New Developments in Tourism and Hotel Demand Modeling and Forecasting

Wu, D., Song, H. and Shen, S.

The final, published version in International Journal of Contemporary Hospitality Management, DOI: 10.1108/IJCHM-05-2015-0249, 2016 is available at:

https://dx.doi.org/10.1108/IJCHM-05-2015-0249

New Developments in Tourism and Hotel Demand Modeling and Forecasting

Abstract

Purpose

The purpose of the study is to review recent studies published from 2007–2015 on tourism and hotel demand modeling and forecasting with a view to identifying the emerging topics and methods studied and to pointing future research directions in the field.

Design/Methodology/approach

Articles on tourism and hotel demand modeling and forecasting published in both Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) journals were identified and analyzed.

Findings

This review finds that the studies focused on hotel demand are relatively less than those on tourism demand. It is also observed that more and more studies have moved away from the aggregate tourism demand analysis, while disaggregate markets and niche products have attracted increasing attention. Some studies have gone beyond neoclassical economic theory to seek additional explanations of the dynamics of tourism and hotel demand, such as environmental factors, tourist online behavior, and consumer confidence indicators, among others. More sophisticated techniques such as nonlinear smooth transition regression, mixed-frequency modeling technique, and nonparametric singular spectrum analysis have also been introduced to this research area.

Research limitations/implications

The main limitation of this review is that the articles included in this study only cover the English literature. Future review of this kind should also include articles published in other languages. The review provides a useful guide for researchers who are interested in future research on tourism and hotel demand modeling and forecasting.

Practical implications

This review provides important suggestions and recommendations for improving the efficiency of tourism and hospitality management practices.

Originality/value

The value of this review is that it identifies the current trends in tourism and hotel demand modeling and forecasting research and points out future research directions.

Keywords: Tourism and hotel demand; modeling and forecasting; methodological development

1. Introduction

Tourism has achieved a sustained expansion and diversification over the past six decades, despite various obstacles such as wars, regional epidemics and financial crises, some of which have had a significant impact on tourist flows in the short term. Accurate demand forecasts are the foundation upon which tourism and hotel-related business decisions depend, in terms of pricing and operation strategies. At the same time, medium- and long-term tourism and hotel demand forecasts are required for the investment decisions of private sector actors and government infrastructure investment.

Demand modeling and forecasting is an important area in tourism and hospitality research. According to Li *et al*. (2005), 420 studies on the topic of tourism demand modeling and forecasting were published between 1960 and 2002. Song and Li (2008) further reviewed 119 studies on the subject published between 2000 and 2007. Goh and Law (2011) reviewed 155 studies on the methodological progress of tourism demand forecasting published between 1995 and 2009. Relatively few studies have focused on hotel modeling and forecasting. Koupriouchina *et al*. (2014) produced an overview of 26 studies on the topic of forecasting in hotels published between 1985 and 2013.

More recently, studies with the theme of tourism and hotel modeling and forecasting have continued to appear in academic journals related to not only tourism and hospitality, but also some other fields, indicating growing interest in the research area. Based on these more recent studies and as a further extension of the existing reviews, this study aims to identify and highlight additional new themes in the field of tourism and hotel demand modeling and forecasting through reviewing the studies published during the period 2007–2015. The articles reviewed in this study were obtained by using the key words ‘tourism forecasting’, ‘hotel forecasting’, ‘tourism modeling’ and ‘hotel modeling’ in both science citation index (SCI) and social science citation index (SSCI) databases, as well as by following up citations in the articles identified. In total, 171 articles were obtained and reviewed. We acknowledge that some studies may have been omitted from the analysis. Nevertheless, the findings based on review of these 171 articles can provide useful insights into the new themes and trends in tourism and hotel demand forecasting.

The remaining sections of this study are organized as follows. Section 2 provides some descriptive statistics on the articles reviewed. Section 3 discusses the measurement of tourism demand, hotel demand and their determinants. Section 4 focuses on the methodological development based on three types of forecasting techniques: non-causal time series methods, econometric methods and artificial intelligence-based methods. Section 5 pays particular attention to some new research interests. Section 6 concludes the review and highlights potential future research directions.

2. Descriptive statistics

The full list of the 171 articles could be supplied on request due to the space limit. It is observed that the majority focus on tourism demand (145), while the remainder deal with hotel demand. In terms of the distribution of these articles, 130 were published in tourism and hospitality journals, and the rest in non-tourism and hospitality journals in such fields as forecasting, economics, statistics and computer sciences. In terms of data frequency for model estimation, 37, 42 and 61 studies employed annual, quarterly and monthly data, respectively. It was also found that five studies employed weekly data and six daily data. Meanwhile, 16 studies employed mixed frequency data and 10 employed cross-sectional data specifically focusing on demand analysis without forecasting.

In general, tourism and hotel demand research is centered around two broad directions. The first is aimed at developing new methodologies with a view of improving accuracy in tourism (or hotel) demand forecasting. This type of study normally uses a number of alternative forecasting models to forecast tourism or hotel demand, and their forecasting performances are then compared and evaluated based on various forecasting error measures. The second is to identify the relationships between tourism (or hotel) demand and their influencing factors based on established econometric models in order to quantify the effects of these factors on demand through demand elasticity analysis. Among the articles reviewed, 107 contained forecasting exercise and the rest focused on demand relationship analysis.

In forecasting exercises, it is common to assess the accuracy by examining the difference between forecasts and the real value of demand. There are a number of measurements for this assessment. The most widely used include the mean absolute percentage error (MAPE), the root mean square error (RMSE), the root mean square percentage error (RMSPE), and the mean absolute error (MAE); 66, 36, 33, and 19 studies adopted these accuracy measurements respectively. Others include the mean square error (MSE), Theil's U-statistics, the mean absolute deviation (MAD), and the mean absolute square error (MASE), among others. Some studies also applied statistical tests to examine the significance of forecasting performance; these include the Diebold-Mariano (DM) test, the Wilcoxon Signed-Rank test, and the Harvey, Leybourne and Newbold (HLN) test.

3. Variables and data

3.1 Measurement and market segmentation of tourism demand

Tourism demand for a particular destination is the quantity of tourism goods and services that consumers are willing to purchase during a specified period under a given set of conditions (Song and Witt, 2000). Tourist arrivals in a destination is the traditional and most widely used measure of tourism demand. Another two popular measures are tourist expenditure (e.g. Cortés-Jiménez and Blake, 2011; Smeral, 2010) and the number of nights stayed (e.g. Athanasopoulos and Hyndman, 2008; Baggio and Sainaghi, 2016). These three variables reflect the overall magnitude of tourism demand from different perspectives and their analysis may contribute directly to policy recommendations for destination governments and managerial decisions in private tourism businesses.

Instead of focusing on the aggregate tourism demand in a destination, some recent studies examine the disaggregate demand either by a particular market segment or for a specific type of tourism. Arrivals-based subcategories often include holiday tourist arrivals, business tourist arrivals, and arrivals for visiting friends and relatives (VFR); expenditure-based subcategories include meal expenditure, sightseeing expenditure, shopping expenditure, gaming expenditure, and so on. For instance, Cortés-Jiménez and Blake (2011) modeled tourist expenditures by four visit purposes: holidays, business, VFR, and study. Zheng *et al*. (2013) examined how recession affected Iowa's gaming volume by employing an autoregressive integrated moving average (ARIMA) with intervention model. Some studies have focused on subcategories according to transportation type, such as cruise tourists (e.g. Cuhadar *et al.*, 2014) and air passengers (e.g. Cazanova *et al*., 2014; Tsui *et al*., 2014).

Other recent studies have displayed interest in the examination of demand for niche tourism products. For example, Rodriguez *et al*. (2012) investigated the academic tourism market by examining higher education students' mobility in Galicia, Spain, and identified special determinants of the demand for this market based on dynamic panel data analysis. Lee *et al*. (2008) forecasted the visitor numbers to an international tourism expo in Korea by combining quantitative techniques and willingness-to-visit surveys. Due to the growth of tourism markets and the increasing diversity of tourist demand, more and more tourists have shifted from mass tourism to alternative tourism. Demand analysis of these markets is valuable and the results will benefit both academics and tourism practitioners. Currently, some niche tourism products (e.g. wine tourism, film tourism, and golf tourism) and market segments (e.g. volunteers, backpackers, and gap-year students) have matured and gained increasing attention by scholars, but these markets have yet to be explored by quantitative demand analysis. Researchers are therefore encouraged to analyze these demands quantitatively once data are available.

3.2 Measurement of hotel demand

Hotel demand modeling and forecasting is often related to hotel revenue management. It has also been used for hotel business operation management, business planning, purchasing decision, and inventory control (Lim *et al*., 2009). According to Song and Li (2008), only three studies on hotel demand forecasting were published during the period 2000–2006. Recently, researchers have paid even more attention to this sector. 25 studies that examine hotel demand modeling and forecasting have been identified in this review. The demand for hotel accommodation is measured by a variety of variables, from different perspectives. Some variables relate to the scale of demand, such as guest arrivals (Guizzardi and Stacchini, 2015), the number of nights stayed (Falk, 2014; Lim *et al*., 2009), the number of rooms sold (such as Corgel *et al*., 2013; Song *et al*., 2011b), and occupancy rates (Koupriouchina *et al*., 2014; Wu *et al*., 2010). Some variables measure hotel demand from a financial performance perspective, such as sales revenue (Chen, 2013), revenue per available room (RevPAR) (Zheng, 2014), and profit per available room (profitPAR) (Croes and Semrad, 2012).

Macro-level hotel demand forecasting provides useful information to the hotel industry as a whole, though the contribution of such studies is limited given that the data are highly aggregated. There has been increasing interest in forecasting the demand for individual hotels based on hotel-specific data (e.g. Ellero and Pellegrini, 2014; Koupriouchina *et al*., 2014). The forecasts for individual hotels will benefit hotel practitioners with operational policy implementation such as reservations by higher-value customers, price discrimination, overbooking policies, late cancelations, and early departures (Koupriouchina *et al*., 2014).

3.3 New explanatory variables

The selection of tourism demand's determinants is far more diverse than its measurement, given the various research objectives of different studies. According to the neoclassical economic theory, price and income are the two key influencing factors of demand for a product. In empirical studies, tourists' income, tourism prices in a destination, and substitute prices in substitute destinations are most often used to explain and predict tourism demand. Tourists' income is expected to influence tourism demand positively and is often measured by the gross domestic product (GDP). Other proxies include the industry production index (Goh *et al*., 2008) and gross disposable income (Onafowora and Owoye, 2012). Tourism prices in a destination are expected to negatively influence tourism demand and are often measured by the relative consumer price index (CPI) between destination and origin, adjusted by exchange rates. Substitute price refers to the tourism price at a substitute destination or a group of substitute destinations, which is often measured by the CPI of the substitute destination or a weighted average of the CPIs of a group of substitute destinations. A positive coefficient thus estimated indicates a substitute relationship, whereas a negative coefficient indicates a supplementary relationship between destination and substitutes.

Other traditional determinants include transportation cost, which is often measured by oil price, advertising expenditure, exchange rate, volume of trade between origin and destination, population in the origin market, unemployment rate, and other social, cultural, geographic, and political factors. In addition, dummy variables are used to capture the impact of seasonality and unique occurrences such as the outbreak of diseases, terrorist attacks, and the Olympics on tourism demand.

When analyzing hotel demand from the macro-level perspective (i.e. demand for hotel accommodation in a destination), the important determinants for hotel demand are similar to those that affect the demand for tourism, which include tourist/guest income, destination tourism price, substitute tourism price, exchange rates, transportation cost, one-off events, and seasonal variables. Other key determinants such as room rate, unemployment rate, inflation rate, money supply, industrial production growth, and stock market return have also been examined (Chen, 2013; Singh *et al.*, 2014).

The above mentioned economic variables still dominate recent studies on econometric modeling and forecasting of tourism and hotel demand. Meanwhile, new explanatory variables have appeared in recent empirical studies and some are particularly strong in explaining tourism and hotel demand trends and changes. These include climate variables and tourist online behavior variables, among others.

3.3.1 Climate variables

Climate is considered to affect tourism and hotel demand in the long term due to tourists' preference for particular climates. This variable is relatively stable and has not shown the high variations required for tourism demand modeling. Climate is therefore seldom considered in earlier studies on tourism and hotel demand modeling and forecasting. Due to increasing concerns about climate change and increasing research interest in climate issues and their impact on tourism, however, some recent empirical studies have included climate variables in tourism and hotel demand models and have identified a significant impact on tourism and hotel demand. The inclusion of temperature alone as a determining climatic variable has tended to be widespread. Nonetheless, it is recognized that temperature alone does not fully represent a destination's climate. There are other climate variables such as relative humidity, heat waves, frost days, sunshine duration, and seasonal variations (Rosselló-Nadal *et al.*, 2011).

One challenge for the inclusion of climate variables into the tourism and hotel demand modeling process is that the relationship between climate variables and tourism and hotel demand may be nonlinear (Rosselló-Nadal *et al*., 2011) or particularly present an inverted u-shape, indicating the existence of an optimal climate for tourist preference. One solution is to establish the tourism climate index (TCI) as a determining variable. The TCI was initially proposed by Mieczkowski (1985) and has been applied in empirical studies of tourism demand by Amelung and Moreno (2012), Eugenio-Martín and Campos-Soria (2011), and Goh (2012). Unlike objective measures of climate such as temperature or humidity, the TCI is a measure of tourist perception of climate comfort and is often measured by a combination of sub-indices. The advantages of the TCI are that tourists' perception of climate comfort can be measured and its impact on tourism and hotel demand is expected to be linear and can be examined directly. The disadvantages are that the TCI cannot identify the optimal physical climate preferred by tourists or how physical climatic conditions affect tourist behavior.

3.3.2 Tourist online behavior variables

As online consumer behavior data have become increasingly available to researchers, the latter have recently started to use such data in conjunction with traditional economic data to improve forecasting performance (Yang *et al.*, 2014b). In the field of tourism and hotel forecasting, two types of online data have been employed: search query data and web traffic data.

Today consumers tend to use search engines such as Google to find travel and accommodation information before they purchase holidays. According to Yang *et al*. (2015), in 2012, 85% of the Americans used the Internet and 91% of those used search engines to find information; while 40% of the people in China used the Internet with 80% of those used search engines to find information. Millions of people utilize online search engines to seek destination-related information as well as to plan their trips. Additionally, Google Trends provides public access to the search data for specific queries on Google. Researchers are beginning to analyze the potential value of these search data to tourism forecasters. Practice indicates that online search engine data provide new insights into tourism and hotel demand forecasting. Bangwayo-Skeete and Skeete (2015) used a mixed-data frequency modeling technique, namely the autoregressive mixed-data sampling (AR-MIDAS) model, to examine the forecasting ability of weekly search query data in predicting monthly overnight tourist arrivals in five Caribbean countries from the US, UK, and Canada. The results show that Google search data significantly improve the forecasting accuracy over benchmark models of seasonal autoregressive integrated moving average (SARIMA) and autoregressive (AR) models. Pan *et al*. (2012) used search volume data from Google Trends on five related queries to predict the demand for hotel rooms in a specific city. An accuracy comparison between three autoregressive moving average (ARMA) family models and their ARMAX counterparts (i.e. ARMA models augmented with search volume data as an explanatory variable) indicates the usefulness of these data in improving forecasting performance. Besides Google search data, Yang *et al*. (2015) also examined the value of Baidu (the largest search engine in China) data and demonstrated the potential of these data in improving forecasting accuracy.

Apart from engine query data, Yang *et al.* (2014b) evaluated the forecasting performance considering the destination marketing organization's web traffic data. In this study, two traffic data are used: the number of users (identified by cookies) who accessed a specific website and the number of visits to the organization's website. The results show that web traffic volume data of a destination marketing organization are capable of improving the accuracy of hotel demand forecasts for a destination.

Compared with the traditional economic variables such as tourists' income and tourism prices, search engine data have their own advantages to generate forecasts of tourism and hotel demand. They are usually free of charge and real-time, which allows forecasters to predict demand betimes. These data are of high frequency, often generated on a daily basis, which allows high-frequency forecasting of tourism and hotel demand. These data are also a direct measure of tourist behavior, and thus are sensitive to changes of tourist behavior. These advantages make such online data an effective supplement to conventional determinants, and studies considering this kind of online data are encouraged in the future.

3.3.3 Other new explanatory variables

In reality, not all variables can be included in a single model because of data availability and research purposes, as well as the consideration of the degrees of freedom for model estimation. Therefore, researchers have attempted to find appropriate determinants of tourism and hotel demand and their optimal proxies according to particular research objectives. For example, Yang *et al.* (2014a) measured relative income using the distance between individual income and the average income of a city/province and identified the significant effect of the variable. Goh *et al*. (2008) incorporated a leisure time index and a climate index into monthly demand forecasting and found that they have a stronger impact on tourist arrivals than economic factors. Lee *et al*. (2010) and Song *et al*. (2012b) revealed that visa restriction has a significant negative impact on inbound tourist flows, using South Korea to Japan, China inbound and Hong Kong inbound as their cases. Using a gravity model, Balli *et al*. (2013) further identified that both the international export of Turkish soap operas and termination of the Turkish government's visa requirement policy have increased tourist inflow to Turkey.

Gounopoulos *et al.* (2012) examined how unemployment and consumer confidence indicators affect demand in Greece. Claveria and Datzira (2010) tested whether consumer confidence indicator is able to improve forecasting accuracy, with mixed results. Guizzardi and Stacchini (2015) introduced subjective supply-side information, such as business sentiment indicators, to forecast hotel guest numbers in Rimini, Italy. These business sentiment indicators were obtained from surveys of hotel owners' and managers' opinions and expectations about the performance of their own hotels and of the market as a whole. The empirical results showed that this subjective information can improve forecasting accuracy over time series models that do not contain such information.

The inclusion of these various explanatory variables into demand models enriches the tourism and hotel demand analysis. The results provide new insights into tourist behaviors and useful management implications for relevant practitioners. However, when some non-traditional variables are included in the models, their effects on tourism and hotel demand should be supported by solid theoretical justifications and verified by statistical testing rather than being tested on an *ad hoc*, trial and error basis. Researchers are therefore encouraged to consider new non-traditional variables to explain tourism and hotel demand with the support of theories from different disciplines.

3.4 Data frequency

As noted in Section 2, a large number of studies have used annual data for tourism and hotel demand modeling and forecasting exercises. The focus of these studies is normally long-term relationships between tourism (or hotel) demand and its influencing factors, and/or medium- to long-term trend forecasting. Using annual data removes the seasonal variability in a tourism (or hotel) demand model; the disadvantage of this is that such an analysis cannot capture the seasonal characteristics or predict seasonal variations of the demand. If the latter are the focus of a study, seasonal data are employed, including quarterly and monthly data, where seasonality needs to be considered during the modeling process. The straightforward and traditional approach to dealing with seasonality is to include seasonal dummies in the model, in which seasonality is treated as deterministic. However, this is an overly restrictive assumption, especially when lengthy time series are considered. Empirical evidence shows that seasonal patterns vary over time (Song and Li, 2008). Hence, some recent studies treat seasonality as stochastic by identifying and eliminating the seasonal unit roots before building a model or decomposing the demand series into a few unobserved components, including the seasonal component, and then specify them in a structure time series model (e.g. Guizzardi and Stacchini, 2015; Song *et al.*, 2011a).

Although more appropriate treatments of seasonality tend to improve the accuracy of seasonal tourism and hotel demand forecasting, business needs cannot be fully satisfied by quarterly or monthly predictions, given the increasingly dynamic nature of the demand system and the growing trend of late booking. Some businesses, such as hotels and airlines, may require forecasts of an even higher frequency, such as weekly or even daily. Some studies have then started to use weekly data (e.g. Yang *et al*., 2014b; Zheng, 2014) or daily data (e.g. Diaz and Mateu-Sbert, 2011; Divino and McAleer, 2010; Medeiros *et al*., 2008) for tourism and hotel demand analysis. Accurate forecasting and analysis based on these high-frequency demand data is especially helpful in planning and scheduling day-to-day operations and achieving higher yield levels through improved matching of demand with capacity. In this situation, time series models are often the option for analysis since some explanatory variables such as price and income are generally unavailable as high-frequency data. Another emerging option is to employ mixed-frequency modeling techniques which allow the inclusion of variables with different frequencies in demand models. Section 4.2 will offer a detailed discussion.

4. Methodological development

It is observed that non-causal time series models, causal econometric approaches, and artificial intelligence-based methods still dominate the tourism and hotel demand forecasting field. In particular, some advanced models, such as the almost ideal demand system (AIDS) and panel data analysis, have received wider application or been introduced to this field and demonstrated their superiority over certain of the traditional methods. In addition, the combination of different techniques has continued to be a key direction of methodological development.

4.1 Non-causal time series methods

Some traditional, commonly used univariate time series models continue to appear in recent studies, including the no change (Naïve I) and constant growth rate (Naïve II) models, different exponential smoothing (ES) models (such as double ES and Holt-Winters ES), ARMA family models (such as ARIMA and SARIMA) (Tsui *et al*., 2014), and the structural time series (STS) model (e.g. Gounopoulos *et al*., 2012). Some of these are often used as benchmark models for accuracy comparison. In the meantme, more new and sophisticated time series methods have emerged in recent studies. For instance, Chu (2009) introduced an autoregressive ARMA (ARARMA) model and a fractionally integrated ARMA (ARFIMA) model to forecast tourist arrivals to nine tourist destinations in the Asia-Pacific region. Unlike the ARIMA model which transforms data by differentiating them, the ARARMA model identifies the transformation by an autoregressive process. On the other hand, the ARFIMA model allows the series to contain fractional order of integration. The empirical results show that the ARFIMA model is superior to the other two ARMA-based models, SARIMA and ARARMA. Assaf *et al*. (2011) also disclosed the fractional degrees of integration in a series of tourist arrivals in Australia and verified that models based on both non-seasonal and seasonal fractional integration outperformed the standard ARIMA and SARIMA models, respectively.

Athanasopoulos and Hyndman (2008) and Athanasopoulos *et al*. (2011) introduced innovations state space (ISS) models for exponential smoothing, which encapsulate the notion of exponential smoothing in a state space framework and allow for maximum likelihood estimation. Although both STS and ISS models are specified in state space form, they deal with the error term of each equation of a state space model differently. ISS only involves a single source of error, while the STS model allows each equation to carry its own independent error term. From the model estimation point of view, the ISS method is more efficient. Athanasopoulos *et al*. (2011) presented empirical evidence of the ISS model's superior forecasting performance in a broad range of tourism forecasting competition exercises.

Time-varying conditional variance is also identified in tourism demand data series. The convention is to apply the autoregressive conditional heteroscedasticity (ARCH) technique to model the demand for tourism/hotel rooms. For example, Divino and McAleer (2010) used generalized ARCH (GARCH) and exponential GARCH to model the growth rate of daily arrivals to Peru. Toma *et al*. (2009) examined the impact of the release of a best-selling book and movie, *Midnight in the Garden of Good and Evil*, set in Savannah, Georgia on the local tourism demand using the ARIMA-ARCH model.

Very recently, a nonparametric forecasting technique, the singular spectrum analysis (SSA), has been introduced into the tourism literature (Beneki *et al*., 2012; Hassani *et al*., 2015). Assuming that a time series consists of signal and noise, unlike traditional time series models which forecast both signal and noise, SSA aims to filter the noise and forecast the signal only (Hassani *et al.*, 2015). Similar to a STS model, SSA decomposes a time series into independent components such as trend, seasonal and business cycle components but, as a nonparametric method, SSA is model-free and data-driven, making no assumptions about the data-generating processes. The above empirical studies, Beneki *et al*. (2012) and Hassani *et al*. (2015), showed that SSA outperforms other time series models such as ES, SARIMA, STS, and a neural network model. So far, only the univariate version of SSA has been applied to tourism and hotel demand forecasting; a multivariate version of SSA has been developed recently, but no empirical work has been carried out to examine its forecast accuracy in the tourism context. Furthermore, researchers should consider other spectral methods such as multi-taper methods and maximum entropy (Ghil *et al*., 2002) and compare their performance against other tourism forecasting methods.

Other nonlinear time series models, such as the self-exciting threshold autoregressive model (Claveria and Datzira, 2010; Claveria and Torra, 2014) and the Markov-switching model (such as Chen, 2013; Valadkhani and O'Mahony, 2015), have also attempted to forecast tourism and hotel demand. To give an instance, Claveria and Datzira (2010) applied both models to forecast tourism demand in France, the UK, Germany, and Italy with two simple time series models (AR and ARIMA) as benchmarks. The results showed that the ARIMA and Markov-switching models outperform the other two.

Another research trend observed is that researchers attempted to extend the non-causal time series models into the econometric framework by augmenting them with additional explanatory variables. Athanasopoulos and Hyndman (2008) found that combining the ISS model with exogenous variables captures time series dynamics well and outperforms the regression models. Song *et al.* (2011a) combined the STS and the time-varying parameter (TVP) technique to forecast quarterly tourist arrivals and demonstrate superior forecast accuracy over six time series and econometric competitors. Guizzardi and Stacchini (2015) incorporated business sentiment indicators in naïve and STS models and noted that forecasting accuracy was improved.

4.2 Econometric methods

Policymakers in tourist destinations, especially those where tourism is the major source of foreign exchange earnings, have made great efforts to understand the key determinants of demand for their tourism products and predict future trends in order to formulate the most effective policies and strategies. Such objectives cannot be achieved by non-causal time series analysis, and so there has been continuous interests in econometric studies of tourism and hotel demand in the past few years. A number of modern econometric models reviewed by Song and Li (2008), especially the dynamic versions of these models, have continued to appear in recent studies under different empirical settings. These models include the autoregressive distributed lag (ADL) model (e.g. Onafowora and Owoye, 2012; Song and Lin, 2010; Song *et al.*, 2012b), the error correction (ES) model (e.g. Goh, 2012; Smeral, 2010) and the VAR model (e.g. Torraleja *et al.*, 2009). In addition, the AIDS model, one of the most theoretically sound approaches to demand, has demonstrated its potential for broader application in tourism. Studies using the AIDS models prior to 2007 mainly focused on the substitution and complementary relationships between tourist destinations. In more recent studies, the AIDS models have been applied to examine the substitution and complementary effects between different tourism consumption categories such as accommodation, restaurants, and shopping (e.g. Wu *et al.*, 2011; 2012a), or the substitution effect between domestic tourism and outbound tourism (e.g. Athanasopoulos *et al.*, 2014). Furthermore, Li *et al.* (2013) extended the AIDS model to examine the competitiveness of an international destination *vis-à-vis* its competitors.

Some lesser-used methods in the pre-2007 studies have started to gain popularity in the more recent literature. For example, Song and Li (2008) only identified four studies that used panel data analysis during the period 2000–2006, while more than a dozen have employed this technique since 2007 (e.g. Falk, 2013; Garín-Muñoz and Montero-Martín, 2007; Gholipour *et al*., 2014; Yang *et al*., 2014a). Panel data analysis incorporates information from both time series and cross-sectional dimensions and is therefore especially efficient when the time series are short but cross-sectional data are available. Besides, the panel data analysis offers a greater degree of freedom in model estimation and reduces the multicollinearity problem (Serra *et al.*, 2014).

Another trend in tourism and hotel demand analysis is the application of spatial econometric models. Although gravity models have been applied in tourism and hotel demand analysis to measure the effect of distance between an origin and a destination on tourism flows, this technique assumes independence among tourist flows once the effect of distance is controlled for (Marrocu and Paci, 2013). This assumption is restrictive and the spatial spillover effect is beyond the consideration of a gravity model. An alternative approach, spatial econometric modeling, takes origin-destination dependence into account and is able to capture the spatial interaction in the modeling process. Given their advantages, a growing interest in spatial econometric techniques has emerged in the recent literature. Marrocu and Paci (2013), for example, employed a spatial autoregressive model to discover the importance of spatial dependency induced by neighboring provinces by analyzing domestic tourism flows for a complete set of 107 provinces of Italy. Deng and Athanasopoulos (2011) applied a spatial lag panel model to capture both temporal and spatial dependence of tourism demand systems based on 83 local areas of Australia. Spatial econometric techniques offer a new perspective from which the changing characteristics of tourism and hotel demand system are examined. The world is increasingly interconnected and it is easier for tourists to move across multiple countries to experience different cultures in a single trip. Spatial econometric models can determine the interdependence of destinations in a region and help governments to establish cooperation through visa access, or help businesses to formulate joint marketing campaigns across borders. Further applications of this approach in tourism and hotel demand analysis are recommended to supply valuable empirical evidence for relevant strategic decision-making.

Despite extensive research into the econometric analysis of tourism and hotel demand, most studies have examined the relationship between tourism (or hotel) demand and its economic determinants under the assumption of a linear relationship. Thus the determinants, such as tourist income or tourism prices, are assumed to have an impact of constant degree on tourism (or hotel) demand over time, which is highly restrictive and does not reflect the reality, given tourists' changes in their preference and attitude. In this view, the TVP model can be regarded as a nonlinear modeling technique since the coefficients are allowed to vary over the sample period in order to trace the evolution of the tourism (or hotel) demand system over time. The TVP technique has been applied to tourism demand analysis (e.g. Page *et al.*, 2012) and in conjunction with other advanced econometric techniques to develop more sophisticated models such as the TVP-STS (Song *et al.*, 2011a) and the restricted TVP-EC-AIDS model

(Wu *et al.*, 2012a).

Though the TVP models can examine the evolution of the impacts of determinants on demand over time, it cannot identify this evolution over different scales of determinants. Economic theory indicates that when economic factors are on different scales, their impact on the demand system may also change. As an illustration, when price is on a higher scale, its impact on demand may be stronger than in cases where price is on a lower scale. Under this circumstance, an alternative nonlinear technique, the smooth transition regression (STR) model, is able to capture the deterministic structural change in a time series regression. In an STR model, the transition between regimes is allowed to take place smoothly over time. In each of the regimes, the demand system can be described adequately by a linear model. In spite of the technical advantages of the STR model and its wide applications in other fields, only one study has applied this method to tourism demand modeling. Wang (2014) applied a panel STR model to measure the impacts of income on tourism expenditures under different savings regimes and found that the effect is more pronounced in a low savings regime. The nonlinear characteristics of tourism (or hotel) demand system would benefit from further research, and the STR model is a useful tool for such analysis.

One more trend is the application of the mixed frequency techniques. In an econometric analysis, if the variables are measured in different frequencies, the conventional method is to transform the higher frequency data into lower frequency ones to keep all variables at the same frequency. An alternative solution is to apply mixed frequency techniques by which researchers can establish models whereby the data for different variables are in different frequencies. This is an effective way to avoid the loss of information included in the higher frequency data. Since more information is taken into consideration, mixed frequency techniques are assumed to describe tourist behavior more precisely and generate more accurate forecasts. The mixed-data sampling (MIDAS) approach has been applied in tourism forecast or nowcast (i.e. real-time forecast) when macroeconomic variables, such as GDP, are used as explanatories, which are often reported at a low frequency (quarterly often). MIDAS models estimate using the parsimonious distributed lag polynomials or nonlinear least squares method (Bangwayo-Skeete and Skeete, 2015).

Besides MIDAS, the mixed-frequency vector autoregressive (VAR) model proposed by Zadrozny (1988) is also well established in the econometric literature as a means to handle unbalanced datasets but has not yet appeared in tourism and hotel demand forecasting. The mixed-frequency VAR method treats all series as being generated at the highest frequency but considers those low frequency variables to be missing values. Given the fact that researchers often encounter the problem that available data are measured at different frequencies, the mixed-frequency techniques should be applied more to tourism and hotel demand forecasting with a view to avoiding information loss. Nowadays more high-frequency data are available, such as tourist online behavior data which contain rich information to describe and predict tourist behavior.

4.3 Artificial intelligence-based methods

The AI techniques have continued to be applied to tourism and hotel demand forecasting and empirical evidences have demonstrated their satisfactory performance.

Many of these studies are published in journals in other disciplines such as computing science and statistics. One possible reason is that the majority of these studies focus primarily on the methodological development and evaluation of forecasting accuracy rather than tourism-specific applications. Also, the setup of an AI-based model lacks strong theoretical foundation and it is difficult to measure the impact of economic factors on tourism and hotel demand using such models. These limit the application of AI-based models to tourism and hotel demand analysis and explain the scarcity of publications on AI methods in tourism and hospitality journals. The AI-based technique that has appeared most frequently in the recent literature is the artificial neural network (ANN) model. Other techniques, such as support vector regression (SVR), the rough set model, fuzzy system methods, genetic algorithms, and Gaussian process regression (GPR), have also been used in tourism and hotel demand forecasting but to a lesser extent.

The ANN, a nonparametric and data-driven technique, has attracted great interest due to its capability of mapping linear or nonlinear function without any assumption imposed by the modeling process. ANN simulates biological neural systems, especially human brains, by including input, hidden, and output layers; each layer containing one or more neurons. These neurons are interrelated in the process of information processing and computing (Cuhadar *et al.*, 2014). Different ANN models have been applied to tourism and hotel forecasting practice, including multi-layer perceptron (MLP), radial basis function (RBF), generalized regression neural network (GRNN), and Elman neural network (Elman NN). MLP is the most widely used ANN model; it contains three or more layers of neurons with nonlinear activation function (e.g. Chen *et al.*, 2012; Claveria and Torra, 2014; Lin *et al.*, 2011). As an alternative, an RBF network contains only one hidden layer and does not need to deal with local minimums but approximates the best solution directly. The RBF training process is shorter than that of the MLP network. Applications include Cang (2014), Claveria *et al.* (2015a), and Cuhadar *et al.* (2014). GRNN is similar to the RBF network, being based on kernel regression. Cuhadar *et al.* (2014) employed GRNN to forecast cruise tourism demand to Izmir, Turkey. Elman NN contains both a three-layer network and a set of context units, and the context units and the hidden layer are connected for processing and computing the information (e.g. Claveria *et al*., 2015b).

Another AI-based model is the SVR. Unlike ANN which adopts the empirical risk minimization principle, SVR minimizes training error by implementing the structural risk principle. SVR solves linear regression problems by nonlinearly mapping the input data to a high-dimensional space. Theoretically, SVR is able to achieve a global optimum, rather than obtaining trapped optima like an ANN model (Hong *et al.*, 2011). SVR has been applied to tourism and hotel forecasting by several studies (e.g. Cang, 2014; Chen and Wang, 2007; Hong *et al*., 2011; Xu *et al*., 2009).

The fuzzy system model is suitable in circumstances where data are linguistic terms or comprise less than 50 data points (Tsaur and Kuo, 2011). Different versions of the fuzzy system model are used for tourism and hotel demand forecasting. For example, Aladag *et al*. (2014) employed a seasonal fuzzy system model to forecast international tourism demand in Turkey. Chen *et al*. (2010) applied the adaptive network-based fuzzy inference system model to forecast tourist arrivals to Taiwan and demonstrated its superior forecasting performance over the fuzzy time series model, grey forecasting model, and Markov residual modified model.

The fuzzy system model is often combined with genetic algorithms, another AI-based technique, to compute data. The idea of genetic algorithms derives from the evolutionary theory of natural selection and genetics. A hybrid method based on the fuzzy system and genetic algorithms has been used by several studies (e.g. Hadavandi *et al.*, 2011; Shahrabi *et al*., 2013; Tsaur and Kuo, 2011). Genetic algorithms have also been applied to a SVR model (e.g. Chen and Wang, 2007; Hong *et al*., 2011). Pai *et al*. (2014) further incorporated the fuzzy system, SVR technique, and genetic algorithms into a new model which has demonstrated superior forecasting performance over a number of other models.

The rough sets model has also been applied to tourism demand forecasting since 2007. For example, Goh *et al*. (2008) applied it to forecast the long-haul demand for Hong Kong tourism among residents of the US and UK. Based on the classical set theory, the rough sets model can handle vague and imprecise data by replacing them with precise lower and upper approximations. The model focuses on generating decision rules on the basis of a list of conditions.

Furthermore, Wu *et al*. (2012b) introduced a new machine learning method, the sparse GPR model, for tourism demand forecasting. GPR uses a nonparametric technique for regressions in high dimensional spaces provides uncertainty estimations, and learns the noise and smoothness parameters from training data. Sparse GPR is capable of reducing the computational complexity of the basic GPR model.

Given the different advantages of these AI-based methods, researchers have done substantial work applying them to forecasting performance and achieved satisfactory results. Even so, one of the limitations of these methods is that the underlying relationships between different variables are unknown, which restricts their applications to impact analysis on demand. A possible future research direction could be to uncover some rules for the nonlinear relationships between the demand variables and their determinants using AI-based techniques.

Although different new methods, whether the non-causal time series ones, multivariate econometric ones, or AI-based ones, have been introduced constantly into tourism and hotel forecasting practices, there is a consensus that no one model can perform best consistently in all conditions, and across all data characteristics and study features such as time period, origin/destination pairs, measurements of tourism demand, purpose of trip, forecast horizon, sample size, and data frequency. All these features may affect the forecasting accuracy of tourism and hotel demand models. Employing the mega-regression analysis, Kim and Schwartz (2013) and Peng *et al.* (2014) empirically verified this finding by examining 32 and 262 studies respectively. The latter also provides suggestions for the choice of appropriate forecasting methods when dealing with different data characteristics. In the future, more evidence is required to identify the forecasting performance of specific models and to highlight the connection between the study features and the models' performance with the aim of providing practical suggestions to tourism and hotel forecasting practitioners.

5. Other new research interests

5.1 Interval estimation and interval forecasting

Point estimation and point forecasting have dominated recent tourism and hotel demand literature. Point estimation gives a single value of the parameter of interest. Similarly, a point forecast is “a single number which is an estimate of the unknown true future value” (Kim *et al*., 2011, p. 888). Point estimates and forecasts do not provide any information as to the degree of variability or uncertainty associated with the estimate or forecast (Kim *et al.*, 2011). Interval estimation and forecasting, on the other hand, are able to overcome this limitation by providing a range instead of a single value for the estimate, given a specified level of confidence. Such an interval provides more useful information to industry practitioners and policymakers and allows them to formulate policies and strategies with more confidence. Interval estimation and interval forecasting have been introduced to tourism and hotel demand studies lately, although the applications are still limited. Song *et al.* (2010a) provided interval estimates of the elasticities of tourism demand in Hong Kong. Kim *et al.* (2010) proposed the use of the bias-corrected bootstrap for interval forecasting of an autoregressive tourism demand series and showed desirable small-sample properties of the proposed interval forecasting method. Kim *et al.* (2011) further evaluated the performance of tourism forecast intervals generated from alternative time series models and found that most models produce satisfactory prediction intervals, and that those based on the bias-corrected bootstrap perform best in general. Bermúdez *et al*. (2009) generated both point and interval forecasts for hotel occupancy in three provinces of Spain based on the Bayesian-based multivariate Holt-Winters model with additive seasonality and errors.

In contrast to point forecast error measurement, interval forecasting often employs coverage rate and interval width for the measurement of forecasting accuracy. Coverage rate refers to the percentage by which the actual demand falls into the prediction intervals; and width refers to the mean width of the prediction intervals. Good interval forecasts offer a coverage rate close to the nominal coverage rate, such as 95% or 99%. When the coverage rates of interval forecasts from two models are equivalent, the model with narrower or tighter width is assumed to have superior forecasting property (Kim *et al.*, 2011).

5.2 Forecast combination and adjustment

Clemen (1989) demonstrated that combining forecasts generated by alternative forecasting models through certain combination methods generally leads to improvement of forecasting accuracy. However, the application of forecast combinations to the tourism context was rare until recently. According to Shen *et al.* (2008), only three studies on tourism forecast combination were published before 2006. More applications of combined forecasting techniques emerged in the tourism literature in the period 2007–2015. It is observed that the individual models to be combined vary, from time series such as ARIMA to the more advanced econometric models such as EC, ADL, VAR, and TVP models (Shen *et al.*, 2011; Song *et al.*, 2009). With reference to the combination methods, in addition to the simple average, in which equal weighting is imposed on the individual forecasts to generate the final forecasts, more sophisticated techniques in which different weighting schemes are applied to each individual forecasting model according to the historical performance of the individual methods. More weighting is given to the forecasts of the models which have produced relatively more accurate individual forecasts in the past.

Examples of these techniques include the variance-covariance method, the discounted mean square forecast error method, the shrinkage method, the Granger and Ramanthan regression method, and the TVP combination method. More recently, AI-based techniques have been used to determine how individual forecasts are combined. For example, Cang (2014) combined individual time series forecasts based on two ANN models and one SVR model and empirical results showed that the combined forecasts based on the three AI-based techniques generate satisfactory forecasting performance.

Regarding the performance of combined forecasts it is generally accepted that the combination of forecasts from different forecasting techniques can help to improve forecasting accuracy. Particularly, Wong *et al*. (2007) demonstrated that combination forecasts cannot beat the best single forecast but always perform better than the worst single one. Hence, it is less risky to adopt combined forecasting techniques. Shen *et al.* (2011) further proved that combined forecasts generally outperform the best single forecast involved. Song *et al.* (2009) provided statistical evidence that although combined forecasts cannot beat the best single forecast, their forecasting accuracy is significantly higher than the average accuracy of single forecasts involved. Andrawis *et al.* (2011) later combined the forecasts derived from tourism demand data to capture information of time series with different frequencies. Their results showed that forecast combination performs better than individual models. Given the potential to reduce forecasting risks and improve forecasting accuracy, more discussions on combination forecasting, such as selection criteria of individual models for the pool, optimal numbers of individual models to be combined, and innovative combination methods, should be considered in future studies. Another direction is to examine the performance of forecast combinations in interval forecasting which has not been studied so far.

In addition to the attempts at combining forecasts generated by different statistical models, there has been recent interests in integrating quantitative methods with qualitative approaches such as expert judgement (e.g. Croce and Wöber, 2011; Lin, 2013; Lin *et al.*, 2014). Judgmental adjustment of statistical forecasts is one of the notable alternatives for integrating statistical and judgmental approaches. Forecasts based on quantitative techniques are produced first, and are then distributed to expert panels for adjustment based on their professional judgment. Following the Delphi procedure, the final forecasts may contain both information from the quantitate methods and judgment from the experts. Using a web-based forecasting system and the Delphi method, Lin *et al.* (2014) invited 11 academics and practitioners to make judgmental adjustment of the forecasts derived from econometric techniques and identified that such adjustment of statistical forecasts can effectively improve the forecast accuracy. Lin (2013) also noted that on average the adjusted forecasts are unbiased, though the adjusted forecasts are not always unbiased when individual markets are examined separately. More in-depth analysis should be conducted in this arena to enhance the accuracy and stability of judgmental forecasting.

5.3 Development of web-based tourism and hotel demand forecasting systems

The rapid development of Internet technology has allowed researchers to build a web-based tourism and hotel demand forecasting system (TDFS), which is defined as "a computerized information system that delivers tourism demand forecasts and provides decision support to policymakers and business strategists via a Web browser" (Song and Li, 2008, p. 446). A web-based TDFS offers an effective bridge between academics and industry practitioners. As a computer-based innovation, web-based TDFS often includes the following functions (Croce and Wöber, 2011; Song *et al*., 2013): (1) systematic storage of a broad range of tourism and hotel demand variables and their determinants, which is demonstrated in user-friendly ways such as graphs and tables; (2) application of quantitative forecasting techniques to generate forecasts for tourism and hospitality; (3) incorporation of forecasters' judgement to adjust demand forecasts derived from the statistical model; and (4) generation of forecasts under different scenarios as requested. Such a web-based TDFS can provide enormous benefits to various stakeholders and support their evidence-based decision-making processes. Web-based TDFS development is still in its early stages and further improvements are necessary. For example, interval forecasts under different nominal coverage rates could be offered and industry practitioners could be more engaged in the process of judgmental adjustment. Furthermore, more forecasting models could be included and combined to generate more stable statistical forecasts and further reduce the risk of forecasting failure.

6. Concluding remarks

6.1 Conclusions

Tourism demand analysis continues to dominate tourism economics studies in terms of research interests and methodological advancements (Song *et al*., 2012a). This review of recent studies identifies a broader research scope. With regard to the diversity of research interests, studies focusing on hotel demand are relatively less than those focusing on tourism demand. It is also observed that more and more studies have moved away from the aggregate tourism demand analysis, while disaggregate markets and niche products have attracted increasing attention. Some studies have gone beyond neoclassical economic theory to seek additional explanations of the recent dynamics of tourism and hotel demand, such as environmental factors, tourist online behavior, and consumer confidence indicators, among others.

Referring to variables, different explanatory ones have been introduced to the tourism and hotel modeling process, such as climate variable, consumer confidence indicators, and business sentiment indicators, amongst others. In particular, the development of Internet technologies provides researchers with newly emerging online data such as engine queries and web traffic data. Empirical studies have also demonstrated that these data are very useful in improving the accuracy of forecasts of tourism and hotel demand. Due to the advantages of using these data which are real-time, high-frequency, and directly measure tourist behavior, further efforts are necessary to improve the performance of the forecasting models by incorporating data of tourist online behavior with traditional, low-frequency economic indicators.

Methodologically, greater diversity has been observed in the range of techniques applied to the domain. Regarding the non-causal time series techniques, new time series models such as nonparametric SSA, the self-exciting threshold autoregressive model, and the Markov-switching model have started to appear in the literature in addition to the traditional methods such as ARIMA, ES, STS, and GARCH. Another trend is the extension of non-causal time series models into the econometric framework by augmenting them with additional explanatory variables. It has been demonstrated that the extension will improve the forecasting performance of tourism and hotel demand models. Another advantage is that the impact of some interventions on tourism and hotel demand can be captured based on the time series techniques.

In terms of econometric methods, new methods have been introduced into tourism and hotel demand analysis apart from those widely used models such as EC, VAR, TVP, and AIDS techniques,. The new methods include models such as the nonlinear STR models, mixed frequency models, and spatial econometric models. The STR technique is capable of identifying the nonlinear relationship between tourism and hotel demand and their determinants. Mixed-frequency modeling technique provides the possibility of including variables with different frequencies in the demand model and its performance deserves further examination. Also, the AIDS model and spatial econometric techniques should be used further in tourism and hotel demand modeling and forecasting given the fact that the demand for different tourism (or hotel) products are interrelated.

Furthermore, the forecast combination technique is an efficient way to avoid serious forecasting failure, given it is widely admitted that no single model can outperform other models in all conditions. Recently, different forecast combination techniques aiming at identifying optimal weights have been introduced and examined in tourism demand forecasting exercises. Empirical results show that forecast combination is able to reduce forecasting risks and improve forecasting accuracy. Besides, judgmental adjusted forecasting is also verified to be able to enhance forecasting performance.

6.2 Theoretical implications

By reviewing the relevant studies published during 2007–2015, this study identifies new trends in tourism and hotel demand modeling and forecasting.

From a methodological perspective, additionally new and innovative models from other disciplines have been introduced to tourism and hotel demand forecasting which contributes to the advancement of tourism forecasting methodologies. For example, Athanasopoulos and Silva (2012) developed a new set of forecasting models dealing with local level and trend, and damped trend with an additive multivariate seasonal components to forecast the demand for tourism. Another example is Shahrabi *et al.* (2013) who proposed a modular genetic-fuzzy forecasting system by combining genetic fuzzy expert and data preprocessing systems. These studies contribute not only to the tourism and hotel demand forecasting literature but to the study of generic forecasting also.

In addition, the development of new forecasting methods has facilitated a better understanding of tourist behavior which in turn provided useful insights for the development of effective tourism demand forecasting systems. For example, studies that employed the STR models allow us to identify nonlinear characteristics of tourist consumption whilst those using the AIDS models explore the substitution effect when tourists choose a destination or a product amongst a number of alternatives.

Moreover, recent studies on tourism demand modeling and forecasting have incorporated subjective variables such as consumer confidence and/or business sentiment indicators in the forecasting models. A growing number of tourism forecasting studies has also used expert judgment to enhance forecasting accuracy. These efforts have clearly demonstrated that the research on tourism forecasting has developed beyond the traditional economic modeling frameworks.

6.3 Practical implications

Tourism and hotel demand modeling and forecasting is directly related to tourism and hotel management practices. The research findings of the published studies provide important suggestions and recommendations on improving the efficiency of tourism and hospitality practices. Closer engagement with key stakeholders will greatly benefit both academic research and tourism practice. The development of web-based forecasting systems is a good example of engaging scientific research in combination with relevant stakeholders.

One of the trends emerged from recent studies is that more attention has been paid to the demand for niche tourism products such as ski tourism and cruise tourism. These studies have made useful information and future directions available for business decision-makers related to these markets. For example, the price elasticity analysis of the demand for such products can help ski and cruise businesses to formulate appropriate pricing strategies. Accurate forecasts of future tourism demand in destinations will help destination governments and businesses to allocate limited resources more effectively and efficiently. Tourism destinations/businesses will be more willing to use their resources on promotions if an overwhelming future demand is forecasted.

Moreover, the discovery of new explanatory variables in the demand modeling process also benefits industry practitioners. Take tourist online behavior variables as an example, their use in tourism and hotel demand modeling and forecasting can help tourism businesses to identify the relationship between the online behavior and actual behavior of tourists. Once this relationship is recognized, businesses can generate accurate forecasts in real time and make prompt operational decisions such as staffing and inventory adjustments. Based on the high-frequency data overserved online, particularly, hotels can adjust their daily demand predictions in near-real time and achieve revenue management objectives. Public event organizers and local authorities can also make effective use of these online search data in real-time or very short-term visitor forecasting in order to support crowd management, such as by providing sufficient facilities and a safe and orderly environment for the events.

6.4 Limitations and future research directions

This review identified the significant theoretical and practical contributions of recent studies to tourism demand modeling and forecasting. The limitations identified in this review form a ground for future research.

Firstly, it is observed that the diversity of the methods applied in studies of hotel demand studies is relatively limited compared with those of tourism demand. The application of advanced models, such as nonlinear modeling technique and dynamic systems of equation modeling technique, is still very rare. This finding is consistent with Mohammed *et al.* (2015) who suggested that more advanced modeling techniques should be used to identify the dynamics of the demand for hospitality products.

Secondly, though increasing interests are identified in demand analysis and forecasts for niche tourism products, data unavailability has limited the quantitative analysis on these demands. Studies focusing on such niche market as wine tourism and film tourism have not yet been seen in the literature and need to be encouraged once these data are available.

Thirdly, although some researchers have employed online data like search query data and web traffic data in forecasting tourism and hotel demands, there is still a huge potential for the use of such data in tourism and hotel demand forecasting.

Fourthly, even though other new non-traditional explanatory variables such as climate variables and consumer confidence indicators have been confirmed to have explanatory power in tourism and hotel demand functions, the theoretical justification for the use of these variables is still relatively weak. Accordingly, researchers are encouraged to employ theories from different disciplines rather than on an *ad hoc* and trial and error basis while considering these new variables to explain tourism and hotel demand in the future.

Lastly, though more diverse techniques have been applied to this area of study, there is still room for exploring new methodologies and applications in tourism demand modeling and forecasting. The AIDS model and spatial econometric techniques can further be explored, for example, given the fact that the world is increasingly interconnected and the demand for different tourism products or destinations are interrelated. Albeit the advantages of the mixed-frequency model, the introduction of the mixed-frequency VAR model has not yet been applied in this field thus its application is encouraged. Nonlinear modeling techniques, such as the STR model, are encouraged to be further applied to tourism and hospitality modeling and forecasting. Another possible future research direction is to develop hybrid models that combine the strengths of both econometric- and AI-based techniques, and uncover the rules for the nonlinear relationship between demand and its determinants. Meanwhile, the studies that generate interval forecasts are far from adequate although interval forecasts can provide industry practitioners more confidence in their decision-making. Specifically, the combination of interval forecasts has not been studied in the current literature and deserves considerable attention from researchers in the field of tourism and hospitality demand modeling and forecasting.

Acknowledgements

The authors would like to acknowledge the financial support of the National Natural Science Foundation of China (Grant No. 71573289) and the Research Grants Council of the Hong Kong SAR (PolyU 5969/13H).

References

Aladag, C.H., Egrioglu, E., Yolcu, U. and Uslu, V.R. (2014), "A high order seasonal fuzzy time series model and application to international tourism demand of Turkey", *Journal of Intelligent & Fuzzy Systems*, Vol. 26 No. 1, pp. 295-302.

Amelung, B. and Moreno, A. (2012), "Costing the impact of climate change on tourism in Europe: results of the PESETA project", *Climatic Change*, Vol. 112 No. 1, pp. 83-100.

Andrawis, R.R., Atiya, A.F. and El-Shishiny, H. (2011), "Combination of long term and short term forecasts, with application to tourism demand forecasting", *International Journal of Forecasting*, Vol. 27 No. 3, pp. 870-886.

Assaf, A.G., Barros, C.P. and Gil-Alana, L.A. (2011), "Persistence in the short- and long-term tourist arrivals to Australia", *Journal of Travel Research*, Vol. 50 No. 2, pp. 213-229.

Athanasopoulos, G., Deng, M., Li, G. and Song, H. (2014), "Modelling substitution between domestic and outbound tourism in Australia: a system-of-equations approach", *Tourism Management*, Vol. 45, pp. 159-170.

Athanasopoulos, G. and Hyndman, R.J. (2008), "Modelling and forecasting Australian domestic tourism", *Tourism Management*, Vol. 29 No. 1, pp. 19-31.

Athanasopoulos, G., Hyndman, R.J., Song, H.Y. and Wu, D.C. (2011), "The tourism forecasting competition", *International Journal of Forecasting*, Vol. 27 No. 3, pp. 822-844.

Athanasopoulos, G. and Silva, A. (2012), "Multivariate exponential smoothing for forecasting tourist arrivals", *Journal of Travel Research*, Vol. 51 No. 5, pp. 640-652.

Baggio, R. and Sainaghi, R. (2016), "Mapping time series into networks as a tool to assess the complex dynamics of tourism systems", *Tourism Management*, Vol. 54, pp. 23-33.

Balli, F., Balli, H.O. and Cebeci, K. (2013), "Impacts of exported Turkish soap operas and visa-free entry on inbound tourism to Turkey", *Tourism Management*, Vol. 37, pp. 186-192.

Bangwayo-Skeete, P.F. and Skeete, R.W. (2015), "Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach", *Tourism Management*, Vol. 46, pp. 454-464.

Beneki, C., Eeckels, B. and Leon, C. (2012), "Signal extraction and forecasting of the UK tourism income time series: a singular spectrum analysis approach", *Journal of Forecasting*, Vol. 31, pp. 391-400.

Bermúdez, J.D., Corberán-Vallet, A. and Vercher, E. (2009), "Multivariate exponential smoothing: a Bayesian forecast approach based on simulation", *Mathematics and Computers in Simulation*, Vol. 79, pp. 1761-1769.

Cang, S. (2014), "A comparative analysis of three types of tourism demand forecasting models: Individual, linear combination and non-linear combination", *International Journal of Tourism Research*, Vol. 16, pp. 596-607.

Cazanova, J., Ward, R.W. and Holland, S. (2014), "Habit persistence in air passenger traffic destined for Florida", *Journal of Travel Research*, Vol. 53 No. 5, 638-655.

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

Chen, K.Y. and Wang, C.H. (2007), "Support vector regression with genetic algorithms in forecasting tourism demand", *Tourism Management*, Vol. 28 No. 1, pp. 215-226.

Chen, M.-H. (2013). "Determinants of the Taiwanese tourist hotel industry cycle", *Tourism Management*, Vol. 38, 15-19.

Chen, M.S., Ying, L.C. and Pan, M.C. (2010), "Forecasting tourist arrivals by using the adaptive network-based fuzzy inference system", *Expert Systems with Applications*, Vol. 37, pp. 1185-1191.

Chu, F.-L. (2009), "Forecasting tourism demand with ARMA-based methods", *Tourism Management*, Vol. 30 No. 5, pp. 740-751.

Claveria, O. and Datzira, J. (2010), "Forecasting tourism demand using consumer expectations", *Tourism Review*, Vol. 65 No. 1, pp. 18-36.

Claveria, O., Monte, E. and Torra, S. (2015a), "A new forecasting approach for the hospitality industry", *International Journal of Contemporary Hospitality Management*, Vol. 27 No. 7, pp. 1520-1538.

Claveria, O., Monte, E. and Torra, S. (2015b), "Tourism demand forecasting with neural network models: different ways of treating information", *International Journal of Tourism Research*, Vol. 17, pp. 492-500.

Claveria, O. and Torra, S. (2014), "Forecasting tourism demand to Catalonia: neural networks vs. time series models", *Economic Modelling*, Vol. 36, pp. 220-228.

Clemen, R.T. (1989), "Combining forecasts: a review and annotated bibliography", *International Journal of Forecasting*, Vol. 5, pp. 559-583.

Corgel, J., Lane, J. and Walls, A. (2013), "How currency exchange rates affect the demand for U.S. hotel rooms", *International Journal of Hospitality Management*, Vol. 35, pp. 78-88.

Cortés-Jiménez, I. and Blake, A. (2011), "Tourism demand modeling by propose of visit and nationality", *Journal of Travel Research*, Vol. 50 No. 4, pp. 408-416.

Croce, V. and Wöber, K.W. (2011), "Judgmental forecasting support systems in tourism", *Tourism Economics*, Vol. 17 No. 4, pp. 709-724.

Croes, R. and Semrad, K.J. (2012), "Discounting works in the hotel industry: a structural approach to understanding why", *Tourism Economics*, Vol. 18 No. 4, pp. 769-779.

Cuhadar, M., Cogurcu, I. and Kukrer, C. (2014), "Modelling and forecasting cruise tourism demand to Izmir by different artificial neural network architectures", *International Journal of Business and Social Research*, Vol. 4 No. 3, pp. 12-28.

Deng, M.F. and Athanasopoulos, G. (2011), "Modelling Australian domestic and international inbound travel: a spatial-temporal approach", *Tourism Management*, Vol. 32 No. 5, pp. 1075-1084.

Diaz, M.A. and Mateu-Sbert, J. (2011), "Forecasting daily air arrivals in Mallorca Island using nearest neighbour methods", *Tourism Economics*, Vol. 17 No. 1, pp. 191-208.

Divino, J.A. and McAleer, M. (2010), "Modelling and forecasting daily international mass tourism to Peru", *Tourism Management*, Vol. 31 No. 6, pp. 846-854.

Ellero, A. and Pellegrini, P. (2014), "Are traditional forecasting models suitable for hotels in Italian cities?", *International Journal of Contemporary Hospitality Management*, Vol. 26 No. 3, pp. 383-400.

Eugenio-Martín, J.L. and Campos-Soria, J.A. (2011), "Income and the substitution pattern between domestic and international tourism demand", *Applied Economics*, Vol. 43 No. 20, pp. 2519-2531.

Falk, M. (2013), “Impact of long-term weather on domestic and foreign winter tourism demand”, *International Journal of Tourism Research*, Vol. 15 No. 1, pp. 1-17.

Falk, M. (2014), “Impact of weather conditions on tourism demand in the peak summer season over the last 50 years”, *Tourism Management Perspectives*, Vol. 9, pp. 24-35.

Garín-Muñoz, T. and Montero-Martín, L.F. (2007), “Tourism in the Balearic Islands: a dynamic model for international demand using panel data”, *Tourism Management*, Vol. 28, pp. 1224-1235.

Ghil, M., Allen, M.R., Dettinger, M.D., Ide, K., Kondrashov, D., Mann, M.E., Robertson, A.W., Saunders, A., Tian, Y., Varadi, F. and Yiou, P. (2002), “Advanced spectral methods for climatic time series”, *Reviews of Geophysics,* Vol. 40 No. 1, pp. 1-41.

Gholipour, H.F., Tajaddini, R. and Al-Mulali, U. (2014), “Does personal freedom influence outbound tourism?”, *Tourism Management*, Vol. 41, pp. 19-25.

Goh, C. (2012), “Exploring impact of climate on tourism”, *Annals of Tourism Research*, Vol. 39 No. 4, pp. 1859-1883.

Goh, C. and Law, R. (2011), “The methodological progress of tourism demand forecasting: a review of related literature”, *Journal of Travel and Tourism Marketing*, Vol. 28 No. 3, pp. 296-317.

Goh, C., Law, R. and Mok, H.M.K. (2008), “Analyzing and forecasting tourism: a rough sets approach”, *Journal of Travel Research*, Vol. 46 No. 3, pp. 327-338.

Gounopoulos, D., Petmezas, D. and Santamaria, D. (2012), “Forecasting tourist arrivals in Greece and the impact of macroeconomic shocks from the countries of tourists’ origin”, *Annals of Tourism Research*, Vol. 39 No. 2, pp. 641-666.

Guizzardi, A. and Stacchini, A. (2015), “Real-time forecasting regional tourism with business sentiment surveys”, *Tourism Management*, Vol. 47, pp. 213-223.

Hadavandi, E., Ghanbari, A., Shahanaghi, K. and Abbasian-Naghneh, S. (2011), “Tourist arrival forecasting by evolutionary fuzzy systems”, *Tourism Management*, Vol. 32 No. 5, pp. 1196-1203.

Hassani, H. Webster, A., Silva, E.S. and Heravi, S. (2015), “Forecasting U.S. tourist arrivals using optimal singular spectrum analysis”, *Tourism Management*, Vol. 46, pp. 322-335.

Hong, W.-C., Dong, Y., Chen, L.-Y. and Wei, S.-Y. (2011), “SVR with hybrid chaotic genetic algorithms for tourism demand forecasting”, *Applied Soft Computing*, Vol. 11, pp. 1881-1890.

Kim, J.H., Song, H. and Wong, K.K.F. (2010), “Bias-corrected bootstrap prediction intervals for autoregressive model: new alternatives with applications to tourism forecasting”, *Journal of Forecasting*, Vol. 29 No. 7, pp. 655-672.

Kim, J.H., Wong, K., Athanasopoulos, G. and Liu, S. (2011), “Beyond point forecasting: evaluation of alternative prediction intervals for tourist arrivals”, *International Journal of Forecasting*, Vol. 27 No. 3, pp. 887-901.

Kim, N. and Schwartz, Z. (2013), “The accuracy of tourism forecasting and data characteristics: a meta-analytical approach”, *Journal of Hospitality Marketing and Management*, Vol. 22, pp. 349-374.

Koupriouchina, L, van der Rest, J. and Schwartz, Z. (2014), “On revenue management and the use of occupancy forecasting error measures”, *International Journal of Hospitality Management*, Vol. 41, pp. 104-114.

Lee, C.-K., Song, H.-J. and Bendle, L.J. (2010), “The impact of visa-free entry on outbound tourism: a case study of South Korean travellers visiting Japan”, *Tourism Geographies*, Vol. 12 No. 2, pp. 302-323.

Lee, C.-K., Song, H.-J. and Mjelde, J.W. (2008), "The forecasting of International Expo tourism using quantitative and qualitative techniques", *Tourism Management*, Vol. 29 No. 6, pp. 1084-1098.

Li, G., Song, H. and Witt, S. F. (2005). "Recent developments in econometric modeling and forecasting", *Journal of Travel Research*, Vol. 44 No. 1, pp. 82-99.

Li, G., Song, H., Cao Z. and Wu, D.C. (2013), "How competitive is Hong Kong against its competitors? An econometric study", *Tourism Management*, Vol. 36 No. 1, 247-256.

Lim, C., Chang, C. and McAleer, M. (2009), "Forecasting h(m)otel guest nights in New Zealand", *International Journal of Hotel Management*, Vol. 28 No. 2, pp. 228-235.

Lin, C.-J., Chen, H.-F. and Lee, T.-S. (2011), "Forecasting tourism demand using time series, artificial neural networks and multivariate adaptive regression splines: evidence from Taiwan", *International Journal of Business Administration*, Vol. 2 No. 2, pp. 14-24.

Lin, V.S. (2013), "Improving forecasting accuracy by combining statistical and judgmental forecasts in tourism", *Journal of China Tourism Research*, Vol. 9 No. 3, pp. 325-352.

Lin, V.S., Goodwin, P. and Song, H. (2014), "Accuracy and bias of experts' adjusted forecasts", *Annals of Tourism Research*, Vol. 48, pp. 156-174.

Marrocu, E. and Paci, R. (2013), "Different tourists to different destinations: evidence from spatial interaction models", *Tourism Management*, Vol. 39, pp. 71-83.

Medeiros, M.C., McAleer, M., Slottje, D., Ramos, V. and Rey-Maquieira, J. (2008), "An alternative approach to estimating demand: neural network regression with conditional volatility for high frequency air passenger arrivals", *Journal of Econometrics*, Vol. 147 No. 2, pp. 372-383.

Mieczkowski, Z. (1985), "The tourism climatic index: a method of evaluating world climates for tourism", *Canadian Geographer*, Vol. 29, pp. 220-233.

Mohammed, I., Guillet, B.D. and Law, R. (2015), "The contributions of economics to hospitality literature: a content analysis of hospitality and tourism journals", *International Journal of Hospitality Management*, Vol. 44, pp. 99-110.

Onafowora, O.A. and Owoye, O. (2012), "Modelling international tourism demand for the Caribbean", *Tourism Economics*, Vol. 18 No. 1, pp. 159-180.

Page, S.J., Song, H. and Wu, D.C. (2012). "Assessing the impacts of the economic crisis and swine flu on inbound tourism demand in the UK", *Journal of Travel Research*. Vol. 51 No. 2, pp. 142-153.

Pai, P.F., Hung, K.C. and Lin, K.P. (2014), "Tourism demand forecasting using novel hybrid system", *Expert Systems with Applications*, Vol. 41 No. 8, pp. 3691-3702.

Pan, B., Wu, D.C. and Song, H. (2012), "Forecasting hotel room demand using search engine data", *Journal of Hospitality and Tourism Technology*, Vol. 3 No. 3, pp. 196-210.

Peng, B., Song, H. and Crouch, G.I. (2014), "A meta-analysis of international tourism demand forecasting and implications for practice", *Tourism Management*, Vol, 45, pp. 181-193.

Rodriguez, X.A., Martinez-Roget, F. and Pawlowska, E. (2012), "Academic tourism demand in Galicia, Spain", *Tourism Management*, Vol. 33 No. 6, pp. 1583-1590.

Rosselló-Nadal, J., Riera-Font, A. and Cárdenas, V. (2011), "The impact of weather variability on British outbound flows", *Climatic Change*, Vol. 105 No. 1-2, pp. 281-292.

Serra, J., Correia, A. and Rodrigues, P.M.M. (2014), "A comparative analysis of tourism destination demand in Portugal", *Journal of Destination Marketing & Management*, Vol. 2, pp. 221-227.

Shahrabi, J., Hadavandi, E. and Asadi, S. (2013), "Developing a hybrid intelligent model for forecasting problems: case study of tourism and hospitality demand time series", *Knowledge-Based Systems*, Vol. 43, pp. 112-122.

Shen, S., Li, G. and Song, H. (2008), "An assessment of combining tourism demand forecasts over different time horizons", *Journal of Travel Research*, Vol. 47 No. 2, pp. 197-207.

Shen, S., Li, G. and Song, H. (2011), "Combination forecasts of international tourism demand", *Annals of Tourism Research*, Vol. 38 No. 1, pp. 72-89.

Singh, A., Dev, C.S. and Mandelbaum, R. (2014), "A flow-through analysis of the US lodging industry during the great recession", *International Journal of Contemporary Hospitality Management*, Vol. 26 No. 2, pp. 205-224.

Smeral, E. (2010), "Impacts of the world recession and economic crisis on tourism: forecasts and potential risks", *Journal of Travel Research*, Vol. 49 No. 1, pp. 31-38.

Song, H. and Li, G. (2008), "Tourism demand modelling and forecasting – a review of recent research", *Tourism Management*, Vol. 29 No. 2, pp. 203-220.

Song, H. and Lin, S. (2010), "Impacts of the financial and economic crisis on tourism in Asia", *Journal of Travel Research*, Vol. 49 No. 1, pp. 16-30.

Song, H. and Witt, S.F. (2000), *Tourism Demand Modelling and Forecasting: Modern Econometric Approaches*, Pergamon: Oxford.

Song, H., Dwyer, L., Li, G. and Cao, Z. (2012a), "Tourism economics research: a review and assessment", *Annals of Tourism Research*, Vol. 39 No. 3, pp. 1653-1682.

Song, H., Gao, B.Z. and Lin, V.S. (2013), "Combining statistical and judgmental forecasts via a web-based tourism and hospitality demand forecasting system", *International Journal of Forecasting*, Vol. 29 No. 2, pp. 295-310.

Song, H., Gartner, W.C. and Tasci, A.D.A. (2012b), "Visa restrictions and their adverse economic and marketing implications – evidence from China", *Tourism Management*, Vol. 33 No. 2, pp. 397-412.

Song, H., Kim, J.H. and Yang, S. (2010), "Confidence intervals for tourism and hospitality demand elasticity", *Annals of Tourism Research*, Vol. 37 No. 2, pp. 377-396.

Song, H., Li, G., Witt, S.F. and Athanasopoulos, G. (2011a), "Forecasting tourist arrivals using time-varying parameter structural time series models", *International Journal of Forecasting*, Vol. 27 No. 3, pp. 855-869.

Song, H., Lin, S., Witt, S.F. and Zhang, X.Y. (2011b), "Impact of financial/economic crisis on demand for hotel rooms in Hong Kong", *Tourism Management*, Vol. 32 No. 1, pp. 172-186.

Song, H., Witt, S.F., Wong, K.F. and Wu, D.C. (2009), "An empirical study of forecast combination in tourism", *Journal of Hospitality & Tourism Research*, Vol. 33 No. 1, pp. 3-29.

Toma, M., McGrath, R. and Payne, J.E. (2009), "Hotel tax receipts and the 'Midnight in the Garden of Good and Evil': a time series intervention seasonal ARIMA model with time-varying variance", *Applied Economics Letters*, Vol. 16 No. 7, pp. 653-656.

Torraleja, F.-A.G., Vazquez, A.M. and Franco, M.J.B. (2009), "Flows into tourist areas: an econometric approach", *International Journal of Tourism Research*, Vol. 11 No. 1, pp. 1-15.

Tsaur, R.-C. and Kuo, T.-C. (2011), "The adaptive fuzzy time series model with an application to Taiwan's tourism demand", *Expert Systems with Application*, Vol. 38, pp. 9164-9171.

Tsui, W.H.K., Balli, H.O., Gilbey, A. and Gow, H. (2014), "Forecasting of Hong Kong airport's passenger throughput", *Tourism Management*, Vol. 42, pp. 62-76.

Valadkhani, A. and O'Mahony, B.O. (2015), "Identifying structural changes and regime switching in growing and declining inbound tourism markets in Australia", *Current Issues in Tourism* (DOI: 10.1080/13683500.2015.1072504z).

Wang, Y.S. (2014), "Effects of budgetary constraints on international tourism expenditures", *Tourism Management*, Vol. 41, pp. 9-18.

Wong, K.K.F., Song, H., Witt, S.F. and Wu, D.C. (2007), "Tourism forecasting: to combine or not to combine?", *Tourism Management*, Vol. 28 No. 4, pp. 1068-1078.

Wu, D.C., Li, G. and Song, H. (2011), "Analyzing tourist consumption: a dynamic system-of-equations approach", *Journal of Travel Research*, Vol. 51 No. 1, pp. 46-56.

Wu, D.C., Li, G. and Song, H. (2012a), "Economic analysis of tourism consumption dynamics: a time-varying parameter demand system approach", *Annals of Tourism Research*, Vol. 39 No. 2, pp. 667-685.

Wu, E.H.C., Law, R. and Jiang, B. (2010), "Data mining for hotel occupancy rate: an independent component analysis approach", *Journal of Travel & Tourism Marketing*, Vol. 27, pp. 426-438.

Wu, Q., Law, R. and Xu, X. (2012b), "A sparse Gaussian process regression model for tourism demand forecasting in Hong Kong", *Expert Systems with Applications*, Vol. 39, pp. 4769-4774.

Xu, X., Law, R. and Wu, T. (2009), "Support vector machines with manifold learning and probabilistic space projection for tourist expenditure analysis", *International Journal of Computational Intelligence Systems*, Vol. 46, pp. 386-397.

Yang, X., Pan, B., Evans, J.A. and Lv, B. (2015), "Forecasting Chinese tourist volume with search engine data", *Tourism Management*, Vol. 46, pp. 386-397.

Yang, Y., Liu, Z.-H. and Qi, Q. (2014a), "Domestic tourism demand of urban and rural residents in China: does relative income matter", *Tourism Management*, Vol. 40, pp. 193-202.

Yang, Y., Pan, B. and Song, H. (2014b), "Predicting hotel demand using destination marketing organization's web traffic data", *Journal of Travel Research*, Vol. 53 No. 4, pp. 433-447.

Zadrozny, P.A. (1988), "Gaussian-likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies", *Econometric Theory*, Vol. 4 No. 1, pp. 108-124.

Zheng, T. (2014), "What caused the decrease in RevPAR during the recession?", *International Journal of Contemporary Hospitality Management*, Vol. 26 No. 8, pp. 1225-1242.

Zheng, T., Farrish, J., Lee, M.-L. and Yu, H. (2013), "Is the gaming industry still recession-proof?", *International Journal of Contemporary Hospitality Management*, Vol. 25 No. 7, pp. 1135-1152.