Annals of Tourism Research 90 (2021) 103277

Contents lists available at ScienceDirect

Annals of Tourism Research

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

Forecasting tourism demand: Developing a general nesting spatiotemporal model

Xiaoying Jiao, Jason Li Chen, Gang Li *

School of Hospitality and Tourism Management, University of Surrey Guildford, GU2 7XH, UK

ARTICLE INFO

Article history:
Received 15 April 2021
Received in revised form 8 July 2021
Accepted 9 July 2021
Available online xxxx

Keywords:
Tourism demand forecasting
Spatiotemporal model
SAC model
GNST model
Panel data

ABSTRACT

This study proposes a general nesting spatiotemporal (GNST) model in an effort to improve the accuracy of tourism demand forecasts. The proposed GNST model extends the general nesting spatial (GNS) model into a spatiotemporal form to account for the spatial and temporal effects of endogenous and exogenous variables as well as unobserved factors. As a general specification of spatiotemporal models, the proposed model provides high flexibility in modelling tourism demand. Based on a panel dataset containing quarterly inbound visitor arrivals to 26 European destinations, this empirical study demonstrates that the GNST model outperforms both its non-spatial counterparts and spatiotemporal benchmark models. This finding confirms that spatial and temporal exogenous interaction effects contribute to improved forecasting performance.

Introduction

Given the perishable nature of tourism services, accurate tourism forecasts are essential to helping tourism businesses and government bodies devise plans and manage resources efficiently. Tourism demand forecasting has been a popular research area since the 1970s. Numerous forecasting models have been developed with the aim of improving forecasting accuracy. Three major categories of models have been applied in the tourism forecasting literature, namely time series models, econometric models, and artificial intelligence (AI) models (Song & Li, 2008; Witt & Witt, 1995).

Time series models generate forecasts based on patterns such as cycles and trends identified in historical data. As reviewed by Song et al. (2019), basic time series models appearing in the tourism demand forecasting literature consist of naïve, autoregressive (AR), exponential smoothing (ETS), moving average (MA), and historical average (HA) models. Many widely used advanced time series models, such as the autoregressive integrated moving average (ARIMA) model, can be seen as extensions of basic models. Econometric models consider the effects of exogenous variables when modelling the endogenous variable. Based on relationships between tourism demand and its influential factors, modern econometric models such as the autoregressive distributed lag model (ADLM) and the error correction model (ECM) are common in tourism demand forecasting. In addition, AI methods, such as artificial neural networks, support vector machine, the rough sets approach, fuzzy time series, and grey series, have been applied in tourism demand forecasting as well (Jiao & Chen, 2019; Wu et al., 2017). Although each of the above methods has shown superior forecasting performance in certain empirical contexts, none has outperformed the others consistently. Therefore, scholars continue to explore new methodological developments to enhance tourism forecasting performance.

In most of the tourism forecasting literature, destinations are treated in isolation. However, in reality many destinations are connected to each other in various ways. For example, long-haul tourists tend to visit multiple destinations in Europe within a

* Corresponding author.
E-mail addresses: x.jiao@surrey.ac.uk, (X. Jiao), l.chen@surrey.ac.uk, (J.L. Chen), g.li@surrey.ac.uk. (G. Li).

https://doi.org/10.1016/j.annals.2021.103277
0160-7383/

single trip. From the supply perspective, factors such as productivity spillover, market access and joint promotion also explain the connection among neighbouring destinations, as demonstrated by Yang and Wong (2012). The spatial connections between cross-sectional units can be exploited if the panel data contain a spatial dimension such as the geographic locations of units. To account for these spatial connections among destinations, spatial econometric models are widely used especially in regional studies by explicitly incorporating spatial interactions into the model.

Many tourism demand modelling studies have confirmed the existence of spatial spillovers in tourist flows by using spatial econometric models (e.g., Majewska, 2015; Yang & Fik, 2014). However, applications in tourism demand forecasting are still underexplored. Only three studies have recently sought to exploit spatiotemporal dependence in data to bolster forecasting performance (Long et al., 2019; Yang & Zhang, 2019; Jiao et al., 2020). These spatiotemporal models generally outperform their non-spatial counterparts in tourism demand forecasting, exemplifying the benefits of accounting for spatial spillover effects to boost forecasting accuracy. Yet as newer forecasting methods, spatiotemporal econometric models remain somewhat underdeveloped; a complete range of spatial effects has not been captured through traditional econometric techniques. Given empirical evidence reflecting the promise of spatiotemporal models in improving the accuracy of tourism demand forecasting, this research endeavours to extend existing forms of this emerging family of methods by accounting for spatial and temporal effects of endogenous and exogenous variables as well as unobserved factors. It formulates the most general form of the spatiotemporal model which has never been applied in any field of forecasting yet. This study also investigates whether the newly developed spatiotemporal model can further improve the accuracy of tourism demand forecasts.

Literature review

Time series models versus econometric models

Despite a consensus that no single forecasting model outperforms others in all circumstances, the advantages of econometric models over time series models have been emphasized in various studies. Song and Li (2008) pointed out that one such advantage is the capability to examine causal relationships between tourism demand and its influencing factors. Econometric analysis can hence provide empirical implications related to interpreting tourism demand from an economic perspective (e.g., by using the tools of price and income elasticities of demand), thereby facilitating evaluations of tourism policies and business strategies (Song & Lin, 2010).

Econometric models have also demonstrated strong forecasting performance. In the general forecasting field, prior research has suggested that econometric models are superior to autoregressive and naïve benchmarks when dealing with nonfinancial series (Allen & Fildes, 2001). Fildes et al. (2011) provided evidence of the superiority of econometric approaches in forecasting by using several econometric models to forecast air traffic flows; these models were compared with time series models such as the autoregressive of order 3 (AR(3)), naïve, and exponential smoothing models. Findings indicated that the ADLM with the "world trade" variable consistently performed better than the two time series benchmarks.

Because relevant exogenous variables contain information about future trends in tourism demand, these variables are often considered as predictors in tourism demand forecasting (Pan & Yang, 2017). Peng et al. (2014) noted that as long as an association exists between tourism demand and exogenous factors, introducing such factors as explanatory variables can increase a model's explanatory power and reduce forecasting error. Therefore, besides practical advantages, many empirical studies have underscored econometric models' improved forecasting performance compared with time series models. For instance, Song et al. (2000)) found that ECM's forecasting performance was superior to simple time series models (e.g., the autoregressive moving average (ARMA) model and naïve method) in forecasting UK tourism demand. Song et al. (2011) combined the causal structural time series model (STSM) with explanatory variables and the time-varying parameter (TVP) model, which allows for time-varying estimation of explanatory variable parameters. The TVP-STSM was benchmarked against other models to forecast tourist arrivals to Hong Kong. Results revealed the TVP-STSMs superior performance over time series benchmark models including the basic STSM without explanatory variables, seasonal ARIMA (SARIMA), and naïve models. Hirashima et al. (2017) used both time series and econometric models to forecast tourism demand to Hawaii. They found that methods incorporating explanatory variables produced more accurate forecasts than either the monthly or quarterly AR model in nearly all cases.

As pointed out by Jiao and Chen (2018), time series models augmented with explanatory variables, which combine the advantages of time series and econometric approaches, have become popular in tourism demand forecasting studies. The autoregressive integrated moving average with exogenous input (ARIMAX) model, an extension of the ARIMA model with explanatory variables, has been widely used and compared with the ARIMA model in tourism demand forecasting. Because the two models have the same specification except for explanatory variables, the value of adding these variables in tourism demand forecasting has become more persuasive. Many empirical studies have shown the superiority of the ARIMAX model over ARIMA in tourism demand forecasting. Pan et al. (2012) adopted three ARMA models and their ARMAX counterparts with search volume data to examine the value of adding explanatory variables when forecasting demand for hotel rooms. Findings indicated that all three ARMAX models outperformed their ARMA counterparts in terms of the mean absolute percentage error (MAPE) and root mean square percentage error (RMSPE). Yang et al. (2014) applied an ARMAX model to forecast hotel demand. Overall, the ARMAX model, which included the volume of web visits as an explanatory variable, generated more accurate 4- and 8-week-ahead forecasts compared to ARMA as evidenced by MAPE and RMSPE values. Tsui et al. (2014) employed SARIMA and ARMAX models to forecast airport passenger traffic at the Hong Kong International Airport. Their results showed that the ARIMAX model surpassed the time series SARIMA model in long-term forecasting. Park et al. (2017) used a seasonal ARIMAX (SARIMAX) model to forecast Japanese tourists' demand for Korean destinations. SARIMAX models were found to exceed the SARIMA and Holt-Winter exponential smoothing models on the bases of mean squared error (MSE) and mean absolute error (MAE). Pan and Yang (2017) also applied the ARMAX model with search engine queries and website traffic data to forecast hotel occupancy. Their ARMAX model containing the two big data explanatory variables demonstrated better forecasting accuracy than an ARMA model with weekly dummies. Rodríguez (2017) used ARMAX models with Google Trends indicators to forecast tourist arrivals to the Balearic Islands, including the UK and Germany; consistent with the results of prior studies, incorporating search queries data appeared to enhance forecasting accuracy. Li et al. (2020) have compared the performance of the ARIMA and ARIMAX models with multisource big data as explanatory variables in forecasting visitor arrivals to Mount Siguniang, China. Again, the ARIMAX model was found to be superior to the ARIMA model as well as some other time series benchmarks including ETS and seasonal naïve models, especially in short-term forecasting.

Another way to accommodate big data variables in tourism demand forecasting is mixed frequency models, given the high frequency of big data and relatively low frequency of tourism demand variables and economic determinants. Various developments of the mixed data sampling (MIDAS) method have been applied to tourism demand forecasting (e.g., Bangwayo-Skeete & Skeete, 2015; Wen et al., 2020) With regard to forecasting performance, the MIDAS models consistently outperformed pure time series models including seasonal naïve, ETS and SARIMA models. Although the SARIMAX model failed to beat some MIDAS models, it still forecast more accurately than its time series counterpart, the SARIMA model, which further confirms that incorporation of exogeneous variables is likely to improve forecasting accuracy.

Apart from the above single-equation models, one stream of tourism forecasting studies extend the single-equation model by using system-of-equations models to capture the interdependency of multiple demand flows. Such models include the vector regressive (VAR) model and vector error correction model (VECM), which have been widely applied in the tourism forecasting literature (Gunter & Önder, 2016; Song & Witt, 2006). But as noted by Song and Li (2008), in many studies, the classical VAR model is outperformed by other modern econometric methods. To improve the forecasting performance of VAR models, the Bayesian VAR (BVAR), global VAR (GVAR) and Bayesian GVAR models have been developed and applied in tourism forecasting, which show improved forecasting performance than the traditional VAR model as well as some univariate time series models (Assaf et al., 2019; Cao et al., 2017; Wong et al., 2006).

Even so, a few empirical studies have showcased the superiority of certain time series models over econometric models. The causal STSM was developed from the basic STSM without explanatory variables to account for causal variables in model specification. Yet Kulendran and Witt (2003) and Turner and Witt (2001) found that the causal STSM produced less accurate forecasts than the basic STSM. Song et al. (2011) offered a possible explanation for this phenomenon, specifically that the coefficients of explanatory variables were treated time-invariant.

The aforementioned empirical studies offer compelling evidence of improved forecasting accuracy when incorporating explanatory variables into forecasting models with the same specifications but without explanatory variables—hence this study's proposition that integrating spatiotemporal dependence between explanatory variables and the dependent variable may boost the accuracy of tourism demand forecasting.

Spatial spillover and spatial heterogeneity

As stated by Fingleton and López-Bazo (2006), initial regional economic growth studies treated heterogeneous countries as isolated economies, thus neglecting interactions across spatial locations. Yet the endogenous growth theory, together with new economic geography models emphasized the interactions across agents which generate spillovers within an economic system (Döring & Schnellenbach, 2006). This spillover effect can be illustrated by the existence of spatial externalities across regions, which implies that regions are neither homogeneous nor independent (Yang & Fik, 2014).

To address this issue, economic studies have often accounted for spatial effects across regions through spatial error models or spatial lag models (Bernat, 1996; Rey & Montouri, 1999). The spatial spillover effect and spatial heterogeneity are two major spatial effects considered in most spatial analyses. Spatial spillover refers to spatial externalities generated by economic activities exerting indirect effects across regions (Yang & Wong, 2012). Spatial spillover in tourism represents the indirect impacts of a tourism destination on that of other destinations. Yang and Wong (2012) noted that tourism generally produces spillover unintentionally, which can either positively or negatively affect neighbouring destinations. Many tourism studies have identified spatial spillover effects on tourism growth (e.g., Capone & Boix, 2008; Li et al., 2016) and tourist flows (e.g., Majewska, 2015; Yang & Fik, 2014) across destinations.

Although studies have highlighted spatial dependence across neighbouring destinations, spatial heterogeneity should not be ignored. Spatial heterogeneity represents the uniqueness of different locations, notably the distinct patterns and dependence of observed and unobserved characteristics across locations. In the global estimation of a spatial model, spatial heterogeneity can be partially captured by including location-specific fixed or random effects. However, spatial spillover effects remain constant across locations in global estimation. To fully account for spatial heterogeneity, local spatial models can be estimated to allow for varying spatial spillover across locations (Jiao et al., 2020; Li et al., 2016).

Spatial methodology development

Model specification is important in spatial econometrics because each specification is accompanied by different interpretations and implications (Fingleton & López-Bazo, 2006). LeSage and Pace (2009) reviewed spatial models applied in spatial econometric studies and stated that models vary in their integration of spatial lags. Among spatial specifications, the spatial autoregressive (SAR) model has received the most attention (Lee & Yu, 2010). This model, as the simplest spatial specification, extends the traditional regression through the integration of spatial lags into the dependent variable. The SAR model can be expanded into a spatial autoregressive combined (SAC) model by integrating spatial lags in the dependent variable and error terms, thereby considering spatial interaction effects in the endogenous variable and unobserved factors. Through the integration of spatial lags into exogenous variables as well as the endogenous variable (i.e., the dependent variable) and error terms, the SAR model can be extended into a generic specification known as the general nesting spatial (GNS) model. The GNS model accounts for all possible spatial interaction effects (Elhorst, 2017), including endogenous interaction, exogenous interaction, and interaction effects among unobserved factors. Imposing different restrictions on this model enables special cases to be specified, such as the spatial error model (e.g., Baltagi et al., 2003; Fildes et al., 2011), spatial Durbin model (e.g., Kim et al., 2021; Mur & Angulo, 2006), and spatial Durbin error model (e.g., Han & Lee, 2013), among others.

Spatiotemporal dependence in the dependent variable

In earlier stages of spatial model development, spatial dependence was solely examined using cross-sectional data. However, panel data have received greater attention in recent forecasting research due to the added time dimension. Compared with cross-sectional data, panel data provide more information by integrating time series and cross-sectional data (Wen, Liu and Song, 2019). When both spatial lags and temporal lags are incorporated, a spatial model becomes a spatiotemporal model and can be used for forecasting based on panel data.

Similar to SAR, the spatiotemporal autoregressive model is the most common specification in empirical studies wherein spatial and temporal lags are integrated in the endogenous variable. Within tourism demand modelling, this specification assumes that a destination's tourism demand is influenced by tourism demand for neighbouring destinations. Only three empirical studies have adopted this approach to forecasting tourism demand thus far. Domestic city arrivals in China (Yang & Zhang, 2019; Wen et al., 2019) and international inbound tourist arrivals in European countries (Jiao et al., 2020) were forecasted respectively using spatiotemporal autoregressive model. Yang and Zhang (2019) considered spatial interactions among tourist arrivals across destinations but did not account for exogenous variables using the space-time autoregressive moving average (STARMA) model; Wen et al. (2019) employed the dynamic SAR model using global GDP as an explanatory variable but neglected its spatial or spatiotemporal effects. In addition, both studies generated forecasts based on the global model, assuming constant spatial interaction effects across all destinations. In other words, spatial heterogeneity was not fully considered. Jiao et al. (2020) extended the SAC model into a spatiotemporal specification by accounting for spatial, temporal, and spatiotemporal interactions in the endogenous variable and unobserved factors. Their work also represented the first attempt to apply a local spatiotemporal model in forecasting, which enables unique estimation of spillover effects for individual destinations. Spatial heterogeneity was captured accordingly.

Even with different specifications, the above three studies demonstrated spatiotemporal models' superior forecasting performance over non-spatial methods. Integrating spatial effects can thus improve tourism demand forecasting performance. However, as mentioned, no research has yet taken the spatial dependence between endogenous and exogenous variables into account. The next section details why exogenous or explanatory variables should be considered in the spatiotemporal context.

Spatiotemporal dependence between endogenous and exogenous variables

Motivated by the superior forecasting performance of econometric models with explanatory variables compared with their time series counterparts (e.g., ARIMAX vs. ARIMA), this study proposes the general nesting spatiotemporal (GNST) model to empirically examine whether including explanatory variables along spatial and temporal dimensions can enhance forecasting performance. As in most tourism demand forecasting research using econometric models, economic explanatory variables will be considered here. Cao et al. (2017) explained the economic interdependence of tourism demand, emphasising the link between a country's tourism industry and other countries' economic circumstances. Especially for destinations such as those in the Schengen area, interdependence can be understood as complementary or substitutive relations between destinations' economies. From an econometric perspective, Elhorst (2014) argued that the cost of ignoring spatial dependence in an independent variable is high because omitting relevant explanatory variables will cause other coefficients' estimators to be biased and inconsistent (LeSage and Pace, 2009). Another reason to consider spatial dependence between the dependent variable and explanatory variables is that many scholars have recommended using more general spatial models instead of improving upon an original model (Burridge, 1981; Manski, 1993). Manski (1993) identified three types of spatial interaction effects in a general spatial model and explained why this type of model should be favoured: (1) the endogenous effect, where the decision of a spatial unit depends on other spatial units; (2) the exogenous interaction effect, where the decision of a spatial unit depends on the independent explanatory variables of other spatial units; and (3) the correlated effect, which indicates that similar unobserved fluctuations produce similar data patterns. The specification of the proposed GNST model explicitly accounts for spatial dependence in the data along with the impacts of spatial spillover associated with explanatory variables and unobserved factors (Pijnenburg & Kholodilin, 2014).

Empirical studies have revealed spatial dependency in exogenous variables. Zhang et al. (2017) stated that some exogenous variables, such as travel time and weather conditions, strongly influence short-term passenger demand and thus demonstrate time dependencies and spatial dependencies. Yang and Fik (2014) examined spatial effects in regional tourism growth and found that four out of six spatial lagged explanatory variables were estimated to be statistically significant. Their findings offered insight into the cross-regional competition/agglomeration effects of these explanatory variables on tourism. However, no scholars have yet exploited explanatory variables' spatial dependence on the dependent variable in tourism demand forecasting. Motivated by enhanced forecasting performance upon adding explanatory variables, the economic interdependence of tourism demand, and the benefits of adding the exogenous interaction effect from an econometric perspective, this study extends the work by Jiao et al. (2020) by including the exogenous interaction effect in model specification; doing so enables the use of more information through "cross-learning" from dependent and explanatory variables as well as unobserved factors over space and time. To the best of our knowledge, the proposed GNST model is the most general and complete model form to have been developed in any forecasting domain.

Methodology

This section proposes a GNST model using global and local specifications. Its forecasting performance will be empirically benchmarked against the spatiotemporal autoregressive combined (STAC) model, where no exogenous interaction effects are considered, as well as five non-spatial models.

STAC model

As Jiao et al. (2020) proposed, the SAC model can be temporally extended into a STAC model through the incorporation of spatial, temporal, and spatiotemporal lags into the dependent variable and the error term. Therefore, a STAC model can be specified as follows:

$$
\begin{aligned}
Y_t &= \lambda W Y_t + \gamma Y_{t-1} + \rho W Y_{t-1} + \mu + S_t \\
S_t &= \phi W S_t + \psi S_{t-1} + \varepsilon_t
\end{aligned}
\tag{1}
$$

where Y_t denotes the value of the dependent variable at time t, in a vector form consisting of N spatial units in the sample; Y_{t-1} denotes the time-lagged dependent variable; and W is an $N \times N$ spatial weight matrix representing geographical relationships among locations in the sample. This study uses the k-nearest neighbours scheme to determine the spatial weight matrix by assigning 1 to neighbouring destinations and 0 otherwise. The optimal k is empirically determined based on in-sample MAPE, following Jiao et al. (2020). The distance between every destination pair is determined by the great-circle distance, measured by the latitude and longitude of each destination's capital. The spatial lag is incorporated via the multiplication of spatial weight matrix W; thus, WY_t denotes the endogenous spatial interaction effect, and WY_{t-1} denotes the endogenous spatiotemporal interaction effect. μ is an $N \times 1$ vector denoting location-specific fixed effects, and S_t denotes the error term of the model in a vector form. Spatial and temporal interaction effects among error terms are specified by WS_t and S_{t-1} respectively; ϵ_t is the disturbance term of S_t; and λ, γ, ρ, ϕ, and ψ are parameters to be estimated.

GNST model

Although the STAC model captures the endogenous interaction and interactions among unobserved factors, exogenous interaction effects are not considered. To fill this gap, the present study proposes a GNST model by extending the GNS model into spatial and temporal dimensions. The resultant model includes all spatial, temporal, and spatiotemporal interaction effects in the explanatory variables as well as the dependent variable and unobserved factors. The global GNST model developed in this study can be expressed as below:

$$
\begin{aligned}
Y_t &= \lambda W Y_t + \gamma Y_{t-1} + \rho W Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W X_t + \beta_4 W X_{t-1} + \mu + S_t \\
S_t &= \phi W S_t + \psi S_{t-1} + \epsilon_t
\end{aligned}
\tag{2}
$$

where X_t denotes an $N \times m$ matrix of explanatory variables in the model (N represents the number of locations in the sample, and m represents the number of explanatory variables); X_{t-1} denotes the vector of exogenous variables lagged by one time period; WX_t denotes spatially lagged explanatory variables; and WX_{t-1} denotes explanatory variables lagged in the spatial and temporal dimensions. Therefore, the GNST model represents the most general form of a spatiotemporal model where all spatial and temporal interactions are included in the dependent variable, explanatory variables, and unobserved factors. Specific variables used in this study are described in Section 0. The model can be estimated using the bias-corrected quasi-maximum likelihood method proposed by Lee and Yu (2010), which generates consistent estimation with properly cantered distributions.

Local model

Following a similar procedure as that adopted by Jiao et al. (2020), a global spatiotemporal model can be respecified as a local model to allow for unique specifications of W and estimated parameters specific to individual locations. For instance, the local GNST model can be written as

$$
\begin{aligned}
U(i) Y_t &= \lambda U(i) W Y_t + \gamma U(i) Y_{t-1} + \rho U(i) W Y_{t-1} + \beta_1 U(i) X_t + \beta_2 U(i) X_{t-1} + \beta_3 U(i) W X_t + \beta_4 U(i) W X_{t-1} + U(i)\mu + U(i) S_t \\
U(i) S_t &= \phi U(i) W S_t + \psi U(i) S_{t-1} + U(i)\varepsilon_t
\end{aligned}
\tag{3}
$$

Similar to W, $U(i)$ represents an $N \times N$ spatial weight matrix, with 1 assigned to the neighbouring destinations and 0 otherwise. $U(i)$ specifies a sub-sample of regions from the whole sample; it is regulated by the region-specific number of neighbours (known as the bandwidth). As such, in the local model, W identifies neighbouring regions within a sub-sample as determined by $U(i)$. Except for $U(i)$, other formulations in the local model are the same as in the global model.

Local models were calibrated first to determine $U(i)$ and W. Different from the global model, which uses the same specification of W for all locations in the sample, the bandwidth and W were each empirically selected for every focal location based on the in-sample MAPE. Afterwards, the two local models were estimated and then used to generate forecasts.

Thus, in this study, four spatial models including the STAC and GNST models in global and local specifications were used to forecast tourism demand in Europe. Their forecasting performance was then contrasted with each other and with five non-spatial counterparts.

Data description

This study used tourist arrivals to measure tourism demand, similar to most tourism demand forecasting studies (Song & Li, 2008). Europe is one of the best examples to illustrate spatial spillover effects, due to tourists' (especially those from long-haul markets) tendency of visiting multiple neighbouring countries in this region. Thus, tourism demand for a European destination is expected to be associated with the demand for its neighbouring destinations (Batista e Silva et al., 2018). Twenty-six European countries were selected as the sample in light of data availability. The sample spanned 2004 Q1–2018 Q4. Monthly tourist arrival data for all countries were collected from Eurostat and aggregated to a quarterly frequency. Given clear patterns of seasonality, seasonal differencing was performed to render the series more stationary. Fig. 1 demonstrates seasonally differenced quarterly tourist arrivals for the 26 chosen countries after log transformation. As shown, the series for most countries became relatively stationary, especially in the first 10-year period. The series became less stable for the last periods from 2015 to 2018. Thus, dummy variables were introduced to account for the impacts of one-off events in the last several periods in the estimation sample. Turning points occurred for most destinations in the first quarter of 2015, the second quarter of 2016, and the second quarter of 2017. These points presumably coincided with the terrorist attack in France in 2015 and in Turkey in 2016, which affected the tourism industry throughout Europe; and Brexit in 2017, respectively. Because arrival data were seasonally differenced, two observations were affected by a single incident. For instance, the terrorist attack in 2015 generated two outliers: (1) the difference in arrivals between the 2015Q1 and 2014Q1 and (2) that between 2016Q1 and 2015Q1. Four dummy variables were hence introduced to capture these effects: D1 equals to 1 if the differenced observations fell from 2014Q1 to 2015Q1; D2 equals to 1 if the differenced observations fell from 2015Q1 to 2016Q1; D3 equals to1 if the differenced observations fell from 2016Q2 to 2017Q2; D4 equals to 1 if the differenced observations fall from 2017Q2 to 2018Q2.

Regarding the model's explanatory variables, tourism product prices and consumers' income have been identified as prime influencing factors of demand and appear frequently in tourism demand research (Li et al., 2005; Lin et al., 2015; Song, Wong, & Chon, 2003). As such, this study included the income variable and price variable as explanatory variables in the model. Because tourists' countries of origin were not specified in arrival data, this study regarded the rest of the world as one source market for a destination country. The income variable was thus measured by global GDP minus the GDP of the focal destination (i.e., the GDP

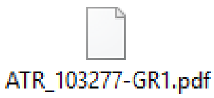

Fig. 1. Seasonally differenced quarterly tourist arrivals of individual countries from 2004 to 2018 (logarithmic scale).

Table 1
Global model specifications

STAC		GNST	
Number of neighbours	In-sample MAPE	Number of neighbours	In-sample MAPE
1	5.925	1	5.481
2	**5.907**	2	5.471
3	5.946	3	5.558
4	5.985	4	5.659
5	6.019	5	5.708
6	5.912	6	**5.365**
7	5.914	7	5.380
8	5.918	8	5.458
9	5.923	9	5.452

Note: Values in bold indicate the number of neighbours that yields the smallest in-sample MAPE.

of the rest of the world). The price variable was defined as $P_t = CPI_t/EX_t$, where CPI_t is the consumer price index of the destination country at time t (base year = 2010), and EX_t represents the exchange rate between the destination country's currency and the US dollar at time t. Spatial effects of tourism income were not considered in this study, as all source markets for a destination were combined into one; only GDP_t and time-lagged GDP_{t-1} were considered in the modelling process. Quarterly GDP, quarterly CPI, and exchange rate data for individual destinations were obtained from the International Monetary Fund.

Empirical results

Model calibration process

To pinpoint optimal specifications of the number of neighbours in spatial models, as well as the optimal size of each sub-sample in the local spatial model, a model calibration process was conducted. Data from 2004 Q1 to 2014 Q4 were applied for model calibration and estimation. In the global model, the selected number of neighbours ranged from 1 to 9, determined by the in-sample MAPE of fitted values in line with Jiao et al. (2020). Table 1 summarises the specification process of the two global models respectively, including the in-sample MAPE of models with 1–9 neighbours. For the global STAC model, the spatial weight matrix with two neighbours yielded the smallest in-sample MAPE; the optimal number of neighbours was six for the global GNST model.

Table 2
Local model specifications.

Destination	STAC			GNST		
	Bandwidth	K	In-sample MAPE	Bandwidth	K	In-sample MAPE
Austria	9	8	5.323	9	7	5.067
Belgium	9	8	3.480	9	8	3.268
Croatia	9	6	7.831	9	7	7.559
Cyprus	9	1	6.478	15	6	6.344
Czechia	10	2	5.144	10	2	5.019
Denmark	15	3	6.834	15	8	6.299
Estonia	9	1	5.119	9	1	4.708
Finland	24	4	5.589	9	7	5.267
Germany	17	1	4.265	16	6	4.034
Greece	9	6	7.897	9	7	7.192
Hungary	10	7	4.344	9	1	3.856
Iceland	10	8	8.887	9	5	8.777
Italy	9	1	4.196	14	7	3.582
Latvia	25	4	7.585	9	4	6.385
Lithuania	25	5	7.053	25	5	6.866
Luxembourg	9	5	5.910	25	3	6.307
Malta	25	2	6.563	9	8	5.553
Netherlands	9	6	5.695	9	1	5.030
Poland	24	4	4.490	25	5	4.116
Portugal	12	5	5.795	11	3	5.195
Romania	25	5	5.465	9	7	5.280
Slovakia	9	4	7.335	9	4	6.764
Slovenia	15	6	5.228	9	7	5.027
Spain	9	1	4.308	9	8	4.112
Sweden	9	1	5.618	9	4	5.321
UK	9	6	4.735	9	2	4.519

For the local model, as explained in Section 0, model calibration was conducted to determine the bandwidth, which controlled the size of the sub-sample for each focal country. The selected bandwidth ranged from 10 to 26 destinations (i.e., the whole sample including the focal destination). The number of neighbours, which defined the spatial weight matrix, was selected simultaneously. Thus, the model for each individual country was estimated as 17 (selections of bandwidth) by 9 (selections of the number of neighbours) totalling 153 times, and the group of selections that yielded the smallest in-sample MAPE was used in the local model specification for forecasting. Table 2 displays local model specifications for the STAC and GNST models. As shown in Table 2, the bandwidth and the number of neighbours determined through the calibration process for each destination varied much from each other, which further confirms the importance of local estimation. Spatial heterogeneity is reflected by a different sub-sample size and a different number of neighbours identified for each destination.

Forecasting

Following the process of model calibration, out-of-sample one-, two-, three-, four-, and eight-step-ahead forecasts were conducted using the four spatial models with the specifications determined above (i.e., STAC and GNST models with global and local specifications) and five benchmark models. Benchmark models included the SARIMA, ETS, and seasonal naïve (SNAIVE) models and an ADLM for individual countries as well as a panel ADLM. Similar to spatial models, the panel ADLM pooled individual countries' tourist arrivals into a single panel for estimation and forecasting. Country-specific fixed effects were incorporated into the panel ADLM to reflect the uniqueness of each country, as in the spatial models, in which fixed effects were used to capture spatial heterogeneity. Thus, the fixed-effects model is generally not as restrictive as the pooled ADLM (i.e., it allows for country-specific intercepts) but capitalizes on the advantage of the panel dimension (Arkadievich Kholodilin et al., 2008). The ADLM model was estimated individually for each destination, as with the other three time series benchmark models. The specifications of explanatory variables and lags of the dependent and independent variables were consistent with the STAC and GNST models for comparability purposes. As such, differences in forecasting performance could be attributed to spatial effects captured by the proposed GNST models. Forecasting accuracy was measured by the MAPE, mean absolute scaled error (MASE) and root mean squared error (RMSE). MAPE and RMSE have been widely used in the forecasting literature (e.g., Song & Witt, 2006; Volchek et al., 2019; Wen et al., 2020), whereas MASE has been proposed to measure forecasting accuracy by Hyndman (2006). MASE is a scale-free error metric which compares the out-sample actual forecasts' mean absolute error (MAE) against the one-step-ahead in-sample MAE of the seasonal naïve method (Chen et al., 2019). The equations for calculating MAPE, MASE and RMSE are as follows:

$$MAPE = \frac{1}{J}\sum_{j=1}^{J}\left|\frac{Y_j - F_j}{Y_j}\right| \tag{4}$$

$$MASE = \frac{\frac{1}{J}\sum_j \left|Y_j - F_j\right|}{\frac{1}{T-4}\sum_{t=5}^{T}\left|Y_t - Y_{t-4}\right|} \tag{5}$$

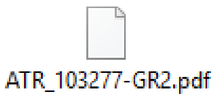

Fig. 2. Average MAPE of different forecasting models by forecasting horizon. Note: forecasting accuracy measures are averaged across the 26 destinations.

$$RMSE = \sqrt{\frac{\sum_j \left(Y_j - F_j\right)^2}{J}} \tag{6}$$

where Y_j and F_j are the actual value and forecast value, respectively; T is the length of the training data set.

The Diebold–Mariano (DM) test is further applied to evaluate whether the forecasting performance of different models are significantly different (Diebold & Mariano, 2002). To compare the forecasting performance between Model 1 and Model 2, the DM test is defined as follows (Sun et al., 2019):

$$DM = \frac{\overline{g}}{\sqrt{\hat{V}_{\overline{g}}/N}} \tag{7}$$

where $\overline{g} = \left(\sum_{t=1}^{N} g_t\right)/N$; $g_t = \sum_{t=1}^{N}(Y_t - F_{1,\,t})^2 - \sum_{t=1}^{N}(Y_t - F_{2,\,t})^2$ and $\hat{V}_{\overline{g}} = \gamma_0 + 2\sum_{l-1}^{\infty}\gamma_l(\gamma_l = COV(g_t, g_{t-l}))$. Y_t denotes the actual observation at time t. $F_{1,\,t}$ and $F_{2,\,t}$ are the forecasting values of Model 1 and Model 2, respectively, at time t. The ensuing discussion of forecasting results focuses on three dimensions: spatial models versus non-spatial models, global models versus local models, and spatial models with explanatory variables (i.e., the GNST model) versus spatial models without explanatory variables (i.e., the STAC model).

Spatial models versus non-spatial models

The average forecasting performance of all nine models is plotted in Figs. 2 and 3. Table 3 presents the average forecasting performance with rankings and Diebold–Mariano (DM) test statistics (Diebold & Mariano, 2002). Overall, the global and local GNST models both exhibited smaller forecasting errors than the other spatial and non-spatial models. In most cases, the four spatial models outperformed the four individual non-spatial benchmark models (i.e., SARIMA, ETS, SNAIVE, and ADLM). Although the one- and two-step-ahead forecasts generated by the SNAIVE model were more accurate than the two STAC models in terms of MAPE and MASE, the superiority of the four spatial models became clearer as the forecasting horizon extended beyond two steps ahead. The panel ADLM with fixed effects outperformed individual ADLM models across all horizons, indicating that pooling destinations within a geographical region (Europe in this study) may improve forecasting results. The superiority of the panel model with fixed effects against individual models was consistent with the results of Arkadievich Kholodilin et al. (2008), which showed that a panel ordinary least squares (OLS) model with fixed effects generated more accurate forecasts for every horizon (one to five steps ahead) than individual OLS when forecasting the GDP of German Länder. As Wen et al. (2019) pointed

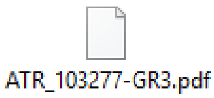

Fig. 3. Average MASE of different forecasting models by forecasting horizon. Note: forecasting accuracy measures are averaged across the 26 destinations.

Table 3
Average performance of different forecasting models.

Horizon	Measure	Global STAC	Global GNST	Local STAC	Local GNST	SARIMA	ETS	SNAIVE	ADLM	ADLM Panel
1 step	MAPE	0.158 (7)	0.144 (2)	0.151 (4)	0.141 (1)	0.159 (8)	0.155 (6)	0.145 (3)	0.168 (9)	0.151 (5)
	MASE	2.504 (8)	2.183 (3)	2.357 (5)	2.131 (1)	2.402 (6)	2.411 (7)	2.206 (4)	2.688 (9)	2.177 (2)
	DM	5.422***	2.448**	4.504***	–	2.888***	1.814*	0.552	4.707***	1.675*
2 steps	MAPE	0.13 (6)	0.118 (2)	0.127 (4)	0.117 (1)	0.137 (7)	0.14 (8)	0.13 (5)	0.144 (9)	0.122 (3)
	MASE	2.346 (7)	2.083 (3)	2.232 (5)	2.026 (2)	2.287 (6)	2.435 (8)	2.197 (4)	2.527 (9)	2.013 (1)
	DM	4.002***	0.971	3.211***	–	2.998***	2.661**	2.46**	5.166***	1.4
3 steps	MAPE	0.126 (3)	0.116 (1)	0.129 (5)	0.116 (2)	0.139 (9)	0.138 (8)	0.134 (6)	0.138 (7)	0.126 (4)
	MASE	2.09 (4)	1.946 (1)	2.14 (5)	1.949 (2)	2.259 (6)	2.279 (8)	2.28 (9)	2.277 (7)	2.08 (3)
	DM	2.847***	−0.303	3.528***	–	3.009***	2.429**	3.781***	5.699***	3.763***
4 steps	MAPE	0.13 (4)	0.126 (2)	0.132 (5)	0.124 (1)	0.144 (7)	0.147 (9)	0.137 (6)	0.145 (8)	0.13 (3)
	MASE	2.114 (3)	2.057 (2)	2.151 (5)	2.03 (1)	2.308 (7)	2.424 (9)	2.28 (6)	2.364 (8)	2.121 (4)
	DM	1.811*	1.466	2.202**	–	2.826***	3.543***	2.878***	5.592***	2.244**
8 steps	MAPE	0.183 (4)	0.18 (3)	0.194 (5)	0.177 (2)	0.219 (9)	0.213 (8)	0.199 (6)	0.203 (7)	0.17 (1)
	MASE	3.025 (4)	2.951 (3)	3.21 (5)	2.9 (2)	3.55 (8)	3.648 (9)	3.464 (7)	3.307 (6)	2.831 (1)
	DM	0.647	1.145	1.574	–	3.161***	2.09**	1.561	2.259**	−1.436

Note: Forecasting accuracy measures are averaged across the 26 destinations. Values in brackets represent rankings among the nine models. The DM test was conducted between the local GNST model and the other eight models. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

out, including fixed effects in a panel model can explain most of the heterogeneity for destinations with similar features. Thus, although the individual ADLM estimated for every destination respectively allowed for unique coefficients compared with the panel ADLM, country-specific fixed effects included in the panel ADLM model could capture the uniqueness of destinations as well, especially for countries with similar features. The global GNST model extended the panel ADLM by adding spatial effects to the dependent variable, explanatory variables, and error terms. The global GNST model surpassed the panel ADLM in most cases in one- to four-step-ahead forecasting, confirming that integrating spatial effects can enhance forecasting accuracy as shown in prior research (Wen et al., 2019). The local GNST model was more comparable with individual ADLMs, as both models were estimated for every country to retrieve country-specific coefficients using the same specifications of explanatory variables and lags. Unsurprisingly, upon incorporating spatial effects, the local GNST model was highly superior across all horizons compared with individual ADLMs. Furthermore, whereas individual ADLMs produced single-country estimates, the local GNST model pooled neighbouring countries' data into a country-specific sub-sample (determined by bandwidth) and included spatial effects among neighbouring countries. Panel advantages and heterogeneity were thus retained. Overall, the local GNST model outperformed the panel ADLM in shorter-term forecasts (one to four steps ahead).

Destination-level forecasting performance was examined as well in order to evaluate the model performance in forecasting each destination's arrivals. To develop a crisper understanding of model performance on each horizon, the nine models were ranked from most to least accurate based on MAPE, MASE and RMSE across all horizons for every destination. Results were generally consistent across horizons: for more than half of the destinations, the global and local GNST models both generated the most accurate and second most accurate forecasts among the seven models based on MAPE. The superiority of the two spatial models with explanatory variables became more apparent as the forecasting horizon extended. Due to space limitations, average forecasting performance rankings for the 26 countries are summarised in Table 4. Findings aligned with the average MAPE and MASE. Ultimately, the local GNST model performed best in short-term forecasts (except for three-step-ahead forecasts).

Table 4
Average performance rankings among the 26 countries of different models.

Horizon	Measure	Global	Global Ex	Local	Local Ex	SARIMA	ETS	Snaive	ADLM	ADLM Panel
1 step	MAPE	6.58	3.58	4.73	**2.96**	6.50	5.65	3.58	6.92	4.46
	MASE	6.81	3.62	5.04	**3.12**	5.85	5.46	3.88	7.50	3.69
	RMSE	7.15	3.85	5.19	**3.23**	5.65	4.15	4.00	7.88	3.85
2 steps	MAPE	5.69	3.12	4.81	**2.46**	6.58	6.65	5.00	6.46	4.19
	MASE	6.42	3.42	4.88	**2.73**	5.77	6.50	4.81	6.73	3.69
	RMSE	7.04	3.81	5.85	**3.27**	4.85	4.46	5.00	7.42	**3.27**
3 steps	MAPE	5.04	**2.31**	5.31	2.42	6.46	6.85	6.31	5.42	4.85
	MASE	4.96	2.62	5.27	**2.46**	5.92	6.38	6.73	5.81	4.81
	RMSE	5.69	3.38	5.73	**3.00**	4.27	3.65	7.12	7.27	4.85
4 steps	MAPE	4.81	3.35	4.73	**2.73**	5.88	7.27	6.23	5.38	4.58
	MASE	4.27	3.50	4.46	**2.92**	5.50	7.19	6.62	5.88	4.62
	RMSE	6.42	4.04	6.62	**3.58**	3.58	4.27	5.81	6.96	3.69
8 steps	MAPE	4.42	3.69	5.35	3.65	5.96	7.04	6.96	4.50	**3.38**
	MASE	4.42	3.88	5.19	3.77	5.77	7.08	6.73	4.62	**3.50**
	RMSE	5.00	4.73	5.62	4.88	5.27	5.96	5.81	4.73	**2.96**

Note: the eight models are ranked and the rankings are averaged across the 26 countries in each horizon. Values in bold indicate the best forecasting performance.

DM tests were carried out to evaluate the difference in predictive accuracy between forecasting models. The local spatial model with explanatory variables was compared with the other eight models; see findings in Table 3. Compared with the four non-spatial individual models, the superiority of the local GNST model was significant in most cases (except for one- and eight-step-ahead forecasts compared with the SNAIVE model). Compared with the panel ADLM, the superiority of the local GNST model was significant in short-term forecasts (except for two-step-ahead forecasts).

Global models versus local models

Although local models did not outperform global models in all cases, the local STAC model produced more accurate short-term (one- to two-step-ahead) forecasts than its global counterpart per MAPE and MASE (Table 3). This pattern coincides with results from Jiao et al. (2020), in which a fuller specification of spatial heterogeneity can improve forecasting accuracy. Upon including explanatory variables, the local GNST model surpassed its global counterpart in most cases, mainly due to the local model's ability to account for spatial heterogeneity; in other words, it enabled country-specific parameter estimations of independent variables' coefficients to further distinguish a single destination's uniqueness. Thus, the local GNST model was notably superior to the global GNST model.

With respect to destination-level performance, the local model outperformed the global model for more than 60% of destinations (62% for the STAC model and 65% for the GNST model). As indicated in Table 4, average forecasting performance rankings were consistent with the average performance. Again, the local GNST model outranked the global GNST model in most cases. Yet when explanatory variables were not included, the local model ranked higher than the global model in short-term (one- and two-step-ahead) forecasts. Finally, DM test statistics showed that the difference between the local and global GNST models was only significant when forecasting one step ahead.

STAC models versus GNST models

One aim of this study was to investigate the effectiveness of including exogenous interactions on space and time dimensions in tourism demand forecasting. Without integrating spatial effects, the individual ADLM with explanatory variables generated worse forecasts than the three time series models. Conversely, after pooling individual countries' data into a panel with fixed effects, the panel model with explanatory variables outperformed the time series benchmarks. Upon considering spatial effects, the superiority of including explanatory variables on both spatial and temporal dimensions was even more evident: the global and local GNST models uniformly outperformed the STAC models (per MAPE and MASE). Thus, this study's hypothesis that adding explanatory variables to STAC models can improve forecasting accuracy was confirmed.

With regard to destination-level performance, for roughly 89% to100% of destinations, the spatial models with explanatory variables outperformed corresponding spatial models without explanatory variables. As the forecasting horizon extended beyond three steps, these percentages declined slightly to around 70%. Table 4 presents the average forecasting performance rankings. The superiority of spatial models including spatial and temporal interactions between the explanatory variables and dependent variables was reflected by better rankings for the global and local GNST models compared with the STAC models. Therefore, for most individual destinations, forecasting results improved at the country level as well after incorporating exogenous interactions into the spatial and temporal dimensions. Table 3 also details the DM test, comparing the forecasting performance of the local GNST model with other models. The local GNST model significantly exceeded the local STAC model in generating one- to four-step-ahead forecasts.

Conclusion

For the first time in the forecasting literature, this study has proposed a general specification of the spatiotemporal model. In terms of the modelling methodology, this study brings spatial and time series models together more fully. From a spatial modelling standpoint, the GNST model represents a complete temporal extension of the GNS model. The proposed model is a full spatial extension of the ARMA model from a time series modelling point of view. It also extends Jiao et al.'s (2020) STAC model by including spatial and temporal exogenous interaction effects in addition to the endogenous interaction and interactions among unobserved factors. Therefore, this study makes a valuable contribution to general methodological development in forecasting beyond the tourism forecasting literature. The built-in parameter optimisation process accompanying this general form of spatiotemporal econometric modelling enables forecasters to identify an optimal model specification from many potential parameter combinations based on the characteristics of available empirical data. This model specification is therefore likely to generate the most desirable forecasting results.

Empirically, this study examined the GNST model's forecasting performance. Jiao et al. (2020) identified the contributions of the endogenous interaction and interactions among unobserved factors to the spatiotemporal model's superior forecasting performance. The current study revealed that the proposed method further improves forecasting accuracy; that is, spatial and temporal exogenous interaction effects can enhance forecasting performance even more. This study also provides some useful practical implications for destination management organisations and the tourism industry. To begin with, the superior forecasting performance of spatial models against non-spatial models indicates that destinations should consider the spatial and temporal effects of neighbouring destinations' tourism demand and relevant economic indicators when making predictions of future tourist arrivals. In practice, where significant spatiotemporal spillover effects are identified by the spatiotemporal model, destinations should closely monitor and forecast the tourism demand and economic situations of neighbouring destinations regularly in order to estimate the impacts on their own tourism demand. Based on the spillover effects, competitive or collaborative strategies can be developed, such as joint promotion and transport connections, to facilitate inbound tourism demand especially in a closely interrelated region such as Europe. Furthermore, when one-off events such as terrorist attacks and Olympic games occur in a particular destination, neighbouring destinations should also estimate the level of spillover effects generated from those events to adjust the forecasts of tourist arrivals in their destinations.

Future applications of the proposed family of GNST models, in both its global and local forms, are encouraged within and beyond tourism forecasting to further assess their forecasting performance in diverse empirical contexts. For example, the GNST model can be applied to multiple regions or cities within a country or in a domestic tourism context. Some destination dummies can be considered as well, such as a dummy variable capturing whether a destination is within the Schengen area or not, to reflect different characteristics of the destinations. In addition, as a general family of models, the GNST model can be further reduced to its special cases (e.g., the STAC and GNS models) by using a model selection process based on criteria such as in-sample MAPE, which can be applied to the global and local model estimation to potentially boost the forecasting performance even further.

This study has several limitations. First of all, factors that are specific to source markets are not fully considered in this study as the proposed GNST model is based on a destination-specific spatial weight matrix. For future research, tourist flows between origin-destination pairs can be used to construct the spatial weight matrix. Therefore, the model can be extended to account for the spatiotemporal relationship between origin-destination pairs if relevant data are available. Secondly, the choice of explanatory variables is worth further exploration in future studies especially those aiming to examine the level of spillover effects and demand elasticities. Moreover, as the application of big data has increasingly attracted attention in recent studies in tourism demand forecasting (e.g., Li et al., 2020), further exploiting the spatiotemporal association of big data variables (e.g., Google trends and hotel booking data) with tourism demand can be another direction for future studies. Finally, due to the unavailability of data, the impact of the Covid-19 pandemic has not been explored in this study. Future research can address this issue by using scenario forecasting as in Wu et al. (2021) and Zhang et al. (2021).

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.

References

Allen, P. G., & Fildes, R. (2001). Econometric forecasting. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners* (pp. 303–362). Springer US. https://doi.org/10.1007/978-0-306-47630-3_15.

Arkadievich Kholodilin, K., Siliverstovs, B., & Kooths, S. (2008). A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spat. Econ. Anal.*, *3*(2), 195–207.

Assaf, A. G., Li, G., Song, H., & Tsionas, M. G. (2019). Modelling and forecasting regional tourism demand using the Bayesian global vector autoregressive (BGVAR) model. *J. Travel Res.*, *58*(3), 383–397.

Baltagi, B. H., Song, S. H., & Koh, W. (2003). Testing panel data regression models with spatial error correlation. *J. Econ.*, *117*(1), 123–150.

Bangwayo-Skeete, P. F., & Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tour. Manag.*, *46*, 454–464.

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tour. Manag.*, *68*, 101–115. https://doi.org/10.1016/j.tourman.2018.02.020.

Bernat, G. A. (1996). Does manufacturing matter? A spatial econometric view of Kaldor's laws*. *J. Reg. Sci.*, *36*(3), 463–477. https://doi.org/10.1111/j.1467-9787.1996.tb01112.x.

Burridge, P. (1981). Testing for a common factor in a spatial autoregression model. *Environ Plan A*, *13*(7), 795–800.

Cao, Z., Li, G., & Song, H. (2017). Modelling the interdependence of tourism demand: The global vector autoregressive approach. *Ann. Tour. Res.*, *67*, 1–13. https://doi.org/10.1016/j.annals.2017.07.019.

Capone, F., & Boix, R. (2008). Sources of growth and competitiveness of local tourist production systems: An application to Italy (1991–2001). *Ann. Reg. Sci.*, *42*(1), 209–224. https://doi.org/10.1007/s00168-007-0133-7.

Chen, J. L., Li, G., Wu, D. C., & Shen, S. (2019). Forecasting seasonal tourism demand using a multiseries structural time series method. *J. Travel Res.*, *58*(1), 92–103. https://doi.org/10.1177/0047287517737191.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *J. Bus. Econ. Stat.*, *20*(1), 134–144. https://doi.org/10.1198/073500102753410444.

Döring, T., & Schnellenbach, J. (2006). What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature. *Reg. Stud.*, *40*(03), 375–395.

Elhorst, J. P. (2014). Dynamic spatial panels: Models, methods and inferences. *Spatial Econometrics* (pp. 95–119). Springer.

Zhang, J., Zheng, Y., & Qi, D. (2017). Deep spatio-temporal residual networks for citywide crowd flows prediction. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 1655–1661http://arxiv.org/abs/1610.00081.

Elhorst, J. P. (2017) Spatial Panel Data Analysis. In S. Shekhar, H. Xiong, & X. Zhou (Eds.), Encyclopedia of GIS (2nd ed.; pp. 2050–2058). Springer International Publishing, Cham, Switzerland.

Fildes, R., Wei, Y., & Ismail, S. (2011). Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *Int. J. Forecast.*, *27*(3), 902–922. https://doi.org/10.1016/j.ijforecast.2009.06.002.

Fingleton, B., & López-Bazo, E. (2006). Empirical growth models with spatial effects*. *Pap. Reg. Sci.*, *85*(2), 177–198. https://doi.org/10.1111/j.1435-5957.2006.00074.x.

Gunter, U., & Önder, I. (2016). Forecasting city arrivals with Google analytics. *Ann. Tour. Res.*, *61*, 199–212.

Han, X., & Lee, L. (2013). Bayesian estimation and model selection for spatial Durbin error model with finite distributed lags. *Reg. Sci. Urban Econ.*, *43*(5), 816–837. https://doi.org/10.1016/j.regsciurbeco.2013.04.006.

Hirashima, A., Jones, J., Bonham, C. S., & Fuleky, P. (2017). Forecasting in a mixed up world: Nowcasting Hawaii tourism. *Ann. Tour. Res.*, *63*, 191–202. https://doi.org/10.1016/j.annals.2017.01.007.

Hyndman, R. J. (2006). Another look at forecast-accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting*, *4*(4), 43–46.

Jiao, E. X., & Chen, J. L. (2019). Tourism forecasting: a review of methodological developments over the last decade. *Tour. Econ.*, *25*(3), 469–492. https://doi.org/10.1177/1354816618812588.

Jiao, X., Li, G., & Chen, J. L. (2020). Forecasting international tourism demand: a local spatiotemporal model. *Ann. Tour. Res.*, *83*, Article 102937. https://doi.org/10.1016/j.annals.2020.102937.
Kim, Y. R., Williams, A. M., Park, S., & Chen, J. L. (2021). Spatial spillovers of agglomeration economies and productivity in the tourism industry: the case of the UK. *Tour. Manag.*, *82*, Article 104201.
Kulendran, N., & Witt, S. F. (2003). Forecasting the demand for international business tourism. *J. Travel Res.*, *41*(3), 265–271. https://doi.org/10.1177/0047287502239034.
Lee, L., & Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *J. Econ.*, *154*(2), 165–185. https://doi.org/10.1016/j.jeconom.2009.08.001.
LeSage, J., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Chapman and Hall/CRC.
Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modelling and forecasting. *J. Travel Res.*, *44*(1), 82–99.
Li, H., Chen, J. L., Li, G., & Goh, C. (2016). Tourism and regional income inequality: evidence from China. *Ann. Tour. Res.*, *58*, 81–99. https://doi.org/10.1016/j.annals.2016.02.001.
Li, H., Hu, M., & Li, G. (2020). Forecasting tourism demand with multisource big data. *Ann. Tour. Res.*, *83*, Article 102912. https://doi.org/10.1016/j.annals.2020.102912.
Lin, V. S., Liu, A., & Song, H. (2015). Modelling and forecasting Chinese outbound tourism: an econometric approach. *J. Travel Tour. Mark.*, *32*(1–2), 34–49. https://doi.org/10.1080/10548408.2014.986011.
Long, W., Liu, C., & Song, H. (2019). Pooling in tourism demand forecasting. *J. Travel Res.*, *58*(7), 1161–1174. https://doi.org/10.1177/0047287518800390.
Majewska, J. (2015). Inter-regional agglomeration effects in tourism in Poland. *Tour. Geogr.*, *17*(3), 408–436. https://doi.org/10.1080/14616688.2014.997279.
Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *Rev. Econ. Stud.*, *60*(3), 531–542. https://doi.org/10.2307/2298123.
Mur, J., & Angulo, A. (2006). The spatial Durbin model and the common factor tests. *Spat. Econ. Anal.*, *1*(2), 207–226. https://doi.org/10.1080/17421770601009841.
Pan, B., Chenguang Wu, D., & Song, H. (2012). Forecasting hotel room demand using search engine data. *J. Hosp. Tour. Technol.*, *3*(3), 196–210. https://doi.org/10.1108/17579881211264486.
Pan, B., & Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *J. Travel Res.*, *56*(7), 957–970. https://doi.org/10.1177/0047287516669050.
Park, S., Lee, J., & Song, W. (2017). Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data. *J. Travel Tour. Mark.*, *34*(3), 357–368. https://doi.org/10.1080/10548408.2016.1170651.
Peng, B., Song, H., & Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tour. Manag.*, *45*, 181–193. https://doi.org/10.1016/j.tourman.2014.04.005.
Pijnenburg, K., & Kholodilin, K. A. (2014). Do regions with entrepreneurial neighbours perform better? A spatial econometric approach for German regions. *Reg. Stud.*, *48*(5), 866–882. https://doi.org/10.1080/00343404.2012.697143.
Rey, S. J., & Montouri, B. D. (1999). US regional income convergence: a spatial econometric perspective. *Reg. Stud.*, *33*(2), 143–156. https://doi.org/10.1080/00343409950122945.
Rodríguez, Ó. G. (2017). Forecasting tourism arrivals with an online search engine data: a study of the Balearic Islands. *PASOS. Revista de Turismo y Patrimonio Cultural*, *15*(4), 943–958.
Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—a review of recent research. *Tour. Manag.*, *29*(2), 203–220. https://doi.org/10.1016/j.tourman.2007.07.016.
Song, H., Li, G., Witt, S. F., & Athanasopoulos, G. (2011). Forecasting tourist arrivals using time-varying parameter structural time series models. *Int. J. Forecast.*, *27*(3), 855–869. https://doi.org/10.1016/j.ijforecast.2010.06.001.
Song, H., & Lin, S. (2010). Impacts of the financial and economic crisis on tourism in Asia. *J. Travel Res.*, *49*(1), 16–30. https://doi.org/10.1177/0047287509353190.
Song, H., Qiu, R. T. R., & Park, J. (2019). A review of research on tourism demand forecasting: launching the annals of tourism research curated collection on tourism demand forecasting. *Ann. Tour. Res.*, *75*, 338–362. https://doi.org/10.1016/j.annals.2018.12.001.
Song, H., Romilly, P., & Liu, X. (2000). An empirical study of outbound tourism demand in the UK. *Appl. Econ.*, *32*(5), 611–624. https://doi.org/10.1080/000368400322516.
Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tour. Manag.*, *27*(2), 214–224.
Song, H., Wong, K. K. F., & Chon, K. K. S. (2003). Modelling and forecasting the demand for Hong Kong tourism. *Int. J. Hosp. Manag.*, *22*(4), 435–451. https://doi.org/10.1016/S0278-4319(03)00047-1.
Sun, S., Wei, Y., Tsui, K. L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tour. Manag.*, *70*, 1–10.
Tsui, W. H. K., Ozer Balli, H., Gilbey, A., & Gow, H. (2014). Forecasting of Hong Kong airport's passenger throughput. *Tour. Manag.*, *42*, 62–76. https://doi.org/10.1016/j.tourman.2013.10.008.
Turner, L. W., & Witt, S. F. (2001). Forecasting tourism using univariate and multivariate structural time series models. *Tour. Econ.*, *7*(2), 135–147. https://doi.org/10.5367/000000001101297775.
Volchek, K., Liu, A., Song, H., & Buhalis, D. (2019). Forecasting tourist arrivals at attractions: search engine empowered methodologies. *Tour. Econ.*, *25*(3), 425–447. https://doi.org/10.1177/1354816618811558.
Wen, L., Liu, C., Song, H., & Liu, H. (2020). Forecasting tourism demand with an improved mixed data sampling model. *J. Travel Res.*, Article 004728752090622. https://doi.org/10.1177/0047287520906220.
Witt, S. F., & Witt, C. A. (1995). Forecasting tourism demand: a review of empirical research. *Int. J. Forecast.*, *11*(3), 447–475. https://doi.org/10.1016/0169-2070(95)00591-7.
Wong, K. K., Song, H., & Chon, K. S. (2006). Bayesian models for tourism demand forecasting. *Tour. Manag.*, *27*(5), 773–780.
Wu, D. C., Cao, Z., Wen, L., & Song, H. (2021). Scenario forecasting for global tourism. *J. Hosp. Tour. Res.*, *45*(1), 28–51. https://doi.org/10.1177/1096348020919990.
Wu, D. C., Song, H., & Shen, S. (2017). New developments in tourism and hotel demand modelling and forecasting. *Int. J. Contemp. Hosp. Manag.*, *29*(1), 507–529. https://doi.org/10.1108/IJCHM-05-2015-0249.
Yang, Y., & Fik, T. (2014). Spatial effects in regional tourism growth. *Ann. Tour. Res.*, *46*, 144–162. https://doi.org/10.1016/j.annals.2014.03.007.
Yang, Y., Pan, B., & Song, H. (2014). Predicting hotel demand using destination marketing organization's web traffic data. *J. Travel Res.*, *53*(4), 433–447. https://doi.org/10.1177/0047287513500391.
Yang, Y., & Wong, K. K. F. (2012). A spatial econometric approach to model spillover effects in tourism flows. *J. Travel Res.*, *51*(6), 768–778. https://doi.org/10.1177/0047287512437855.
Yang, Y., & Zhang, H. (2019). Spatial-temporal forecasting of tourism demand. *Ann. Tour. Res.*, *75*, 106–119. https://doi.org/10.1016/j.annals.2018.12.024.
Zhang, H., Song, H., Wen, L., & Liu, C. (2021). Forecasting tourism recovery amid COVID-19. *Ann. Tour. Res.*, *87*, Article 103149. https://doi.org/10.1016/j.annals.2021.103149.

Xiaoying Jiao is a PhD student at the University of Surrey. Her research interests include tourism demand forecasting and applications of time series and econometric models.
Jason Li Chen is a senior lecturer at the University of Surrey. His research interests include tourism economics and applications of quantitative research methods.
Gang Li is a Professor of Tourism Economics at the University of Surrey. His research interests include economic analysis and forecasting of tourism demand.