ELSEVIER

Forecasting international tourism demand: a local spatiotemporal model

Xiaoying Jiao, Gang Li*, Jason Li Chen

[a] School of Hospitality & Tourism Management, University of Surrey, Guildford, Surrey GU2 7XH, UK
[b] School of Hospitality and Tourism Management, University of Surrey, Guildford GU2 7XH, UK

ABSTRACT

This study investigates whether tourism forecasting accuracy is improved by incorporating spatial dependence and spatial heterogeneity. One- to three-step-ahead forecasts of tourist arrivals were generated using global and local spatiotemporal autoregressive models for 37 European countries and the forecasting performance was compared with that of benchmark models including autoregressive moving average, exponential smoothing and Naïve 1 models. For all forecasting horizons, the two spatial models outperformed the non-spatial models. The superior forecasting performance of the local model suggests that the full reflection of spatial heterogeneity can improve the accuracy of tourism forecasting.

Introduction

With the rapid development of the tourism industry, tourism demand forecasting has become increasingly important for governments and business. According to the United Nations World Tourism Organization (UNWTO) (2019), growth in international tourism continues to outpace the global economy, which makes the tourism sector a global driving force for economic growth and development. Tourism demand forecasting has therefore attracted increasing attention in recent decades, especially in light of the perishable nature of tourism services (Li, Song, & Witt, 2005). More accurate forecasts would enable stakeholders to plan ahead and allocate resources more efficiently, and businesses could adjust their strategies to enhance their performance. To improve forecasting accuracy, new methods have been continually developed in the tourism demand forecasting field. Time series models and econometric models are most frequently used, and artificial intelligence (AI) models have started to gain popularity in the past decade (Jiao & Chen, 2019). Reviews of tourism demand forecasting models have identified a number of new trends, including the use of spatial models (Wu, Song, & Shen, 2017).

Spatial econometrics, which concerns the spatial interactions of economic units (Lee & Yu, 2010b), has recently become a popular research area and has been applied in many different fields of study. Because locations and destinations are highly relevant to the tourism industry, spatial analysis has been applied in some tourism studies to account for the spatial spillover effect among destinations or between origins and destinations (Yang & Wong, 2012). For example, given the Schengen visa arrangements in the European Union, long-haul tourists in particular tend to visit multiple destinations during a single trip to Europe. Previous tourism studies (e.g., Romão, Guerreiro, & Rodrigues, 2017; Romão & Nijkamp, 2019) have confirmed the existence of significant spatial dependency among European countries. This study attempts to examine if the incorporation of both spatial dependence and spatial heterogeneity can improve forecasting performance in 37 European countries.

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

Although the spatial effects on tourism growth and tourism flows have been analysed in many studies, only two have incorporated the spatial spillover effect into tourism demand forecasting. Long, Liu and Song (2019) and Yang and Zhang (2019) used spatio-temporal models to forecast domestic tourism demand in China and compared the forecasting performance with other benchmark models: Naïve 1, Naïve 2, ordinary least squares (OLS), autoregressive integrated moving average (ARIMA) and other models. The spatiotemporal spatial autoregressive (SAR) model used in these studies has the advantage of incorporating both spatial and temporal lags in a panel data model. Following these initial attempts in introducing spatiotemporal models into tourism forecasting, the present study aims to further develop the spatiotemporal model and examine the accuracy improvement in tourism forecasting. It should be noted that, although the empirical setting of this study is in a tourism context, its contributions go beyond the tourism forecasting literature, because to the best of our knowledge the models developed in this study have not been applied in any field of forecasting yet.

The structure of this paper is as follows. Section 0 presents a literature review on tourism demand forecasting and spatial econometrics. Section 0 explains the methodology. Section 0 discusses the empirical results and Section 0 concludes the study.

literature review

Tourism demand forecasting

Tourism demand forecasting is a longstanding stream of tourism research and has continuously attracted scholars' attention since the 1970s. Tourism demand modelling and forecasting models can be broadly classified into three categories: time series models, econometric models and AI models. Time series models identify patterns and trends from historical data to predict the future. Among the time series models, ARIMA and seasonal ARIMA models are most frequently used in tourism forecasting (Song & Li, 2008). They are commonly used as benchmark models along with other popular time series models (e.g., exponential smoothing and Naïve models) to examine the forecasting performance of newly developed models (Jiao & Chen, 2019). Unlike time series models, econometric models incorporate explanatory variables into a forecasting model. Commonly used economic variables associated with origin and destination countries include tourism prices and tourist income (e.g., Ayeh & Lin, 2011; Lee, 2011). Many advanced econometric models have been developed for tourism demand forecasting, such as the error correction model (ECM) (Kulendran & King, 1997), time-varying parameter (TVP) models (Song & Wong, 2003), almost ideal demand system (AIDS) model (Li, Song, & Witt, 2004) and global vector autoregressive (GVAR) model (Cao, Li, & Song, 2017). In addition, a few nonlinear models have been used in tourism forecasting, such as the smooth transition autoregressive (STAR) model (Saayman & Botha, 2015), the Markov switching VAR (MSVAR) model (Chaitip & Chaiboonsri, 2014) and the Markov switching dynamic regression (MSDR) model (Pan & Yang, 2017). Nonlinear models allow for nonlinear combinations of parameters in a way that the values of parameters can switch across different regimes, thus capturing the nonlinear pattern of tourism demand. Compared with time series and econometric models, the use of AI models in tourism demand forecasting is relatively new. As data volume is increasing and data characteristics are becoming more complex, AI models have gained popularity because they can capture complex relations and patterns in a large volume of data. Popular AI models used in tourism demand forecasting include artificial neural networks (ANN) (Chen, Lai, & Yeh, 2012), support vector machines (SVM) (Chen, Liang, Hong, & Gu, 2015) and fuzzy time series models (Chen, Ying, & Pan, 2010), along with some new deep learning approaches (Law, Li, Fong, & Han, 2019). However, without theoretical underpinning, AI models largely remain a 'black box', which limits their application (Gareta, Romeo, & Gil, 2006).

Spatial spillover

Spatial analysis begins with an introduction to the concept of spatial spillover. 'Spillover' is a frequently used economic term representing externalities generated by economic events or processes exerting indirect effects (Yang & Wong, 2012). The addition of the term 'spatial' takes geographical effects and interactions into account. Fingleton and López-Bazo (2006) stated that spatial spillover is related to spatial externalities across regions according to new economic geographies, and to endogenous growth models. Spatial spillover effects have been widely and frequently analysed in the fields of economics and economic geography, in which spatial dependence and regional interactions have been comprehensively discussed (Fingleton & López-Bazo, 2006; Huang, Zhou, Wang, Chang, & Ma, 2017). Spatial effects have also been widely considered in other fields in which location is highly relevant, including networking and transport infrastructures (Álvarez, Barbero, & Zofío, 2016; Condeço-Melhorado, Tillema, de Jong, & Koopal, 2014) capital inflow and outflow analysis (Miranda, Martínez-Ibañez, & Manjón-Antolín, 2017).

In the context of tourism, spatial spillover is the indirect effects that a destination's tourism industry exerts on the tourism growth and tourism flows of other destinations (Yang & Wong, 2012). Empirical studies applying spatial spillover analysis in relation to tourism mainly focus on two things: regional tourism growth and tourist flows to specific destinations (Yang & Fik, 2014). Research on regional tourism growth suggests that regional growth is significantly influenced by tourism development in both local and neighbouring regions (Capone & Boix, 2008). Ma, Hong, and Zhang (2015) confirmed that domestic tourism has a significantly positive spillover effect in China. Li, Chen, Li, and Goh (2016) examined how tourism development contributes to reducing regional income inequality, using global and local spatiotemporal models to capture spatial spillover effects across different regions in China. Zhang, Xu, and Zhuang (2011) discovered a significant neighbouring effect in the distributions of both international and domestic tourists in 299 cities in mainland China.

Spillover effects in tourist flows have also been widely discussed in tourism studies, although the focus has been on spatial interactions between origins and destinations. More recently, Yang and Wong (2012), followed by Yang and Fik (2014), made initial

attempts to analyse tourist flows and discuss spatial spillover effects across destinations. Majewska (2015) measured the inter-regional effects of spatial agglomeration by the spatial dependence and heterogeneity of tourism flows and supply in neighbouring tourist attractions. That is, the spatial dependence among destinations with regard to tourist flows means that tourist flows to one destination are influenced by the tourist flows to another geographically related destination. The reasons for spatial dependence among destinations have been discussed by Yang and Wong (2012). Their study divided the factors and determinants contributing to the spatial spillover effect in tourist flows into the supply-side effect and the demand-side effect. Supply-side factors include productivity spillover, market access, joint promotion and negative events; the demand-side factor is the multi-destination travel pattern of tourists. Applications of spatial modelling in tourism demand forecasting are still rare. Only two studies (i.e., Long et al., 2019; Yang & Zhang, 2019) have been found that applied spatial models to forecast domestic tourism demand in Chinese cities.

Spatial heterogeneity

Spatial heterogeneity is another spatial effect considered in many empirical spatial studies. According to Zhang et al. (2011), spatial heterogeneity reflects the uniqueness of different regions, and can be identified in the form of distinct distributions, means, variance and patterns among subsets of observations. For instance, Capone and Boix (2008) found that different cities in Italy had different patterns with regard to regional tourism growth.

Geographically weighted regression (GWR) (Páez, Uchida, & Miyamoto, 2002) and spatial autoregressive local estimation (SALE) (Pace & LeSage, 2004) have been used in tourism modelling to measure spatial heterogeneity. However, before 2016, spatial heterogeneity was only captured for cross-sectional analysis in the tourism literature. Li et al. (2016) filled the gap by extending the SALE method to examine tourism and regional income inequality in China based on panel data. Although spatial heterogeneity has been captured in tourism modelling studies, it has not been considered in tourism forecasting yet. In the studies of Long et al. (2019) and Yang and Zhang (2019) on domestic tourism demand forecasting, only the global model was used for forecasting, which considered only the average spatial spillovers among all 341 Chinese cities in forecasting. Further to the spatial spillover effect mentioned above, this study will fill the gap and take both spatial spillover effect and spatial heterogeneity into account in tourism forecasting.

Spatial methodology development

Spatial econometrics has received much attention in regional and geographical economic studies because of its ability to capture spatial interactions among geographically related regions. Spatial regression models for cross-sectional data can be seen as an extension of traditional regression models, with the incorporation of spatial lags into the model specification. The spatial lag is defined as the weighted average of neighbouring values of a location (Kelejian and Prucha, 1998). Various spatial models have been developed that differ in the way they incorporate spatial lags. In this section, spatial models for cross-sectional data are presented first, and then the addition of the time dimension to spatial models for panel data is discussed.

Spatial models for cross-sectional data

Initially, most spatial models were designed to handle cross-sectional data. According to Lee and Yu (2010a, 2010b), the spatial autoregressive (SAR) model introduced by Cliff and Ord, 1973 received the most attention. Many Bother spatial regression models have also been applied to model and capture spatial spillover effects based on traditional simple regression models. Anselin (1988) provided a comprehensive analysis and review of spatial econometric models. Based on Anselin's study, LeSage and Pace (2009) further considered spatial dependence and summarised the spatial econometric models used in spatial modelling. However, those models share similar formulations based on traditional regression models. Spatial models extend regression models by incorporating a spatial lag to take spatial effects into account. The differences between spatial models are reflected in the way they incorporate spatial lags. The first way is to add a spatially lagged variable as an additional regressor in the regression model, which is referred to as a spatial lag model. In SAR, the spatial lag is incorporated into the dependent variable only, and the endogenous effect is reflected by the lagged dependent variable. That is, the value of the dependent variable is influenced by the dependent variables of neighbouring regions. Exogenous effects can also be reflected in spatial regression models by imposing spatial lags onto independent variables, which means that the value of the dependent variable of a certain region is influenced by the values of the independent variables of neighbouring regions. A model incorporating both endogenous and exogenous interaction effects is the spatial Durbin model (SDM), which was thoroughly explained by LeSage and Pace (2009). The SDM includes both a spatially lagged dependent variable and spatially lagged independent variables. Other types of spatial model impose a spatial lag on the error term including the spatial error model (SEM) and spatial autoregressive combined (SAC) model, which combines the SAR and SEM, meaning that spatial lags are incorporated into both the dependent variable and the error term.

Spatial models for panel data

Besides cross-sectional data, panel data, defined as multidimensional data containing observations about multiple objects over multiple time periods, have attracted increasing attention recently, especially for spatial econometrics, because it is now easier to access large data sets that include a time dimension since the development of technology to sort and save data. By incorporating temporal panel data with time variants, Elhorst (2014) introduced static and dynamic spatial panel data models. The use of panel data extends modelling possibilities to multiple equations and more variations compared with cross-sectional data sets, which are limited to only one equation. Static spatial panel data models resemble the formulation of spatial dependence models for cross-

sectional data, except that spatial panel data models include observations over multiple time periods for the same set of cross-section observations. Dynamic spatial panel data models incorporate time lags into both the dependent variable and the spatial lagged dependent variable. With the addition of the time dimension, spatial models for cross-sectional data can be extended to spatial models for panel data, as in the spatiotemporal models applied in this study. The mathematical formulations of the models used are presented in the methodology section.

In the general field of forecasting, spatial forecasts have been found to outperform non-spatial forecasts in many studies. Spatial autocorrelation has been taken into account when using spatial panel data models in forecasting. As discussed by Dormann et al. (2007), spatial data analysis is complicated by spatial autocorrelation phenomena, which occur when variables sampled within a region are not independent of each other (Tobler, 1970). Although it is complicated to account for spatial autocorrelation, many studies have implemented it in spatial models, usually leading to improved forecasting performance. For instance, Baltagi and Li (2006) predicted the demand for liquor based on panel data for 43 states in the period from 1965 to 1994, and the results showed that estimators accounting for spatial correlation and heterogeneity across states yield the best one-year-ahead forecasting results. Following Baltagi and Li (2006), Longhi and Nijkamp (2007) compared different methods of forecasting regional labour market developments to analyse whether the consideration of spatial autocorrelation improves forecasting performance. Their results showed that spatial models outperform their non-spatial counterparts, and the spatial error model with spatial autocorrelation outperforms models that ignore spatial autocorrelation. This is consistent with the findings of Giacomini and Granger (2004), which show that ignoring even weak spatial autocorrelation can lead to inaccurate forecasts.

The literature review reveals several research gaps. First, in the field of spatial econometrics, SAR and SDM are most frequently used when spatial lags are incorporated into the dependent and independent variables. For spatial panel data models with a time dimension, SEM and SAC are rarely used because of the difficulty in estimating the spatial moving average term. However, as Yesilyurt and Elhorst (2017) noted, if countries in the same regions face similar unobserved fluctuations, then the spatial lag on the error term might be pertinent. Thus, in this study, the spatiotemporal autoregressive model uses full specifications of spatial and time lags on both the independent variable and the moving average term. For the first time in any field of forecasting, this fully specified spatiotemporal model is employed for a forecasting exercise. Second, local spatial models allowing for different specifications for different geographical locations have not been applied in any field of forecasting either. This study fills the gap by locally estimating the spatial model for every country to fully reflect spatial heterogeneity. Thus this study contributes not just to the tourism forecasting literature, but the broader field of forecasting in general. In terms of the context of forecasting, previous studies applied spatiotemporal models to forecasting tourism demand at the city level in a context of domestic tourism, whereas this study forecasts tourist arrivals in 37 European countries in a context of international tourism. From a demand perspective, the spatial effect of international tourism demand can be different from that of domestic tourism demand. Multi-destination travel is more common among international, especially long-haul, tourists. Thus, it is worth examining if capturing the spatial effect may contribute to accuracy improvement in international tourism demand forecasting. Moreover, previous studies using spatiotemporal models generate forecasts up to two steps ahead whereas in this study, the forecasting horizons are extended to three steps ahead.

Methodology

Global model

To account for both spatial dependence and time variants, a spatiotemporal autoregressive model is used in this study. The global model is developed first, which assumes consistent model specifications for all countries, and spatial heterogeneity can only be reflected through spatial fixed effects.

The global dynamic spatiotemporal autoregressive model proposed in this study can be written as follows:

$$Y_t = \lambda WY_t + \gamma Y_{t-1} + \rho WY_{t-1} + \mu + S_t$$
$$S_t = \phi WS_t + \psi S_{t-1} + \epsilon_t \qquad (1)$$

where Y_t denotes an $N \times 1$ vector consisting of one observation of the dependent variable for every spatial unit $i = (1, \cdots, N)$ in the sample at time t $(t = 2, \cdots, T)$, where N and T are the numbers of spatial units and time periods respectively. Y_{t-1} denotes the dependent variable lagged in time. W is the $N \times N$ spatial weight matrix determined by the geographical locations of the destinations. It assigns 1 to the destination's k nearest neighbours and 0 otherwise. Distances are measured by great circle distance using the latitude and longitude of the destination country's capital; WY_t denotes the dependent variable lagged in space, and WY_{t-1} denotes dependent variable lagged in both space and time, μ is an $N \times 1$ vector representing country-specific spatial fixed effects, which indicates that the model is a panel model instead of a pure time series model; S_t is an $N \times 1$ vector denoting the error specification of the model; ϵ_t is a vector of independent and identically distributed (i.i.d.) disturbance with zero mean, and λ, γ, ρ, ϕ and φ are parameters to be estimated.

To determine the number of neighbours k in the spatial weight matrix W, a process known as calibration will be carried out by estimating the model recursively with an incrementing k ($k = 1$ to 9). The optimal k that is associated with the smallest in-sample mean absolute percentage error (MAPE) will be selected with the reason discussed in Section 0.

The model proposed in this study extends the ARIMA model by further incorporating spatial lags into both the autoregressive and the moving average terms. Compared with the model of Long et al. (2019), which incorporates spatial lag (λWY_t) and time lag (γY_{t-1}) separately, this study proposes a more general specification that incorporates all possible spatial and time lags on both Y and

ϵ (Elhorst, 2014), including the spatiotemporal lag (ρWY_{t-1}), spatial error lag (ϕWS_t) and error time lag (ψS_{t-1}), thus further taking into account the spatial dependence of the unobserved fluctuations among tourist arrivals in neighbouring countries. The spatial lag refers to the weighted average of the neighbouring dependent variable and the spatial error lag refers to the weighted average of neighbouring unobserved fluctuations.

Local model

As noted in the literature review, spatial heterogeneity, which can be defined as the uniqueness of different geographical regions, has been identified in previous studies. In the global model, the specification of W and the regression coefficients are assumed to be the same for all different countries, which can be unrealistic considering the existence of spatial heterogeneity. As a remedy, the local model, which allows for different specifications for different countries, can fully reflect spatial heterogeneity and provide more specific implications for individual countries (Wheeler & Tiefelsdorf, 2005). A spatial autoregressive local estimation (SALE) method was proposed by Pace and LeSage (2004) to estimate the local model, but the time dimension was not considered and thus the limited localised subsample size makes the local model very sensitive to model specifications. To address this issue, Li et al. (2016) and Long et al. (2019) expanded the SALE method to incorporate both spatial and time lags into a local SDM and SAR model, respectively. And this study further extends the local spatiotemporal model by incorporating spatial lags into both the autoregressive and the moving average terms.

In line with the specification of the global model, the local model can be written as follows:

$$U(i)Y_t = \lambda U(i)WY_t + \gamma U(i)Y_{t-1} + \rho U(i)WY_{t-1} + U(i)\mu + U(i)S_t$$
$$U(i)S_t = \phi U(i)WS_t + \psi U(i)S_{t-1} + U(i)\epsilon_t \tag{2}$$

where $U(i)$ denotes an $N \times N$ spatial weight matrix for country i, with m nearest neighbours identified. It assigns 1 to the neighbouring countries of country i and 0 otherwise. Therefore, $U(i)$ extracts a country-specific subsample from the whole sample for every country based on geographical distance, $k + 1$ countries are included in every subsample and the specification of W determines the number of neighbours within the subsample for country i. Thus, unlike the global model, which only needs one estimation after the specification of W, the local model needs to be estimated separately N times for every country, where N is the number of countries in the sample.

Similar to the global model, the number of neighbours k in the local model is determined by a calibration procedure described in Section 0. The calibration process also selects the bandwidth – a hyperparameter that governs the size of the subsample. Further details of the calibration process are presented in Section 0.

Data description

According to Song and Li (2008), tourist arrivals at a destination country from an origin country are the most popular measure of tourist demand. Thus, this study uses tourist arrivals as the measure of tourist demand. A panel data set consists of annual international tourist arrivals in 37 European countries from 1995 to 2018 (except eight countries for which tourist arrivals data are available only up to 2017) were collected from the World Bank and international tourism highlights published by UNWTO. Examination of the performance of annual tourism demand forecasting has useful implications on medium- to long-term strategic planning of tourism-related businesses and destination governments. As the scale of tourist arrivals in different countries varies considerably, which makes model estimation as a panel difficult, a logarithmic transformation was undertaken.

Fig. 1 shows the logged tourist arrivals for the 37 European countries, which are classified into Southern Europe, Central and Eastern Europe, Northern Europe and Western Europe. The shared patterns of neighbouring countries visualise the existence of spatial dependence, whereas the different patterns for some of the countries (e.g., Poland in Central and Eastern Europe and Luxembourg in Western Europe) indicate that spatial heterogeneity should not be neglected.

Empirical results

Model calibration

As explained in Sections 0 and 0, a calibration procedure is carried out to search the optimal number of neighbours k in the spatial weight matrix W. Data from 1995 to 2012 are used for model calibration and estimation. For the local model, the calibration also selects the bandwidth, which regulates the number of countries included the subsample for each focal country, in a range from 10 to 37 countries (the whole sample) including the focal country. Therefore, for every local country, the model is estimated 28 × 9 (choices of bandwidth × choices of neighbours) times, totalling 252 times, and the specifications with the best model fit are selected for forecasting. All models are estimated with a maximum of one lag in time, which is common for estimation based on annual data. The model selection criterion used in this study is the in-sample MAPE. In previous studies (e.g., Elhorst, Zandberg, & De Haan, 2013; Li et al., 2016), the residual variance was used as the selection criterion for model calibration. However, the residual variance measures the average model fit across all countries included in the subsample. As the aim of this study is to generate forecasts for individual countries, the model fit should be measured on the basis of the focal country rather than its neighbours. Therefore, the in-sample MAPE of the fitted values is chosen as the model selection criterion. The same criterion is applied to the global model
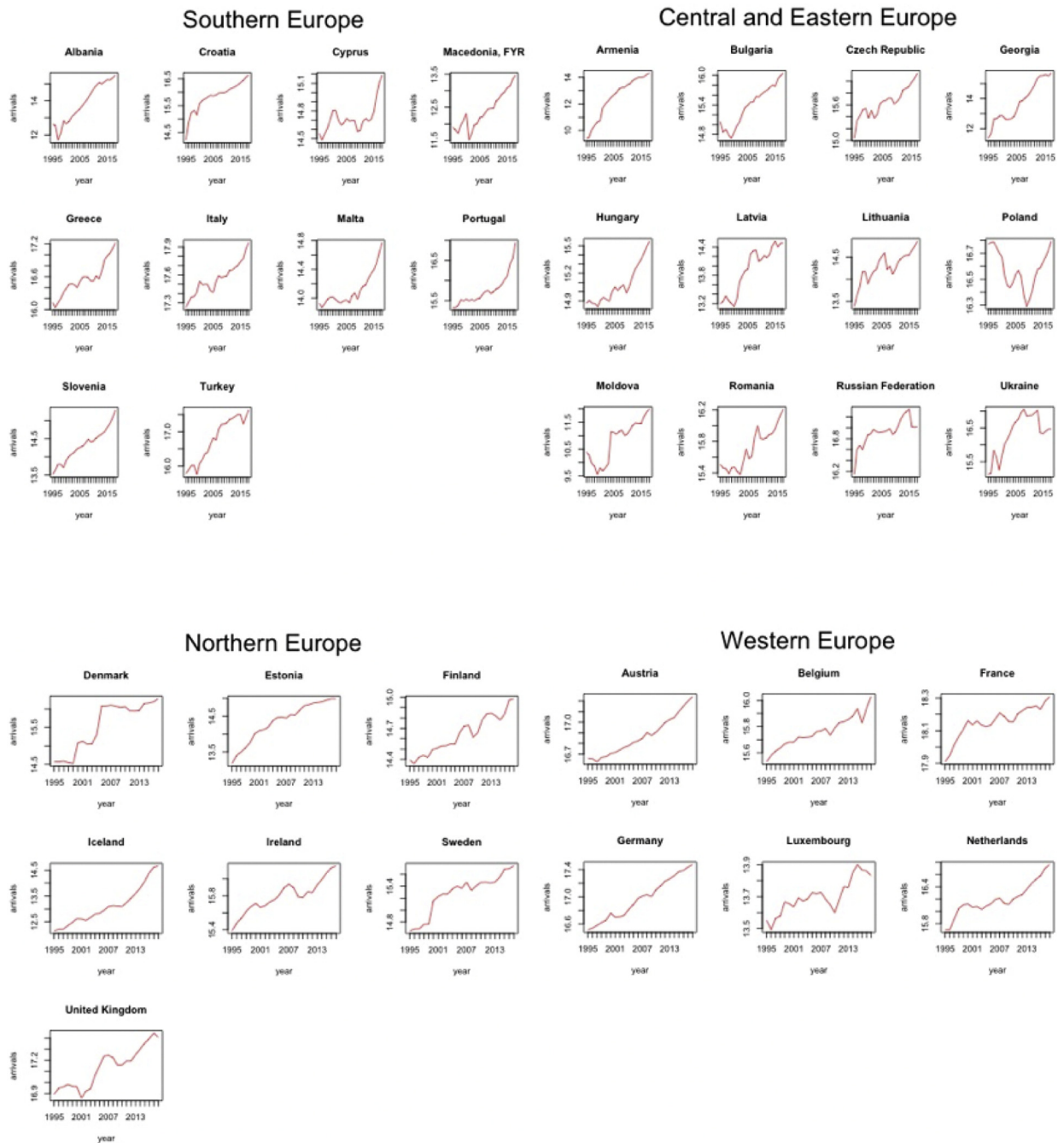
Fig. 1. Tourist arrivals in 37 European countries from 1995 to 2018 (logarithmic scale).

calibration for consistency. The choice of MAPE is also supported by the forecasting results which show that specifications selected by the in-sample MAPE outperform those selected by residual variances in this study. For the global model, the optimal number of neighbours (k) is one, indicating that for each country, one neighbouring country is identified. And the hyperparameters for each local model including the bandwidth and the number of neighbours (k) are shown in Table 1.

The average subsample size across all local models is 21 and the average number of neighbours within each subsample is 3.62. The different bandwidths and numbers of neighbours for different local countries indicate the spatial heterogeneity mentioned in the previous sections.

Table 1

Local model specifications.

Destination	Bandwidth	Neighbour	In-sample MAPE
Albania	34	1	13.821
Armenia	11	1	7.435
Austria	16	8	1.602
Belgium	24	1	2.281
Bulgaria	24	3	2.796
Cyprus	22	1	4.154
Czech Republic	9	2	4.635
Germany	36	6	3.670
Denmark	13	8	3.001
Spain	13	9	3.292
Estonia	29	1	5.634
Finland	21	3	4.848
France	10	1	2.413
United Kingdom	19	9	2.068
Georgia	14	1	11.856
Greece	33	1	4.501
Croatia	32	2	3.174
Hungary	12	7	3.741
Ireland	17	2	4.495
Iceland	18	6	7.018
Italy	11	1	2.265
Lithuania	18	2	14.463
Luxembourg	28	6	4.673
Latvia	36	6	7.870
Moldova	17	4	9.815
Macedonia, FYR	16	2	5.703
Malta	30	2	4.917
Netherlands	36	1	4.152
Norway	18	2	4.602
Poland	9	2	7.360
Portugal	21	8	3.420
Romania	20	4	8.364
Russian Federation	11	1	6.368
Slovenia	29	7	4.269
Sweden	13	2	5.351
Turkey	9	5	6.324
Ukraine	12	6	8.206

Forecasting

After model calibration and estimation, one- to three-step-ahead forecasts were generated in a rolling window from 2013 to the end of the sample period for both the global and local models. In total, six one-step-ahead forecasts, five two-step-ahead forecasts and four three-step-ahead forecasts were produced. In cases for which the data for 2018 are not yet available, five one-step-ahead, four two-step-ahead and three three-step-ahead forecasts were generated. Thanks to the multiple forecast points for each step, the influence of potential outliers can be moderated. To produce the forecasts, Eq. 1 can be rewritten as follows:

$$(I - \lambda W)Y_t = \gamma Y_{t-1} + \rho W Y_{t-1} + \mu + S_t$$
$$(I - \varnothing W)S_t = \varphi S_{t-1} + \epsilon_t \tag{3}$$

where I is an $N \times N$ identity matrix. And Eq. 3 can be further arranged as follows:

$$Y_t = (I - \lambda W)^{-1}[\gamma Y_{t-1} + \rho W Y_{t-1} + \mu + (I - \phi W)^{-1}(\varphi S_{t-1} + \epsilon_t)] \tag{4}$$

As ϵ_t is a vector of i.i.d. disturbance with zero mean, the one-step-ahead forecasts can be computed as:

$$\widehat{Y}_{t+1} = (I - \lambda W)^{-1}[\gamma Y_t + \rho W Y_t + \mu + (I - \phi W)^{-1}\varphi \, \widehat{S}_t]$$
$$\widehat{S}_t = Y_t - \widehat{Y}_t \tag{5}$$

To evaluate the forecasting performance of the proposed spatiotemporal models, ARIMA, ETS and Naïve 1 models were chosen as benchmark models, following the common practice in the recent tourism forecasting literature (e.g., Chen, Li, Wu, & Shen, 2019; Wen, Liu, Song, & Liu, 2020). Similarly, one- to three-step-ahead forecasts were generated using these models for every country. The forecasting results were evaluated using two criteria: MAPE and mean absolute scaled error (MASE). MAPE is the most popular error evaluation method in the forecasting literature (Athanasopoulos & de Silva, 2012) and can be expressed as follows:

Table 2
Average performance of different forecasting models.

Horizon	Measure	Global	Local	ARIMA	ETS	Naïve
1 step	MAPE	0.066 (2.68)	**0.062 (1.92)**	0.066 (2.22)	0.079 (3.86)	0.079 (4.05)
	MASE	1.765(2.68)	1.629 **(1.86)**	**1.617** (2.27)	1.951 (3.89)	2.114 (3.95)
2 steps	MAPE	0.118 (2.54)	**0.107 (2.11)**	0.136 (2.78)	0.145 (3.86)	0.142 (3.65)
	MASE	3.240 (2.30)	**2.900 (2.22)**	3.426 (2.65)	3.700 (3.76)	3.960 (3.90)
3 steps	MAPE	0.163 (2.65)	**0.143 (2.03)**	0.201 (2.68)	0.204 (3.84)	0.195 (3.62)
	MASE	4.694 (2.81)	**4.073 (1.82)**	5.053 (2.65)	5.443 (3.70)	5.719 (3.89)

Note: The values in bold represent the most accurate model among the five models. Values in brackets represent average rankings among 37 countries.

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Y_t - \hat{Y_t}}{Y_t} \right|$$

(6)

Hyndman and Koehler (2006) proposed MASE as another error measurement method which scales the error based on the one-step-ahead in-sample mean absolute error (MAE) generated by the Naïve forecasting method. A scaled error can be expressed as follows:

$$q_t = \frac{Y_t - F_t}{\frac{1}{T-1} \sum_{i=2}^{T} |Y_i - Y_{i-1}|}$$

(7)

where Y_t is the real output and F_t is the forecast, and the denominator is the one-step-ahead in-sample MAE from Naïve forecasts. MASE = mean ($|q_t|$). The performances after conducting the same forecasting practice as for the spatiotemporal forecasting models for the three benchmark models for every country are compared and shown in Table 2.

Overall, the global model outperforms the other three benchmark models for two- and three-step-ahead forecasts, which confirms the conclusions in the previous literature that the incorporation of spatial spillover effects can improve forecasting accuracy compared with non-spatial models. Although the ARIMA model outperforms the global spatial model for one-step-ahead forecasting, the global model is estimated only once as a whole whereas the ARIMA model is estimated separately for every single country. Thus, the performance of the ARIMA model is more comparable with that of the local spatial model, which generates better one-step-ahead forecasts than the ARIMA model if evaluated by MAPE. The MASE values for the two models are quite similar. The forecasting performance of the two spatial models is more consistent and stable as the forecasting horizon extends, compared with the ARIMA model. Meanwhile, the local model performs better than the global model across all forecasting horizons, which suggests that fully reflecting spatial heterogeneity can improve forecasting accuracy compared with the global model, which only reflects heterogeneity through country-specific fixed effects.

Country-level forecasting performance was also compared among the five models. Average MAPE values from one-step-ahead to three-step-ahead forecasting were calculated for every country and the results are shown in Fig. 2. As only two MAPE values are above 0.5 (ARIMA MAPE of Georgia and Ukraine), for clearer presentation, the scale is from 0 to 0.5. For most countries, the global model and local models perform better than the other benchmarks. For many countries, such as Bulgaria, the United Kingdom and Malta, notable improvements of the local models over the global model can be observed in the figure. The MAPE and MASE values of each forecasting method were ranked and the average rankings of the one- to three-step-ahead forecasting results at the country level are shown in Table 2. At the country level, the local models have the highest average ranks for all forecasting horizons using both error measures, again confirming that incorporating both spatial dependence and spatial heterogeneity can improve individual destinations' forecasting performance compared with the global model and the other benchmark models. Among the five forecasting models, on average for 26 countries out of 37 countries in total, the local spatial models generate the best or second-best forecasting results at the country level according to both error measures, which further confirms the superiority of the local spatial model in forecasting.

To confirm the accuracy improvement of the local spatial model, the Harvey-Leybourne-Newbold (HLN) test (Harvey, Leybourne, & Newbold, 1997) is conducted to examine statistically significant differences in forecasting accuracy between the local spatial model and the benchmark models (Song, Li, Witt, & Athanasopoulos, 2011). As shown in Table 3, across all three forecasting horizons, the accuracy improvements of the local spatial model against all benchmark models are significant at the 1% level with only one exception.

To visualise the improvement forecasting performance of the local spatial models, dynamic forecasts for some countries using the local spatial models were conducted and compared with the real tourist arrivals and the ARIMA model, the best performing model among the benchmarks. The two models were estimated based on data from 1995 to 2012 and one- to six-step-ahead dynamic forecasts were generated. The historical tourist arrivals series of Armenia is relatively stable with steady growth. Fig. 3 displays the fitted value and dynamic forecasts of tourist arrivals in Armenia generated by the two models. The ARIMA model tends to over-estimate the increasing trend, whereas the local spatial model generates more steady forecasts with the influence of the spatial spillover effect from the neighbouring countries. The forecasts generated by the local spatial model clearly outperforms ARIMA for Armenia.
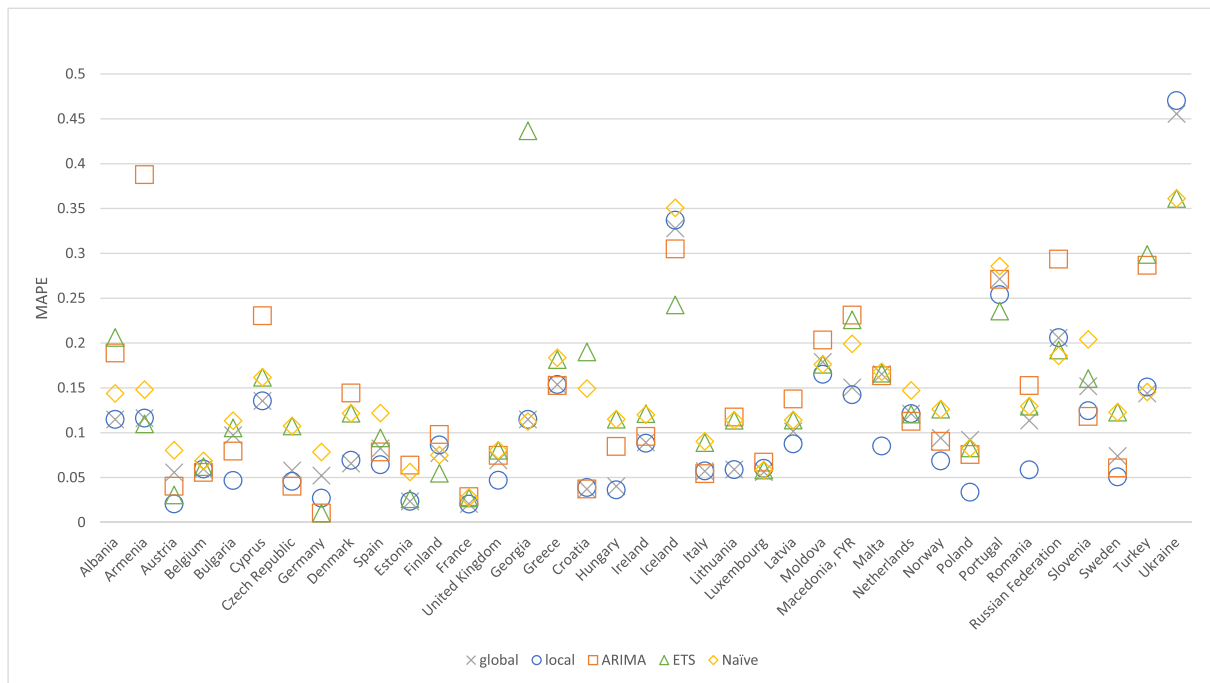
Fig. 2. Average out-of-sample mean absolute percentage error (MAPE) at country level.

Table 3
HLN tests for statistical difference between forecasts of the local spatial model and those of other models.

Horizon	Global	ARIMA	ETS	Naïve
1 step	3.904***	1.194	5.08***	6.94***
2 steps	4.408***	4.489***	5.453***	7.329***
3 steps	5.478***	4.784***	5.366***	6.228***

Note: *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively. Positive values indicate that the local spatial model forecasts more accurately.
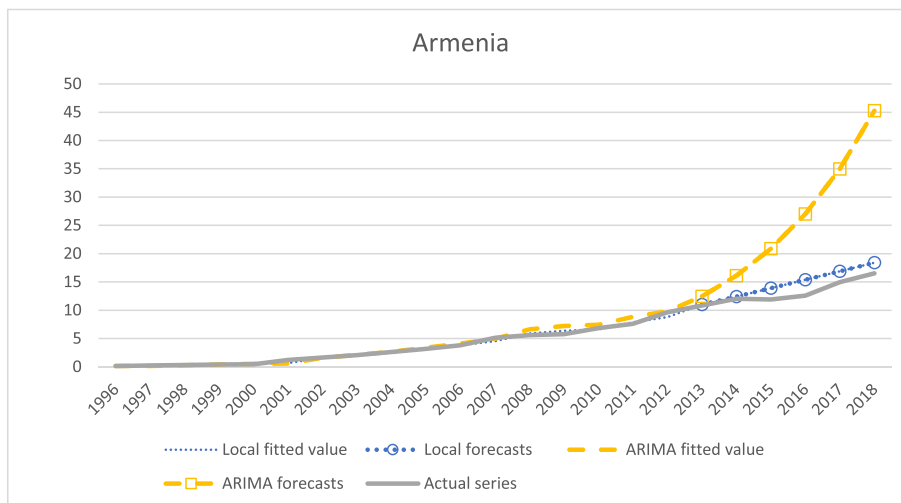

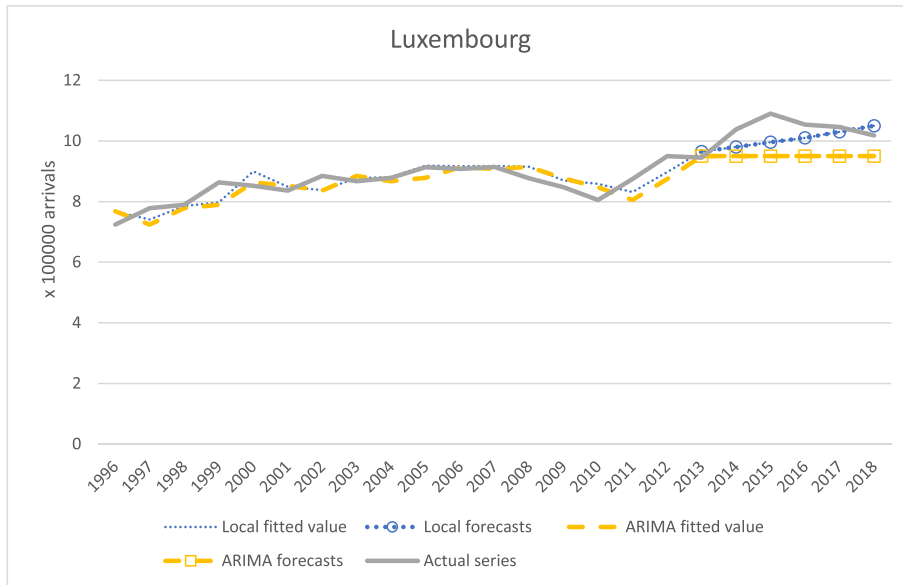**Fig. 3.** Fitted values (1996–2012) and dynamic forecasts (2013–2018) of tourist arrivals in Armenia.

Fig. 4. Fitted values (1996–2012) and dynamic forecasts (2013–2018) of tourist arrivals in Luxembourg.

Luxembourg, another country with fluctuating and unstable historical data, is also selected for illustration purposes. Again, dynamic forecasts were generated using the local spatial model and the ARIMA model and compared with the actual values, as shown in Fig. 4. As the data series of Luxembourg tourist arrivals has a few turning points, the ARIMA model selects p, d, q specifications as (0,0,1), which is simply a random walk (Naïve model). Although the local spatial model cannot capture the turning point of Luxembourg in 2015, eight nearest neighbours of Luxembourg were identified and included in the estimation, and thus the forecasting results are moved closer towards the actual series with the consideration of neighbouring countries' tourist arrivals. Again, the local spatial model significantly outperforms the ARIMA model. Across the two cases, as the forecasting horizon extends, the superiority of the local model becomes more evident, which confirms our previous results that for two- and three-step-ahead forecasting, the spatial model clearly stands out with regard to both average and individual forecasting accuracy.

Conclusion

This study proposes a spatiotemporal autoregressive model incorporating both spatial dependence and spatial heterogeneity to improve forecasting accuracy. The proposed model spatially extends the ARIMA model in the most general form by applying spatial lags to both the response variable and the moving average term. Based on this specification, a local model was developed to account for spatial heterogeneity. The forecasting performance of the proposed approach was empirically tested using international tourist arrivals data for 37 European countries. The results were compared with those of the three benchmark models including the ARIMA, ETS and Naïve 1 models based on MAPE and MASE criteria. The superior forecasting performance of both the global and local spatial models compared with non-spatial models suggests that incorporating spatial effects can improve forecasting performance, which is consistent with previous studies (Long et al., 2019;Yang and Zhang, 2019). Forecasting accuracy can be further improved by taking account of spatial heterogeneity, which allows for individual specifications for each country. At the country level, the local spatial model outperforms all of the other models with respect to average rankings. In over 70% of cases, the local spatial model generates the most or second-most accurate results. Thus, incorporation of both spatial dependence and spatial heterogeneity can improve forecasting accuracy compared with non-spatial models. Moreover, as the forecasting horizon extends, the superiority of the local spatial model clearly stands out, which indicates the advantage of the local spatial model, especially in terms of three-year-ahead tourism demand forecasting.

In the field of tourism research, spatiotemporal models have rarely been used in forecasting, with only two exceptions (Long et al., 2019; Yang and Zhang, 2019). The present study for the first time extends the spatiotemporal model by incorporating spatial lags into both the response variable and the moving average term, thus taking account of all types of spatiotemporal effect. And since this full specification has never been applied in any field of forecasting, this study contributes to the general literature of forecasting. Second, both global and local spatiotemporal models are developed for tourism forecasting in this study, which represents the first attempt at using local spatiotemporal models in forecasting not just in a tourism context, but in the general field of forecasting. In terms of the context of forecasting, previous studies have focused on city-level domestic tourism. This study is the first to present empirical evidence on the benefit of accounting for spatial effects in forecasting tourist arrivals at an international level. Additionally, this study extends the forecasting horizons to three-step-ahead, and the superiority of the local model becomes more obvious as the forecasting horizon extends. Thus, with an extended spatiotemporal model, this study makes a useful addition to the existing tourism forecasting literature by expanding the evidence base of the advantages of spatiotemporal forecasting with an extended spatiotemporal model.

This study has several limitations. The first limitation is related to the short data period, which is a universal problem for spatial panel data analysis. In this study, only one- to three-step-ahead short-term forecasts were generated. The lack of data makes it difficult to generate and evaluate long-term forecasting performance. Secondly, for some countries, the turning points of the data series appeared in the forecasting period. Thus, models that behave more like a Naïve model would perform better, but they may not fit the training data well. The local models tend to be more sensitive to changes in the data, so the specification chosen by any criteria based on model fit would perform badly in forecasting due to the reverse of trends. Therefore, in future studies, tourist arrivals data with higher frequencies should be considered to guarantee the sufficiency of the data for estimation and forecasting. Lastly, exogenous explanatory variables can be incorporated to test if the effect of turning points can be accounted for, and if the forecasting performance can be further improved. The limitations of the present study point to possible directions for future research.

References

Álvarez, I. C., Barbero, J., & Zofío, J. L. (2016). A spatial autoregressive panel model to analyze road network spillovers on production. *Transportation Research Part A: Policy and Practice, 93*, 83–92.

Anselin, L. (1988). *Spatial econometrics: Methods and models.* Dordrecht: Kluwer Academic.

Athanasopoulos, G., & de Silva, A. (2012). Multivariate exponential smoothing for forecasting tourist arrivals. *Journal of Travel Research, 51*(5), 640–652.

Ayeh, J. K., & Lin, S. (2011). Estimating tomorrow's tourist arrivals: Forecasting the demand for China's tourism using the general-to-specific approach. *Tourism and Hospitality Research, 11*(3), 197–206.

Baltagi, B. H., & Li, D. (2006). Prediction in the panel data model with spatial correlation: The case of liquor. *Spatial Economic Analysis, 1*(2), 175–185.

Cao, Z., Li, G., & Song, H. (2017). Modelling the interdependence of tourism demand: The global vector autoregressive approach. *Annals of Tourism Research, 67*, 1–13.

Capone, F., & Boix, R. (2008). Sources of growth and competitiveness of local tourist production systems: An application to Italy (1991–2001). *The Annals of Regional Science, 42*(1), 209–224.

Chaitip, P., & Chaiboonsri, C. (2014). International tourists arrival to Thailand: Forecasting by non-linear model. *Procedia Economics and Finance, 14*, 100–109.

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

Chen, J. L., Li, G., Wu, D. C., & Shen, S. (2019). Forecasting seasonal tourism demand using a multiseries structural time series method. *Journal of Travel Research, 58*(1), 92–103.

Chen, M., Ying, L., & Pan, M. (2010). Forecasting tourist arrivals by using the adaptive network-based fuzzy inference system. *Expert Systems with Applications, 37*(2), 1185–1191.

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

Cliff, A. D., & Ord, J. K. (1973). *Spatial autocorrelation: Monographs in spatial environmental systems analysis.* London: Pion Limited.

Condeço-Melhorado, A., Tillema, T., de Jong, T., & Koopal, R. (2014). Distributive effects of new highway infrastructure in the Netherlands: The role of network effects and spatial spillovers. *Journal of Transport Geography, 34*, 96–105.

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., ... Daniel Kissling, W. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography, 30*(5), 609–628.

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

Elhorst, J. P., Zandberg, E., & De Haan, J. (2013). The impact of interaction effects among neighbouring countries on financial liberalization and reform: A dynamic spatial panel data approach. *Spatial Economic Analysis, 8*(3), 293–313.

Fingleton, B., & López-Bazo, E. (2006). Empirical growth models with spatial effects. *Papers in Regional Science, 85*(2), 177–198.

Gareta, R., Romeo, L. M., & Gil, A. (2006). Forecasting of electricity prices with neural networks. *Energy Conversion and Management, 47*(13–14), 1770–1778.

Giacomini, R., & Granger, C. W. (2004). Aggregation of space-time processes. *Journal of Econometrics, 118*(1–2), 7–26.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting, 13*(2), 281–291.

Huang, H., Zhou, H., Wang, J., Chang, F., & Ma, M. (2017). A multivariate spatial model of crash frequency by transportation modes for urban intersections. *Analytic Methods in Accident Research, 14*, 10–21.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688.

Jiao, E. X., & Chen, J. L. (2019). Tourism forecasting: A review of methodological developments over the last decade. *Tourism Economics, 25*(3), 469–492.

Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics, 17*(1), 99–121.

Kulendran, N., & King, M. L. (1997). Forecasting international quarterly tourist flows using error-correction and time-series models. *International Journal of Forecasting, 13*(3), 319–327.

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.

Lee, K. N. (2011). Forecasting long-haul tourism demand for Hong Kong using error correction models. *Applied Economics, 43*(5), 527–549.

Lee, L., & Yu, J. (2010a). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics, 154*(2), 165–185.

Lee, L., & Yu, J. (2010b). Some recent developments in spatial panel data models. *Regional Science and Urban Economics, 40*(5), 255–271.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics.* Chapman and Hall/CRC.

Li, G., Song, H., & Witt, S. F. (2004). Modeling tourism demand: A dynamic linear AIDS approach. *Journal of Travel Research, 43*(2), 141–150.

Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research, 44*(1), 82–99.

Li, H., Chen, J. L., Li, G., & Goh, C. (2016). Tourism and regional income inequality: Evidence from China. *Annals of Tourism Research, 58*, 81–99.

Long, W., Liu, C., & Song, H. (2019). Pooling in tourism demand forecasting. *Journal of Travel Research, 58*(7), 1161–1174.

Longhi, S., & Nijkamp, P. (2007). Forecasting regional labor market developments under spatial autocorrelation. *International Regional Science Review, 30*(2), 100–119.

Ma, T., Hong, T., & Zhang, H. (2015). Tourism spatial spillover effects and urban economic growth. *Journal of Business Research, 68*(1), 74–80.

Majewska, J. (2015). Inter-regional agglomeration effects in tourism in Poland. *Tourism Geographies, 17*(3), 408–436.

Miranda, K., Martínez-Ibañez, O., & Manjón-Antolín, M. (2017). Estimating individual effects and their spatial spillovers in linear panel data models: Public capital spillovers after all? *Spatial Statistics, 22*, 1–17.

Pace, R. K., & LeSage, J. P. (2004). Spatial autoregressive local estimation. In J. Mur, H. Zoller, & A. Getis (Eds.). *Recent advances in spatial econometrics* (pp. 31–51). Palgrave Publishers.

Páez, A., Uchida, T., & Miyamoto, K. (2002). A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. *Environment and Planning A, 34*(4), 733–754.

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

Romão, J., Guerreiro, J., & Rodrigues, P. M. (2017). Territory and sustainable tourism development: A space-time analysis on European regions. *Region, 4*(3), 1–17.

Romão, J., & Nijkamp, P. (2019). Impacts of innovation, productivity and specialization on tourism competitiveness – A spatial econometric analysis on European regions. *Current Issues in Tourism, 22*(10), 1150–1169.

Saayman, A., & Botha, I. (2015). Evaluating non-linear approaches in forecasting tourist arrivals. *Economic Research Southern Africa. Tourism Economics, 23*(3), 594–613.

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

Song, H., Li, G., Witt, S. F., & Athanasopoulos, G. (2011). Forecasting tourist arrivals using time-varying parameter structural time series models. *International Journal*

of Forecasting, 27(3), 855–869.

Song, H., & Wong, K. K. (2003). Tourism demand modeling: A time-varying parameter approach. *Journal of Travel Research, 42*(1), 57–64.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*(sup1), 234–240.

Wen, L., Liu, C., Song, H., & Liu, H. (2020). Forecasting tourism demand with an improved mixed data sampling model. *Journal of Travel Research.* https://doi.org/10.1177/0047287520906220 In press.

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems, 7*(2), 161–187.

Wu, D. C., Song, H., & Shen, S. (2017). New developments in tourism and hotel demand modeling and forecasting. *International Journal of Contemporary Hospitality Management, 29*(1), 507–529.

Yang, Y., & Fik, T. (2014). Spatial effects in regional tourism growth. *Annals of Tourism Research, 46*, 144–162.

Yang, Y., & Wong, K. K. (2012). A spatial econometric approach to model spillover effects in tourism flows. *Journal of Travel Research, 51*(6), 768–778.

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

Yesilyurt, M. E., & Elhorst, J. P. (2017). Impacts of neighboring countries on military expenditures: A dynamic spatial panel approach. *Journal of Peace Research, 54*(6), 777–790.

Zhang, Y., Xu, J., & Zhuang, P. (2011). The spatial relationship of tourist distribution in Chinese cities. *Tourism Geographies, 13*(1), 75–90.

United Nations World Tourism Organization (UNWTO). (2019). *International Tourism Highlights, 2019 Edition.* Madrid: UNWTO. https://doi.org/10.18111/9789284421152.

Xiaoying Jiao is a PhD student in the School of Hospitality and Tourism Management at the University of Surrey in the UK.

Gang Li is Professor of Tourism Economics at the University of Surrey in the UK (g.li@surrey.ac.uk).

Jason Li Chen is a Senior Lecturer at the University of Surrey in the UK.