Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *Journal of Travel Research, 56*(7), 957–970.

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review, 33*, 1–39.

Seabra, C., Reis, P., & Abrantes, J. L. (2020). The influence of terrorism in tourism arrivals: A longitudinal approach in a Mediterranean country. *Annals of Tourism Research, 80*, Article 102811.

Shabri, A. (2015, February). A novel hybrid ensemble learning paradigm for tourism forecasting. In AIP Conference Proceedings (Vol. 1643, No. 1, pp. 192-200). American Institute of Physics.

Škare, M., Soriano, D. R., & Porada-Rochoń, M. (2021). Impact of COVID-19 on the travel and tourism industry. *Technological Forecasting and Social Change, 163*, Article 120469.

Silva, E. S., Hassani, H., Heravi, S., & Huang, X. (2019). Forecasting tourism demand with denoised neural networks. *Annals of Tourism Research, 74*, 134–154.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management, 29*(2), 203–220.

Song, H., Qiu, R. T., & Park, J. (2019). A review of research on tourism demand forecasting: Launching the Annals of Tourism Research Curated Collection on tourism demand forecasting. *Annals of Tourism Research, 75*, 338–362.

Sun, S., Wei, Y., Tsui, K. L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management, 70*, 1–10.

Sun, S., Li, M., Wang, S., & Zhang, C. (2022). Multi-step ahead tourism demand forecasting: The perspective of the learning using privileged information paradigm. *Expert Systems with Applications, 210*, Article 118502.

Timmermann, A. (2006). Forecast combinations. Handbook of economic forecasting. *Vol.* 1 (pp. 135-196).

Wen, Q., Gao, J., Song, X., Sun, L., Xu, H., & Zhu, S. (2019). RobustSTL: A robust seasonal-trend decomposition algorithm for long time series. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 5409-5416).

Wu, J., Li, M., Zhao, E., Sun, S., & Wang, S. (2023). Can multi-source heterogeneous data improve the forecasting performance of tourist arrivals amid COVID-19? Mixed-data sampling approach. *Tourism Management, 98*, Article 104759.

Wu, T. P., & Wu, H. C. (2019). Causality between European economic policy uncertainty and tourism using wavelet-based approaches. *Journal of Travel Research, 58*(8), 1347–1356.

Xie, G., Qian, Y., & Wang, S. (2020). A decomposition-ensemble approach for tourism forecasting. *Annals of Tourism Research, 81*, Article 102891.

Xing, G., Sun, S., Bi, D., Guo, J. E., & Wang, S. (2022). Seasonal and trend forecasting of tourist arrivals: An adaptive multiscale ensemble learning approach. *International Journal of Tourism Research, 24*(3), 425–442.

Yan, R., Liao, J., Yang, J., Sun, W., Nong, M., & Li, F. (2021). Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. *Expert Systems with Applications, 169*, Article 114513.

Yang, Y., & Zhang, H. (2019). Spatial-temporal forecasting of tourism demand. *Annals of Tourism Research, 75*, 106–119.

Yao, Y., & Cao, Y. (2020). A neural network enhanced hidden Markov model for tourism demand forecasting. *Applied Soft Computing, 94*, Article 106465.

Zhang, C., Jiang, F., Wang, S., & Sun, S. (2021). A new decomposition ensemble approach for tourism demand forecasting: Evidence from major source countries in Asia-Pacific region. *International Journal of Tourism Research, 23*(5), 832–845.

Zhang, C., Li, M., Sun, S., Tang, L., & Wang, S. (2022). Decomposition methods for tourism demand forecasting: A comparative study. *Journal of Travel Research, 61*(7), 1682–1699.

Zhang, C., & Tian, Y. X. (2022). Forecast daily tourist volumes during the epidemic period using COVID-19 data, search engine data and weather data. *Expert Systems with Applications, 210*, Article 118505.

Zhang, M., Li, J., Pan, B., & Zhang, G. (2018). Weekly hotel occupancy forecasting of a tourism destination. *Sustainability, 10*(12), 4351.

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.

Zheng, W., Huang, L., & Lin, Z. (2021). Multi-attraction, hourly tourism demand forecasting. *Annals of Tourism Research, 90*, Article 103271.