Tourism Management 45 (2014) 181–193

Contents lists available at ScienceDirect

Tourism Management

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

A meta-analysis of international tourism demand forecasting and implications for practice

Bo Peng [a], Haiyan Song [b,*], Geoffrey I. Crouch [c]

[a] Cass Business School, City University London, London EC1Y 8TZ, UK
[b] School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Hong Kong, China
[c] Department of Marketing, Tourism and Hospitality, La Trobe Business School, La Trobe University, Melbourne, Australia

HIGHLIGHTS

- The relationship between forecast accuracy, data characteristics and study features is studied.
- The results show that key study features and data characteristics influence forecasting accuracy.
- The findings provide suggestions for the choice of appropriate forecasting methods in practice.

ARTICLE INFO

Article history:
Received 6 October 2012
Accepted 11 April 2014
Available online 7 May 2014

Keywords:
Forecasting accuracy
International tourism demand
Meta-analysis

ABSTRACT

Numerous studies on tourism forecasting have now been published over the past five decades. However, no consensus has been reached in terms of which types of forecasting models tend to be more accurate and in which circumstances. This study uses meta-analysis to examine the relationships between the accuracy of different forecasting models, and the data characteristics and study features. By reviewing 65 studies published during the period 1980–2011, the meta-regression analysis shows that the origins of tourists, destination, time period, modeling method, data frequency, number of variables and their measures and sample size all significantly influence the accuracy of forecasting models. This study is the first attempt to pair forecasting models with the data characteristics and the tourism forecasting context. The results provide suggestions for the choice of appropriate forecasting methods in different forecasting settings.

1. Introduction

Since international tourism has become increasingly important to worldwide economic development, both the public and private sectors have channeled a significant amount of resources and investment into the industry. As both governments and businesses need accurate forecasting to develop efficient public policy and make good business investment decisions, considerable efforts have been made to improve the accuracy of tourism demand forecasting. Before the 1990s, traditional regression approaches dominated the tourism forecasting and modeling literature. After incorporating up-to-date developments in econometric methodologies in recent years, the reputation of econometric forecasting models for improved accuracy has grown (Song & Li, 2008). Other quantitative methods, such as gravity models, artificial neural networks (ANN), and univariate time series models, have also played important roles in tourism demand forecasting. However, conflicting conclusions still exist in terms of which models generate the most accurate forecasts under different conditions. Each method has its own advantages in dealing with a particular problem, but none has been shown to be universally superior.

In addition to research on accuracy improvement, an emerging area of work is the synthesis of tourism demand forecasting techniques. In their literature review of empirical research, Witt and Witt (1995) found that it is not possible to build a single econometric model that is appropriate for all origin-destination pairs. They also showed that the performance of forecasting models varies according to the time interval of the data, the destination–origin pair, and the forecasting horizon. However, until now, little effort has been made to identify optimal forecasting models according to the data characteristics and study features or forecasting contexts, by learning comprehensively from the lessons of this now

* Corresponding author. Tel.: +852 3400 2286; fax: +852 2362 9362.
E-mail addresses: gracepeng213@163.com (B. Peng), haiyan.song@polyu.edu.hk (H. Song), Geoffrey.Crouch@latrobe.edu.au (G.I. Crouch).

http://dx.doi.org/10.1016/j.tourman.2014.04.005
0261-5177/

large body of work. Meta-analysis is a statistically rigorous technique for synthesizing the empirical findings from previous studies. Its aim is to explain what data or study features best account for differences in study findings. Through a comprehensive review and integration of 65 articles and 3198 estimates of tourism demand forecasting accuracy measures over the period 1980–2011, this study sets out to identify such explanations of international tourism demand forecasting accuracy. Where explanations are found, superior forecasting models can be identified, based on the characteristics of the study and the data used in the model, to assist practitioners in making better choices for forecasting methodology, leading to more effective policy and business decisions.

2. Review of quantitative tourism forecasting methods

Time series models, econometric approaches, and artificial intelligence (AI) models are three main categories of quantitative forecasting methods. Time series methods extrapolate from previous data in the series to predict future trends, which require no more than one data series. According to the complexities of the models, time series forecasting methods can be further subdivided into basic and advanced subcategories. The former includes the Naive, Simple Moving Average (SMA), and Single Exponential Smoothing (SES) models. The advanced approaches include the double exponential smoothing (DES), exponential smoothing adjusted by trend, autoregressive moving average (ARMA), and basic structural time series (BSM) models.

Although time series approaches are useful tools in tourism demand forecasting, their major limitation is that their construction is not based on any economic theory that underlines tourists' decision-making processes. Therefore, not only can they not be used to analyse tourists' behavior, they are also incapable of assisting policymakers in evaluating the effectiveness of their strategies and policies. From this perspective, then, econometric models are superior (Song, Witt, & Li, 2009). Rather than relying on extrapolation, econometric approaches seek to find dependent relationships between tourism demand and a set of explanatory variables. Tourism forecasts can then be produced as a function of the values taken by these explanatory variables in the future. This approach permits forecasting for different scenarios (e.g., different exchange rate outcomes).

2.1. Basic time series methods[1]

The *Naive 1 (or no change)* model is the simplest method and has often been shown to generate more accurate one-year-ahead forecasts than other more sophisticated models (Martin & Witt, 1989; Witt, Witt, & Wilson, 1994). However, the performance of the Naive 1 model declines when it has to deal with sudden structural change and longer-term forecasting (Chan, Hui, & Yuen, 1999; Witt et al., 1994). The *Naive 2 (or constant change)* model is another widely used but simple model employed when there is a constant trend present in the data. Chan et al. (1999) use the Gulf War as an example of a sudden shock and show that the Naive 2 model performs better than the autoregressive integrated moving average (ARIMA), exponential smoothing (ES), and quadratic trend curve models when dealing with unstable data.

The *SMA* model allows the past values of a variable to determine the forecast values with equal weights assigned to the former. If a time series shows wide variations around a trend, including more lagged observations, the SMA approach will help the model to pick up the trend. However, its main limitation is that it gives equal weight to all the lagged observations (Hu, Chen, & McChain, 2004), which may not be realistic, as more recent lagged values tend to have a much bigger impact on the current values of a time series. Therefore, the SMA method normally generates more accurate forecasts where the time series is less volatile (Makridakis, Wheelwright, & Hyndman, 1998). Systematic errors may occur when the SMA model deals with a time series that has a linear trend. To overcome this problem, researchers can use the double moving average method to further smooth the series (Hu et al., 2004; Lim & McAleer, 2008).

The *SES* model is used to forecast a time series when there is no trend or seasonal pattern. According to Chen, Bloomfield, and Cubbage (2008), SES is more suitable for a time series with seasonality removed. Witt, Newbould, and Watkins (1992) show that the SES model generates forecasts with relatively lower error magnitudes than the Naive 1 model for domestic tourism demand, which is less volatile than international tourism demand.

2.2. Advanced time-series methods

Brown's DES model (Brown, 1963) was developed to deal with time series with linear trends. Geurts and Ibrahim (1975) were the first to apply Brown's DES model to forecast tourist arrivals in Hawaii and suggest that it is cheaper and easier to use than the Box–Jenkins approach for forecasting domestic tourism demand. Sheldon (2008) shows that Brown's DES model and the Naive 1 model also perform well in forecasting international tourism expenditure. The disadvantage of DES is that it does not track nonlinear trends well and often fails to pick up structural breaks in the time series (Frechtling, 1996).

Holt's DES model (Holt, 1957) is more flexible in selecting the smoothing constants (Makridakis et al., 1998). However, according to Chen et al. (2008), Brown's DES models outperform Holt's model based on mean absolute percentage error (MAPE) in forecasting tourist arrivals to US national parks. Holt–Winter's model (the triple ES method) adds seasonal variation to Holt's model, which captures both the seasonal pattern and trend of the time series and usually outperforms other ES methods (Lim & McAleer, 2001). Grubb and Mason (2001) prove that adding a damped trend to Holt–Winter's method greatly improves long-run forecasts compared with the Box–Jenkins and BSM in the case of UK air passengers.

The *Box–Jenkins* model, with the *ARMA* process as its basic form is the most frequently used time series approach in tourism demand forecasting. Researchers' evaluations of the Box–Jenkins models are mixed. Makridakis and Hibon (1979) argue that they produce little improvement in forecasting accuracy, and Kim, Wong, Athanasopoulos, and Liu (2011) conclude that the Seasonal ARIMA (SARIMA) model tends to underestimate the future uncertainty in interval forecasting. Other studies suggest that the ARIMA and SARIMA approaches are preferred in tourism demand forecasting when the time series does not exhibit structural breaks (see for example, Chu, 2008; Goh & Law, 2002; Gustavsson & Nordström, 2001). Preez and Witt (2003) show that the ARIMA approach performs best in terms of forecasting accuracy and goodness of fit.

The BSM model is constructed by decomposing a time series into its trend, seasonal, cyclical, and irregular components. Greenidge (2001) successfully applied BSM to forecasting tourist arrivals to Barbados, and showed that it offered valuable insights

[1] Our use of two categories for classification of the time series models into *basic* and *advanced time series* models is based on the three categories of Frechtling (1996). For the purpose of parsimony, however, we collapsed these to two categories in order to maintain a tractable number of explanatory variables in the meta-analysis.