International Journal of Forecasting 29 (2013) 295–310

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

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

Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system

Haiyan Song, Bastian Z. Gao, Vera S. Lin *

School of Hotel and Tourism Management, The Hong Kong Polytechnic University, 17 Science Museum Road, Kowloon, Hong Kong

A R T I C L E I N F O

Keywords:
Tourism demand forecasting
Econometric model
Judgmental adjustment
Domain knowledge
Hong Kong

A B S T R A C T

This paper introduces a web-based tourism demand forecasting system (TDFS) that is designed to forecast the demand for Hong Kong tourism, as measured by tourist arrivals, total and sectoral tourist expenditures, and the demand for hotel rooms. The TDFS process comprises three stages – preliminary data analysis, the generation of quantitative forecasts and judgmental adjustments – which correspond to the three key system components: the data module, the quantitative forecasting module and the judgmental forecasting module, respectively. These stages (modules) interact with one another. This paper focuses on a recent case study that illustrates the functional ability of the TDFS as a support system, providing accurate forecasts of the demand for Hong Kong tourism. Specifically, the quantitative forecasts are generated by the autoregressive distributed lag model, then adjusted by a panel of experts comprising postgraduate students and academic staff. The results show that this combination of quantitative and judgmental forecasts improves the overall forecasting accuracy.

1. Introduction

Because tourism products and services are perishable, accurate forecasts of tourism demand are essential for the effective formulation and implementation of tourism strategies or plans by both public and private sector organizations that are involved in tourism. The special characteristics of tourism demand present a number of special difficulties to tourism forecasters that do not afflict forecasters in other industries. Frechtling (2001) summarized five major challenges in tourism demand forecasting, namely the lack of historical data; the high volatility of tourism demand (e.g., high seasonality); the high sensitivity of demand to external shocks such as terrorism, earthquakes, floods and diseases; the complexity of tourist behavior; and the wide choice of forecast variables (e.g., tourist arrivals, tourist expenditure, hotel room nights). Successful tourism managers from both public and private organizations need to find ways of reducing the risk of future failures in tourism demand forecasting. As no consensus has been achieved as to a sound theoretical foundation for tourism demand, a knowledge of the patterns and future trends of tourism demand is of great importance on both theoretical and practical grounds.

There is broad consensus that no one forecasting model outperforms all others on all occasions, and that environment-specific conditions determine which method suits a specific forecasting task best. Combining forecasts based on different methods or data has emerged as one of the most important ways of improving forecasting performances. The combination of quantitative and judgmental forecasts adds a promising dimension to forecasting, and has actually been a key research area over the past three decades. In light of the complementary nature of quantitative and judgmental forecasting methods, the forecasts generated by integrating statistical and judgmental findings are likely to be more accurate than those generated by either of the two methods alone (Blattberg & Hoch, 1990).

* Corresponding author. Tel.: +852 3400 2339; fax: +852 2362 9362.
E-mail address: vera.shanshan.lin@connect.polyu.hk (V.S. Lin).

0169-2070/$ – see front matter
doi:10.1016/j.ijforecast.2011.12.003

Studies have identified the following benefits of combining quantitative methods and expert judgments. First, experts are able to provide contextual information, which can be used to explain anomalies in the data and to identify future events that may cause the data to deviate from historical patterns (Armstrong & Collopy, 1998). Second, forecasting failure is caused mainly by the inappropriate use of statistical models that do not fit the data well. In such situations, judgment is important for identifying the appropriate statistical method(s) needed for forecasting (Bunn & Wright, 1991). Third, evidence suggests that judgmental adjustments of statistical forecasts improve the forecast accuracy if the forecasters have sufficient domain knowledge that is not available to the statistical forecasting method(s) (Goodwin, Fildes, Lawrence, & Nikolopoulos, 2007). Judgmental forecasting is expected to address the problem of forecasting uncertainty associated with the statistical methods. Finally, this forecasting combination enables forecasts to be updated in a timely fashion by incorporating the most recent and most relevant information into the quantitative models (Sanders & Ritzman, 2004), which might be important in ensuring the effectiveness of future planning.

Evidence also suggests, however, that the use of forecasters' judgments can be inefficient, biased and inconsistent. Empirical studies have shown that forecasters frequently adjust their quantitative forecasts in real-world forecasting contexts (Klassen & Flores, 2001; McCarthy, Davis, Golicic, & Mentzer, 2006; Sanders & Manrodt, 1994). Goodwin et al. (2007) revealed that statistical forecasts generated by inappropriate methods tend to be adjusted on a large scale, generally leading to less accurate forecasts than those produced by suitable forecasting models. Fildes, Goodwin, Lawrence, and Nikolopoulos (2009) suggested that when a suitable statistical forecasting method is selected, relatively large and negative adjustments are likely to produce a greater accuracy, whereas smaller and positive adjustments often decrease the accuracy.

Despite a considerable amount of research into and progress in general forecasting methods, research into tourism forecasting that combines quantitative forecasting with expert judgmental forecasting remains limited. The study described herein is a follow-up to the study carried out by Song, Witt, and Zhang (2008), and its aim was to develop their proposed web-based tourism demand forecasting system (TDFS) further by combining the statistical forecasts generated by an autoregressive distributed lag model (ADLM) with the expert judgmental forecasts of an experimental panel comprising postgraduate students and academic staff.

The remainder of this paper is structured as follows: Section 2 reviews the literature on forecasting system architectures, quantitative forecasting methods, judgmental adjustment and combined forecasting approaches. Section 3 briefly introduces the TDFS and its three main functional modules. Section 4 presents an experimental case study in which the forecasting ability of the TDFS is evaluated using Hong Kong tourism demand data. Section 5 summarizes the study's findings and provides some recommendations for future research.

2. Literature review

Armstrong (2001, p. 784) defines a forecasting support system (FSS) as "a set of procedures (typically computer based) that supports forecasting. It allows the analyst to easily access, organize and analyze a variety of information. It might also enable the analyst to incorporate judgment and monitor forecast accuracy". The key features of such an FSS include (i) a database that stores historical time series data, (ii) a set of quantitative forecasting methods, and (iii) functions that allow managerial judgments to be implemented (Fildes, Goodwin, & Lawrence, 2006). The system forecasting model, which includes a set of quantitative methods, is used to produce the initial system forecasts, and the human interactive approach is then employed to incorporate any contextual information (e.g., significant events) that is not reflected in the historical data. Studies of tourism demand forecasting by FSSs have focused mainly on quantitative forecasts based on statistical models. Only a handful of studies, such as those of Song et al. (2008) and Croce and Wober (2011), have combined quantitative forecasting methods and human judgments in tourism.

2.1. Quantitative forecasting methods

Quantitative forecasting methods, including univariate time series methods and causal econometric approaches (Armstrong, 2006; Li, Song, & Witt, 2005), are used widely in tourism demand forecasting. Due to their simplicity, low cost and user-friendliness, univariate time series models, such as the naïve, moving average, exponential smoothing and Box–Jenkins models, are frequently used as the main quantitative forecasting models in FSSs. Clements and Hendry (1998) suggest that econometric models are capable of analyzing the causal relationships between an outcome variable and influential factors, thereby generating forecasts and verifying existing economic theories. Econometric models are used widely in empirical analyses of economic relationships and in forecasting. The current tourism demand forecasting literature indicates that modern econometric models, such as the ADLM, the error correction model (ECM), the vector autoregressive (VAR) model, the time varying parameter (TVP) model, and the almost ideal demand system (AIDS), are well represented in the empirical research. However, none of these models outperforms all of the others on all occasions (Song & Li, 2008).

2.2. User intervention in the forecasting process

Empirical studies show that the forecast accuracy is likely to increase with user interventions during the forecasting process. The size of the adjustment, direction of the adjustment (i.e., positive or negative), and characteristics of the data series to be forecast also affect the forecasting performance (either increasing or decreasing the forecast accuracy) (Bonaccio & Dalal, 2006; Croce & Wober, 2011; Harvey, 1995; Kottemann, Davis, & Remus, 1994; Lim & O'Connor, 1995; Sanders & Ritzman, 1992; Syntetos, Nikolopoulos, Boylan, Fildes, & Goodwin, 2009; Vokurka, Flores, & Pearce, 1996).

- *Finding* 1: *The user's judgment in identifying characteristics of the series to be forecast and the appropriate data processing approach is beneficial for reducing forecast errors.*

In order to obtain reliable forecasts, the user's judgment is normally incorporated at the preliminary data analysis stage, which includes data feature detection and the purging of noise from the raw data, before the statistical models are estimated (Collopy & Armstrong, 1992; Vokurka et al., 1996).

- *Finding* 2: *Judgmental adjustments improve the forecast accuracy when forecasters have important information about the outcome variable being forecast that is not available to the statistical model.*

Empirical evidence also suggests that the forecast accuracy is increased when forecasters take the effects of special events (e.g., a forthcoming sales promotion) into account in adjusting the statistical forecasts. Such contextual information, however, is not usually available to forecasters (Donihue, 1993; Goodwin & Fildes, 1999; Lim & O'Connor, 1996; McNees, 1990; Turner, 1990). If important contextual information is lacking, then forecasters may identify false patterns in the noise associated with the past data series, and adjustments made accordingly are likely to decrease the accuracy (O'Connor, Remus, & Griggs, 1993).

2.3. Computer-based innovation

A tourism demand forecasting system (TDFS) can be defined, at the functional level, as "a forecasting support system that is capable of providing quantitative tourism demand forecasts and allowing users to perform scenario analyses or make their own 'what-if' forecasts" (Song et al., 2008, p. 446). Specifically, a web-based TDFS is "a computerized information system that delivers tourism demand forecasts and provides decision support to policymakers and business strategists via a Web browser" (Song et al., 2008, p. 446). A web-based system is easier to use and update than a stand-alone forecasting system, as well as being cheaper and enabling better collaboration among stakeholders (Song et al., 2008). Croce and Wober (2011) point out other benefits of a web-based forecasting system: using the Internet removes geographical barriers to accessing the database, reduces the costs of information dissemination, enables collaboration among people in different geographic locations and different knowledge domains, and prevents the influence of leading personalities. In addition, users can specify their forecast for each origin–destination pair and replicate the series-specific behaviour in their estimates via a web-based TDFS, which is particularly relevant for tourism analyses (Croce & Wober, 2011).

Delen, Sharda, and Kumar (2007) designed a Movie Forecast Guru (MFG) which supports a process that is repeated at regular intervals and allows forecasts and actual outcomes to be compared. Petropoulos, Patelis, Metaxiotis, Nikolopoulos, and Assimakopoulos (2003) proposed a web-based decision support system specifically for tourism demand forecasting. The proposed four-tier architecture and its computer-based techniques provide the user with convenient tools for tourism data analysis, and the automatic time series forecasting process makes the forecasting easier. However, a comparison of the forecasts generated by the system with those generated by four alternative methods showed that the use of the system did not significantly improve the forecast accuracy, as measured by the mean absolute percentage error (MAPE). Indeed, large forecast errors are still associated with the existing FSSs, for the following reasons:

(1) These systems consist only of pure time series methods, which ignore changes in the outcome variables resulting from explanatory variables.
(2) Most of the systems require their users to have a strong mathematics/statistics background. However, tourism practitioners often lack such a background, which tends to result in the selection of inappropriate forecasting models.
(3) None of these systems provides suggestions or guidelines for users during the forecasting process, and no evaluation of forecasting performances is provided.

These drawbacks can affect the effective use of an FSS and reduce the forecast accuracy. Fildes et al. (2006) stated that improvements in FSSs should focus on the forecaster's ability to discern when to use judgmental interventions and how to apply them. Therefore, future research on the development of FSSs for tourism should focus on a design of the system architecture which will allow effective user interventions and performance evaluation during the forecasting process through appropriate computer-based algorithms. One goal is to automate the quantitative forecasting procedure without any loss of forecast accuracy (Hannan & Rissanen, 1982; Liu, 1989; Mélard & Pasteels, 2000). Another goal is the effective incorporation in the forecasting process of either the forecaster's technical knowledge about the statistical models used in the system or his or her professional knowledge about the industry (Collopy & Armstrong, 1992; Petropoulos et al., 2003; Vokurka et al., 1996).

3. TDFS design

3.1. Overview

The features added to the TDFS originally developed by Song et al. (2008) include the following:

- User-friendliness: the new system architecture can help users to generate tourism forecasts in a more efficient and effective way.
- Modularity: the components of the system are designed as stand-alone modules to reduce the cost of system maintenance.
- Flexibility: the system modules, particularly the application modules, can now be updated and redesigned easily when new technologies and algorithms become available.
- Enhanced website administration system: system administrators and authorized users can log on to the administration module via a web-based interface and perform routine administrative tasks, such as user account management, file sharing and database management.

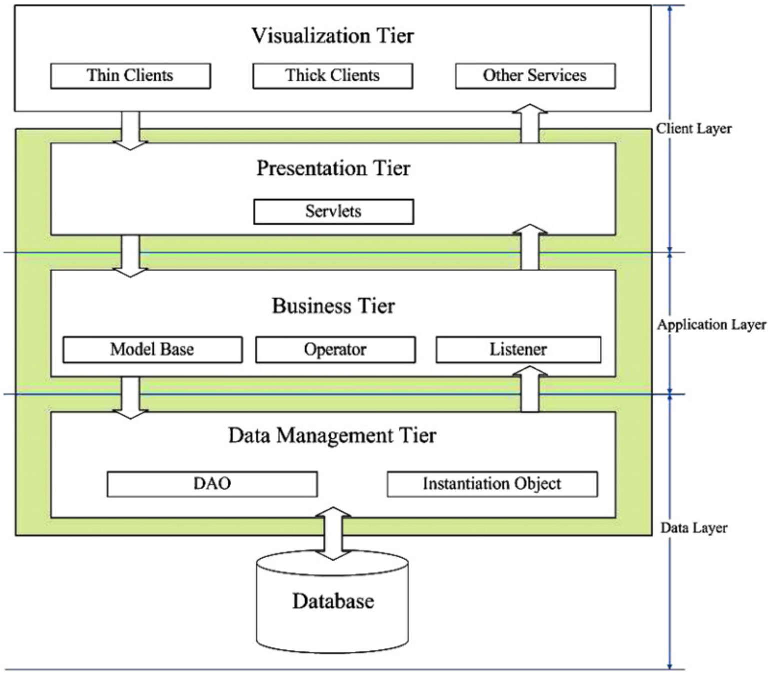


Fig. 1. TDFS architecture.

- Java Server Pages (JSP) and R-based applications: the JSP web language is used to develop the system, as it provides stable interfaces with external software and languages, such as the statistical language and the R environment.
- Implementation of open source R code: R provides a wide variety of statistical and graphical options, including linear and nonlinear modelling, classic statistical tests, time series analysis, classification, clustering, and many other statistical applications. With the TDFS, the quantitative forecasts are generated in the R environment.
- Improvements in judgmental inputs: the system includes a dynamic online Delphi survey module, which allows the integration of statistical and judgmental forecasts.

Fig. 1 illustrates the four-tier Client/Server (C/S) architecture of the TDFS web platform. The first two tiers are traditional components of the C/S architecture, by which users interact with the system for any specific application. In the third (business) tier, the core functions, including all operational logistics, are hosted on an Apache Tomcat Web Server. In particular, an interface known as REngine, an abstract base class for all implementations of R engines, is deployed in this tier to allow communication between the web platform and the R environment. To begin with, a dataset (in Excel format) is supplied by the user. Once a request to estimate the model is given, the system connects itself to the REngine client and runs the Model Estimation module. The estimation results, including the diagnostic tests and tourism demand elasticities (e.g., income and price elasticities), are then presented on the web pages. The results are stored in the database simultaneously (see Fig. 1).

The fourth tier is the database tier (see Fig. 2). Two different databases are available in this tier. The first is the MySQL database, which is used to store three types of data: (i) historical time series of all tourism demand measures and their influencing factors, such as GDP, own price and substitute price; (ii) estimation results (e.g., diagnostic statistics and elasticities); and (iii) forecasts. The second database contains the R program codes that are used to run the econometric model (or ADLM) embedded in the system.

Although most tourism managers/forecasters have rich industry experience, some may have very little knowledge about quantitative forecasting methods, and particularly about advanced econometric forecasting methods. To make the use of the TDFS easier, the system is designed to automate the forecasting process, and thus requires little modelling knowledge. The system also makes full use of the forecasters' domain knowledge and integrates it with the econometric forecasts in order to achieve a greater forecast accuracy.

The TDFS forecasting procedure involves three stages (see Fig. 3). The first stage is the pre-modelling data analysis, which is performed outside the TDFS via a number of statistical analysis/software packages such as SPSS, EViews, Matlab and Excel. The tasks in this stage are to examine and identify the properties of the data by testing for unit roots and co-integration, for example, and to introduce dummy variables that take seasonality and the impacts of special events into account. Following the preliminary data analysis, the processed dataset can then be imported into the TDFS. In the second stage, once the data have been input into the system and are ready for the forecasting tasks to be performed, the system runs the Model Estimation module automatically after receiving the

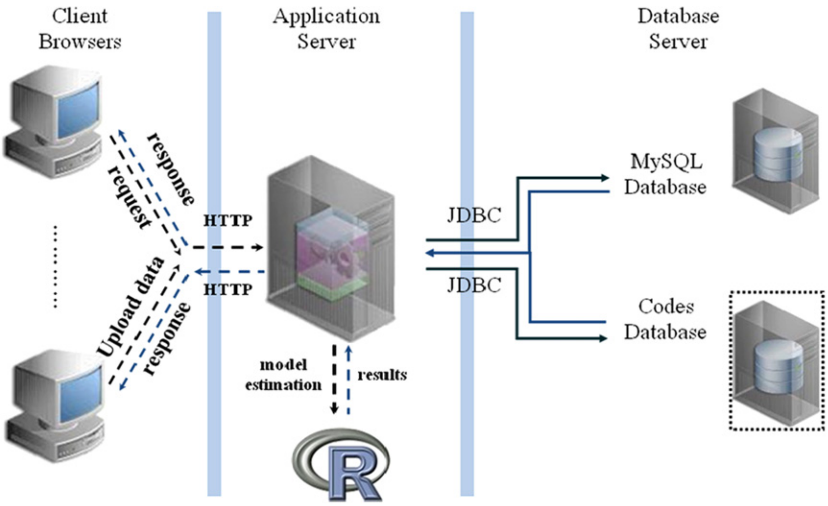


Fig. 2. TDFS flowchart.

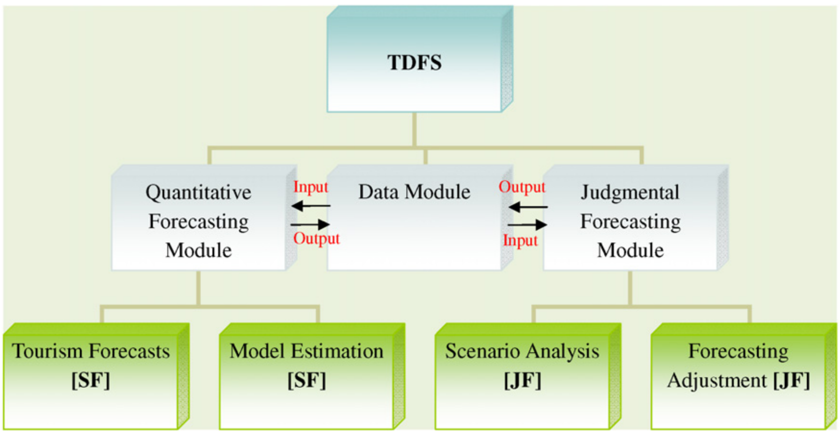


Fig. 3. TDFS components.

user's HTML instructions. The user can select the model by choosing different dependent and independent variables and conducting diagnostic checks on the model adequacy. The general-to-specific methodology is followed to obtain the final ADLM, which passes most of the diagnostic tests. After the final model has been confirmed by the user, forecasts of the dependent variables are generated and stored in the system database. In the third stage, users can adjust the statistical forecasts based on their domain knowledge if they believe that there is important information which is not captured by the econometric model. The TDFS consists of three functional modules: the data module, the quantitative forecasting module and the judgmental forecasting module (see Fig. 3).

3.2. Data module

Four types of tourism forecasts for Hong Kong were provided by the TDFS: tourist arrivals, tourist expenditure, the demand for hotel rooms (i.e., High Tariff A and B hotel rooms, Medium Tariff hotel rooms and Tourist Guesthouses) and expenditure by sector (i.e., hotels, shopping, meals, entertainment and tours). Quarterly data were collected for the first three types of tourism demand measures, whereas annual data were obtained for the sectoral demand over the period 1985–2009. Four main types of explanatory variables were included: (i) income, as measured by the real gross domestic product (GDP) index (2005 = 100); (ii) own price, as calculated from the exchange-rate-adjusted consumer price index (CPI); (iii) substitute price, defined as the weighted *CPI* in six competitive tourism destinations (i.e., mainland China, Malaysia, Singapore, South Korea, Taiwan and Thailand), adjusted to their share of international tourist arrivals; and (iv) dummy variables, including seasonal dummies and one-off event dummies (e.g., SARS, swine flu, the Olympic Games). The data and variables are described in detail by Chon, Li, Lin, and Gao (2010) and Song, Lin, Witt, and Zhang (2011); Song, Lin, Zhang, and Gao (2010).

To run the Model Estimation module, the user must first upload the data in Excel format. After the data have been uploaded, the system converts them into a template by listing all of the variables, such as those shown in Fig. 4. The data can be previewed in spreadsheet form (see Fig. 5).

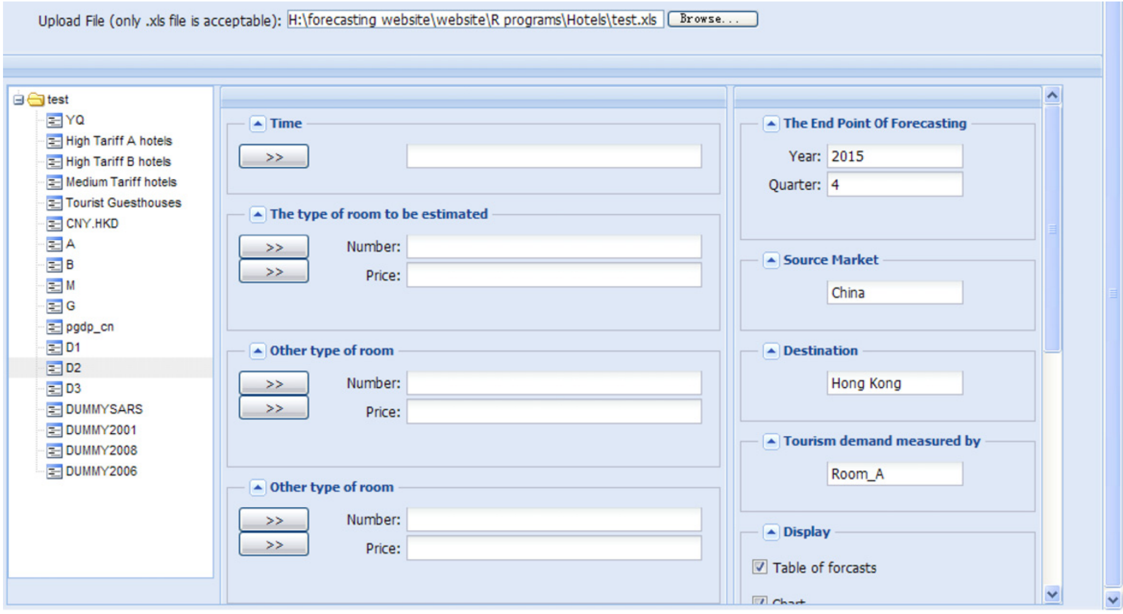
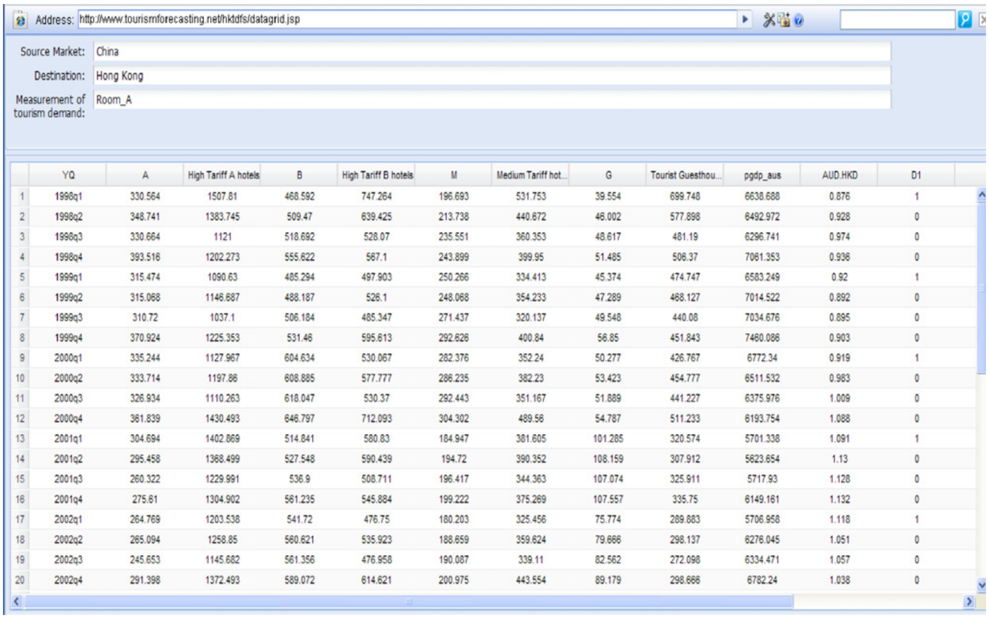


Fig. 4. Screen shot of uploaded data.



	YQ	A	High Tariff A hotels	B	High Tariff B hotels	M	Medium Tariff hot...	G	Tourist Guesthou...	pgdp_aus	AUD.HKD	D1
1	1998q1	330.564	1507.81	468.592	747.264	196.693	531.753	39.554	699.748	6638.688	0.876	1
2	1998q2	348.741	1383.745	509.47	639.425	213.738	440.672	46.002	577.898	6492.972	0.928	0
3	1998q3	330.664	1121	518.692	528.07	235.551	360.353	48.617	481.19	6296.741	0.974	0
4	1998q4	393.516	1202.273	555.622	567.1	243.899	399.95	51.485	506.37	7061.353	0.936	0
5	1999q1	315.474	1090.63	485.294	497.903	250.266	334.413	45.374	474.747	6583.249	0.92	1
6	1999q2	315.068	1146.687	488.187	526.1	248.068	354.233	47.289	468.127	7014.522	0.892	0
7	1999q3	310.72	1037.1	506.184	485.347	271.437	320.137	49.548	440.08	7034.676	0.895	0
8	1999q4	370.924	1225.353	531.46	595.613	292.626	400.84	56.85	451.843	7460.086	0.903	0
9	2000q1	335.244	1127.967	604.634	530.067	282.376	352.24	50.277	426.767	6772.34	0.919	1
10	2000q2	333.714	1197.86	608.885	577.777	286.235	382.23	53.423	454.777	6511.532	0.983	0
11	2000q3	326.934	1110.263	618.047	530.37	292.443	351.167	51.889	441.227	6375.976	1.009	0
12	2000q4	361.839	1430.493	646.797	712.093	304.302	489.56	54.787	511.233	6193.754	1.088	0
13	2001q1	304.694	1402.869	514.841	580.83	184.947	381.605	101.285	320.574	5701.338	1.091	1
14	2001q2	295.458	1368.499	527.548	590.439	194.72	390.352	108.159	307.912	5623.654	1.13	0
15	2001q3	260.322	1229.991	536.9	508.711	196.417	344.363	107.074	325.911	5717.93	1.128	0
16	2001q4	275.61	1304.902	561.235	545.884	199.222	375.269	107.557	335.75	6149.161	1.132	0
17	2002q1	264.769	1203.538	541.72	476.75	180.203	325.456	75.774	289.883	5706.958	1.118	1
18	2002q2	265.094	1258.85	560.621	535.923	188.659	359.624	79.666	298.137	6276.045	1.051	0
19	2002q3	245.653	1145.682	561.356	476.958	190.087	339.11	82.562	272.098	6334.471	1.057	0
20	2002q4	291.398	1372.493	589.072	614.621	200.975	443.554	89.179	298.666	6782.24	1.038	0

Fig. 5. Screen shot of the data presentation.

Users are then asked to select the dependent variable to be forecast, as well as the independent variables included in the model.

3.3. Quantitative forecasting module

Empirical studies suggest that econometric models are preferable to pure time series models in forecasting tourism demand because of their ability to (i) reflect the long-term relationships between tourism demand and its influencing factors, and (ii) interpret the changes in tourism demand by evaluating the effectiveness of tourism policies (Song & Li, 2008; Song & Witt, 2000; Song, Wong, & Chon, 2003; Witt, Song, & Wanhill, 2004). In this study, the initial statistical forecasts are produced using the ADLM. Full discussions of this method are provided by Chon et al. (2010) and Song et al. (2011). Briefly, a general-to-specific approach is employed in order to arrive at the final model. The models are estimated automatically by the R codes embedded in the system. A number of diagnostic tests, including the normality, autocorrelation, heteroscedasticity and misspecification tests, are run to

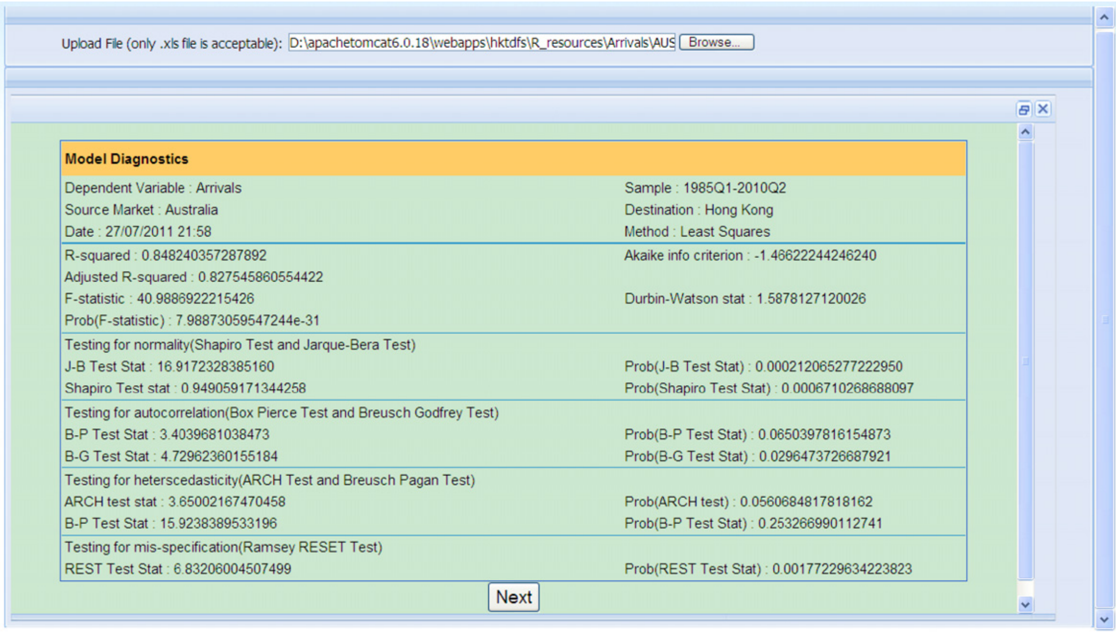


Fig. 6. Diagnostic statistics.

determine the final model to be used for forecasting (see Fig. 6).

As has been suggested by the tourism demand forecasting literature, reliable resources for forecasting the explanatory variables can be obtained from (i) international organizations such as the International Monetary Fund (IMF) and the UN World Tourism Organization (UNWTO) (Song & Li, 2008; Song & Witt, 2003), and (ii) statistics bureaus and/or government departments of the destination and source markets of interest. Alternatively, it is common to forecast the explanatory variables using time series methods. One of the forecasting methods most frequently adopted for such a purpose is the exponential smoothing approach (Hyndman, Koehler, Ord, & Snyder, 2005; Khandakar & Hyndman, 2008; Taylor, 2003). Here, we used two methods for forecasting the explanatory variables: (i) projections of real GDP growth rates obtained from the IMF were used to forecast the income variables from each of the selected source markets, (ii) and exponential smoothing via the state-space model (Hyndman, Koehler, Ord, & Snyder, 2008), available in the forecast package for R (Hyndman, 2011), was employed to forecast the price variables. These forecasts were then used in conjunction with the estimated relationships to generate forecasts of the dependent variables.

3.4. Judgmental forecasting module

After the statistical forecasts have been produced, the system allows users to incorporate their domain knowledge in them. Two modules, Scenario Analysis and Statistical Adjustment, are available to users for entering their judgmental inputs into the system (see Fig. 3). The Scenario Analysis module takes the statistical forecasts provided by the ADLM as the baseline forecasts, and these forecasts are then used as benchmarks for the scenario forecasts created by the user. This component offers four baseline scenarios (5% or 1% higher or lower than the benchmark growth rates), plus a customized scenario where users can input their own estimates (see Fig. 7(a)). When a specific scenario is submitted, the system will present the scenario forecasting results and the baseline statistical forecasts (see Fig. 7(b)) in both tabular and graphic formats. The system also allows users to revise the statistical forecasts by going back to the Model Estimation module.

Unlike the Scenario Analysis module, the Forecasting Adjustment module allows users to adjust the forecasts of both the dependent and independent variables. This component is also responsible for a dynamic Delphi survey procedure and includes a few features which are not included in the earlier TDFS version described by Song et al. (2008). These features include the following. First, the system presents both the historical data series and the forecasts, to permit experts to compare the historical trends of the time series with the forecasts easily. According to Benson and Önkal (1992) and Fildes et al. (2006), giving experts access to the latest observations of the time series can improve the accuracy of the adjusted forecasts. Second, the system allows the experts to give the reasons for their adjustments of the statistical forecasts. Previous studies have suggested that recording these reasons is an effective way to structure the adjustments and improve the accuracy of judgmental forecasts (Armstrong, 2006; Goodwin, 2000; Rowe & Wright, 1999). Third, the system also allows group feedback to be recorded, permitting the experts to refer to it during later rounds of forecasting adjustments. Lawrence, Goodwin, O'Connor, and Önkal (2006), O'Connor (1989) and Rowe and Wright (1999) concluded that feedback improves the accuracy of statistical forecasts.

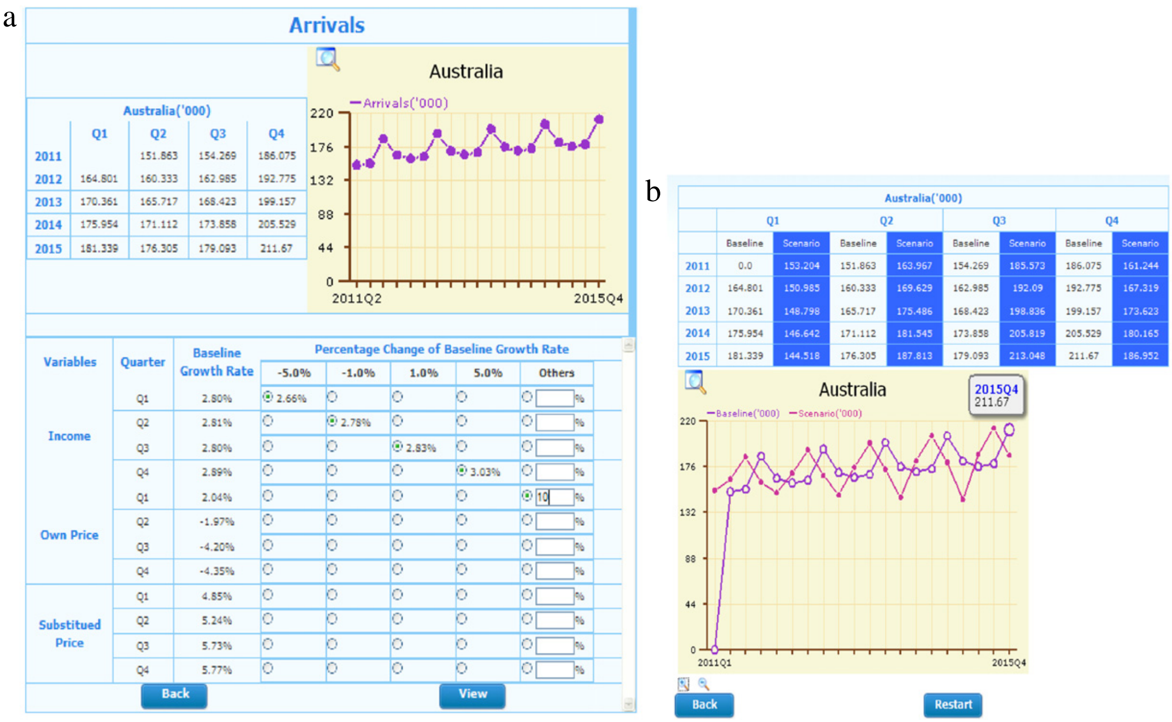


Fig. 7. Screen shots of the scenario analysis.

Only Delphi experts, authorized users and full subscribers are able to access the Forecasting Adjustment module. In each round of the Delphi survey, this module provides panellists with two alternative ways of making their judgmental adjustments: (i) by changing the point forecasts of the dependent variables by year or by individual quarters over the specified forecasting period, and (ii) by changing the growth rates of the determinant variables of tourism demand, as in the Scenario Analysis module. Upon the completion of each round of the Delphi survey, the module publishes the final group forecasts (or median forecasts), which can be accessed by all of the experts. Additional details and screen shots of the Forecasting Adjustment module are included in the following section.

4. A case study

4.1. The dynamic Delphi survey via TDFS

To test the reliability of the TDFS Statistical Adjustment module, an experiment was carried out with the involvement of postgraduate students and staff from the School of Hotel and Tourism Management at The Hong Kong Polytechnic University. This two-round survey was undertaken over the period June 5–11, 2011. Letters seeking participants were sent to a group of postgraduate students enrolled in a doctoral-level quantitative methods course and to researchers working on tourism-related projects. The final panel consisted of 21 students and five research staff members. All participants were asked to self-rate their level of expertise in tourism demand forecasting on a 7-point Likert scale, ranging from very little (1) to excellent (7); the 16 participants who were included in the survey had a mean self-rating score of 4.07. Of these 16 participants, 19% rated themselves as having very little expertise, whereas 6% rated themselves as having level 6 experience (see Fig. 8). About 38% rated themselves as level 5%, 32% fell in levels 3 and 4, and the remaining 30% were in other levels.

Before carrying out the forecasting tasks, the participants were asked to read the introductory document, to familiarize themselves with the procedure for using the TDFS. Students who confirmed their desire to participate were each assigned a user ID and a password to enable them to access the TDFS. Participants were invited to make adjustments to the quarterly forecasts of tourist arrivals from three short-haul markets (China, Taiwan and Japan) and three long-haul markets (the US, the UK and Australia) served by the Hong Kong tourism industry over the period 2010Q1–2015Q4. Statistical forecasts were produced by the ADLM using the sample 1985Q1–2009Q4. Participants were asked to consider the effects of two special events which were not taken into account by the ADLM, namely the 2011 earthquake in Japan and the launch of the Beijing–Shanghai high-speed railway. Annual projections of real GDP growth rates and the exchange rates obtained from the IMF for the six source markets were also provided to the participants.

One major characteristic of the TDFS is that it is user-oriented, in that it allows the user to make a wide variety of interventions through judgmental adjustments, incorporating the effects of special events. Two options were available to users: making annual or quarterly adjustments. In addition, users could also make adjustments for different

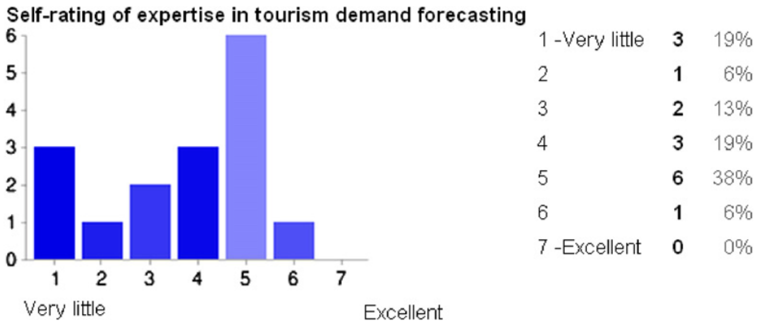


Fig. 8. Self-rating of expertise by participants.

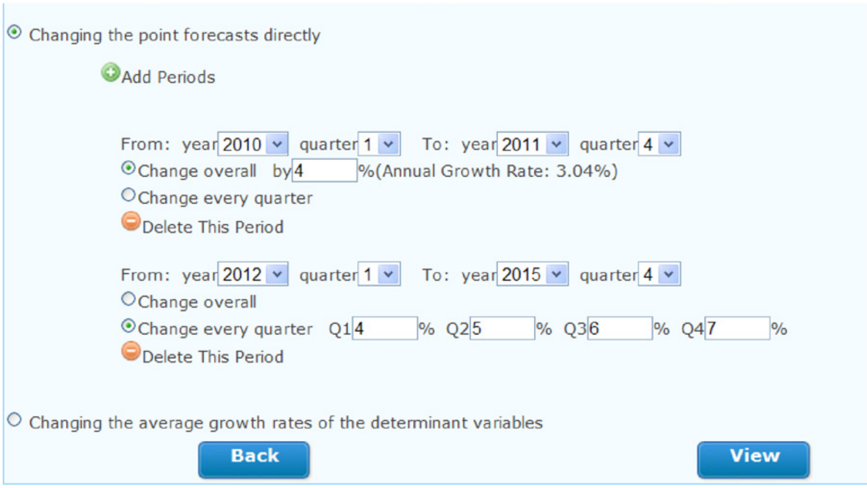


Fig. 9. Options for adjustments.

forecasting periods (see Fig. 9). The statistical forecasts for the period 2010Q1–2015Q4 were presented, together with the individual forecasts in both tabular and graphical form (see Fig. 10(a)). Individual forecasts were set to match the statistical forecasts. Users could choose different output options (with both historical data and forecasts or with forecasts only), as shown in Fig. 10(a). Moreover, users' justifications for their adjustments could be stored within the TDFS for subsequent reference (see Fig. 10(b)).

In the first round, positive responses were received from 56.5% of the selected panelists. In the second round, the median forecasts from Round 1 were presented as the baseline forecasts (see Group Adjustments in Fig. 11(a)). The participants were then required to verify and adjust these group forecasts. The system also allows access to the statistical summary report and written justifications (see Fig. 11(b)), with a view to informing the participants of the adjustments made by other participants and the reasons for these adjustments.

4.2. Evaluation of forecast accuracy

Three error measures, namely the absolute percentage error (APE), the mean absolute percentage error (MAPE) and the root mean square percentage error (RMSPE), were used to evaluate the forecast accuracy. Out-of-sample forecast errors were generated for the period 2010Q1–2011Q2. As anticipated, the judgmentally adjusted forecasts were more accurate than the statistical forecasts (i.e., the average MAPEs decreased from 8.86% to 8.02% and the average RMSPEs from 10.41% to 9.33%). Furthermore, the accuracy also increased over the rounds in terms of both the MAPE and RMSPE (see Table 1). Even though the forecast adjustments improved the forecast accuracy on average, the level of improvement varied across source markets: the accuracy improved for all three short-haul markets and one long-haul market (the US), but decreased for Australia and Taiwan (see Table 1). In other words, the overall improvement came largely from the contribution of the three short-haul markets. The MAPEs of the statistical forecasts from the long-haul markets were far smaller than those from the short-haul markets; however, the forecasting performance deteriorated for Australia and the UK after judgmental adjustments. When the statistical forecasts were highly accurate, judgmental adjustment seemed to have little impact on the accuracy, or even reduced it. One possible explanation for this finding may lie in the capacity of econometric models to make accurate extrapolations and identify established patterns and existing relationships, and thus produce highly accurate forecasts (e.g., all MAPEs for the three long-haul markets were less than 8%). Under such conditions, judgmental revisions of the statistical forecasts may tend to overreact to fluctuations in the arrivals series (Sanders & Ritzman, 1995).

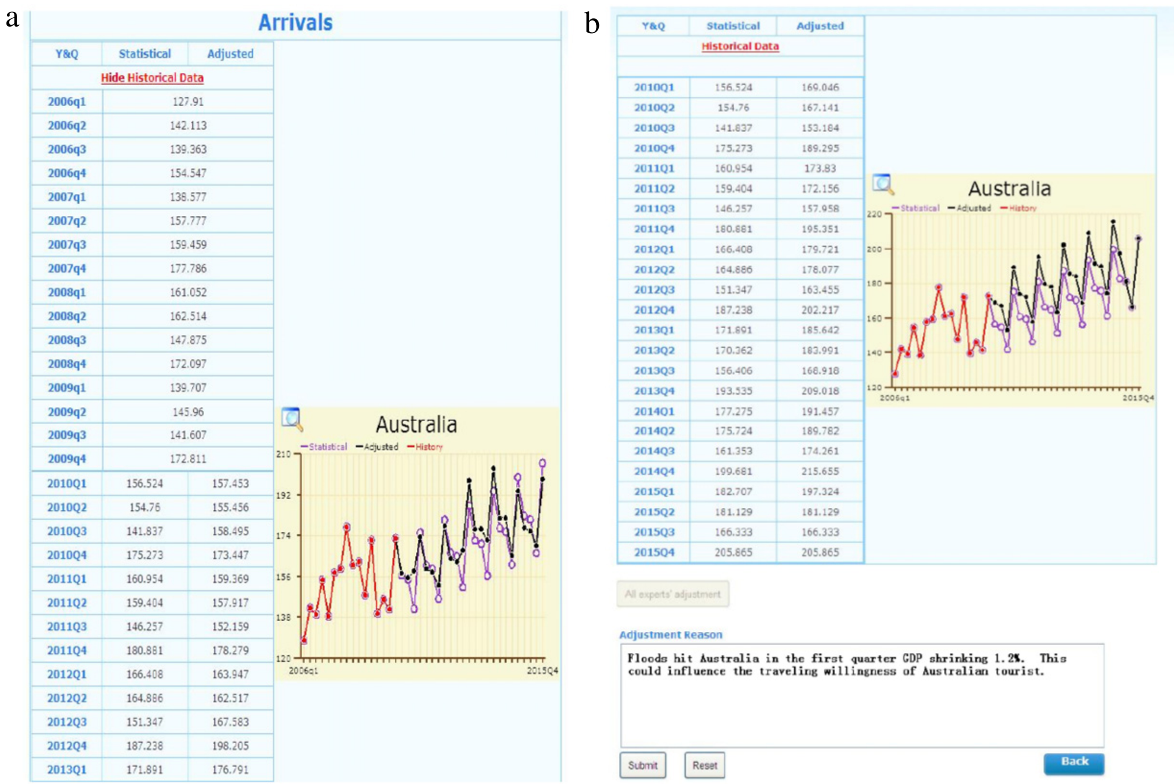


Fig. 10. Screen shot from TDFS (first round survey).

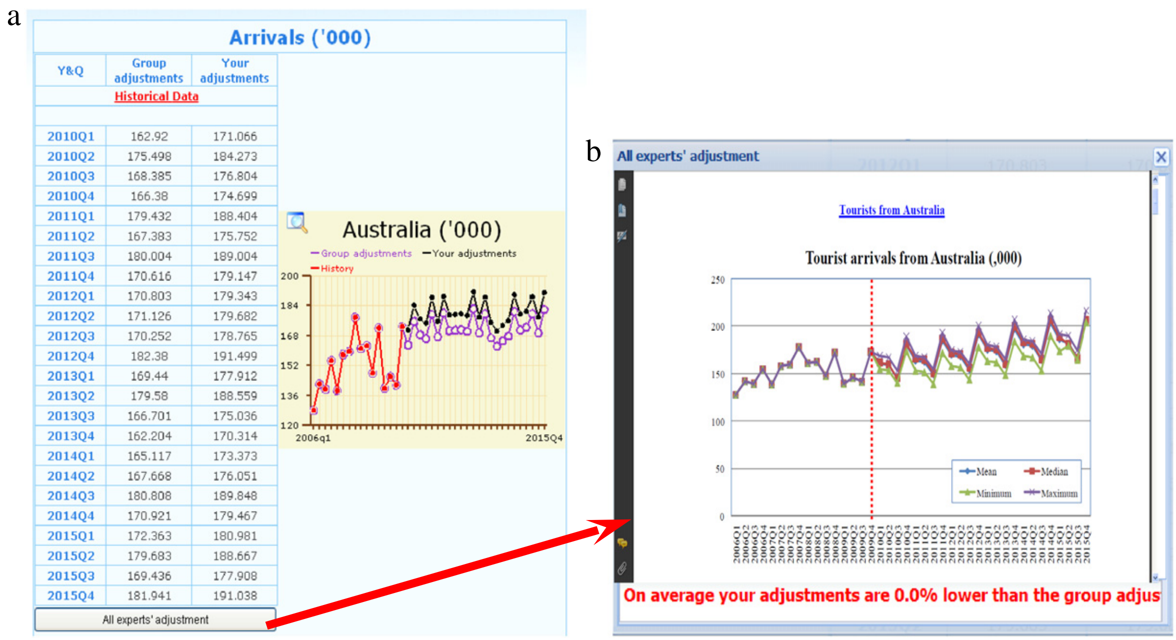


Fig. 11. Screen shots from TDFS (second round survey).

Another factor which contributed to the forecasting performance is the volatility of the time series being forecast. As is shown in Fig. 12, the historical data series of tourist arrivals for the short-haul markets are more volatile (or less stable) than those for the long-haul markets. This is reflected in the coefficient variation for the six markets: 1.1, 0.39 and 0.32 respectively for China, Taiwan and Japan, and 0.33, 0.25 and 0.40 respectively for the UK, the US

Table 1
MAPE and RMSPE.

Country	MAPE (%)			RMSPE (%)		
	SF	GF1	GF2	SF	GF1	GF2
China	16.19	11.91	11.29	17.64	13.78	13.03
Taiwan	10.74	8.80	9.02	12.20	10.02	10.26
Japan	8.45	7.91	7.87	11.61	10.49	10.65
Australia	2.94	3.30	4.88	4.64	4.26	5.29
UK	7.58	11.54	10.47	8.95	12.47	11.55
US	7.25	4.69	4.39	7.41	4.93	4.64
Mean_{Short-haul}	11.79	9.54	9.39	13.82	11.43	11.31
Mean_{Long-haul}	5.93	6.51	6.58	7.00	7.22	7.16
Mean_{Total}	***8.86***	***8.02***	***7.99***	***10.41***	***9.33***	***9.24***
Error reduction (%)	GF1–SF	GF2–SF	GF2–GF1	GF1–SF	GF2–SF	GF1–GF2
China	−4.28	−4.90	−0.62	−3.86	−4.61	−0.75
Taiwan	−1.94	−1.72	0.22	−2.19	−1.94	0.24
Japan	−0.54	−0.58	−0.04	−1.12	−0.96	0.16
Australia	0.35	1.94	1.59	−0.38	0.66	1.03
UK	3.95	2.88	−1.07	3.52	2.59	−0.92
US	−2.57	−2.87	−0.30	−2.48	−2.77	−0.29
Mean_{Short-haul}	*−2.26*	*−2.40*	*−0.14*	*−2.39*	*−2.50*	*−0.12*
Mean_{Long-haul}	*0.58*	*0.65*	*0.07*	*0.22*	*0.16*	*−0.06*
Mean	***−0.84***	***−0.87***	***−0.04***	***−1.08***	***−1.17***	***−0.09***
Percentage reduction (%)	GF1–SF	GF2–SF	GF2–GF1	GF1–SF	GF2–SF	GF2–GF1
China	−26.44	−30.25	−5.18	−21.87	−26.13	−5.45
Taiwan	−18.09	−16.03	2.52	−17.92	−15.93	2.43
Japan	−6.45	−6.89	−0.48	−9.62	−8.27	1.48
Australia	12.05	65.95	48.11	−8.15	14.16	24.28
UK	52.09	37.98	−9.28	39.27	28.97	−7.39
US	−35.39	−39.52	−6.39	−33.46	−37.41	−5.94
Mean_{Short-haul}	−16.99	−17.72	−1.05	−16.47	−16.78	−0.51
Mean_{Long-haul}	9.58	21.47	10.81	−0.78	1.91	3.65
Mean	***−3.71***	***1.87***	***4.88***	***−8.62***	***−7.44***	***1.57***

Note: SF, GF1 and GF2 represent the econometric (statistical) forecasts, group 1 forecasts and group 2 forecasts, respectively. MAPE = mean absolute percentage error; RMSPE = root mean square percentage error.

and Australia. Sanders and Ritzman (2004) suggest that less emphasis should be placed on contextual knowledge when making combination forecasts if the variability is low.

Error reductions between each of the pairs of forecasts, namely the statistical forecasts and the two rounds of judgmental forecasts, were computed and are shown in Table 1. A negative change in either the MAPE or RMSPE indicates an improvement in accuracy. The greatest improvement in accuracy over the statistical forecasts was in the predictions of tourist arrivals from the US, followed by those from mainland China. The big improvement in forecasting American visitors to Hong Kong may be due to the provision of useful feedback from participants. For example, one panellist pointed out that "with some signs of recovery from the global financial crisis, arrivals from the USA can improve faster than the statistical trend", and this turned out to be the case after comparing the statistical forecasts with the actual arrivals over the period 2010Q1–2011Q2. As was discussed earlier, the series of Chinese tourist arrivals was the most unstable one, with the largest coefficient of variation (1.1). Judgmental inputs for these series could significantly improve the forecast accuracy (Sanders & Ritzman, 1995). The smallest improvement over the statistical forecasts made by judgmental interventions was in predicting tourist arrivals from the UK. However, as is shown in Table 1, the greatest improvement in accuracy over the two rounds was achieved in the case of the UK (9.28% and 7.39% reductions in MAPE and RMSPE, respectively), followed by the US and China.

A more detailed analysis of the performance statistics (see Table 2) reveals that the group forecasts from the second round were more accurate than either the statistical forecasts or the group forecasts from the first round. Taking China as an example, the APEs of the three sets of forecasts were calculated for each quarter between 2010Q1 and 2011Q2. The cumulative frequencies of negative error differences between the two forecasts, measured by the APE, are given as percentages. As Table 2 shows, an improvement in forecast accuracy as a result of using the combination method (versus statistical forecasting alone) was observed for all quarters in the cases of China, Taiwan and the US; however, the accuracy of the combination method decreased in the case of the UK.

Fig. 13 shows the forecasting performances of 13 of the individual participants involved in the survey for both rounds. Seven of the 13 produced out-of-sample forecasts that were better than the statistical forecasts (according to both the MAPE and RMSPE measures). In order to test whether the overall group performance improved over the rounds, the differences were analyzed using a paired t-test. The judgmental group forecasts in the second round were shown to be significantly more accurate than those in the first round at the 10% significance level, as measured by the MAPE ($t(12) = -1.418$, $p =$

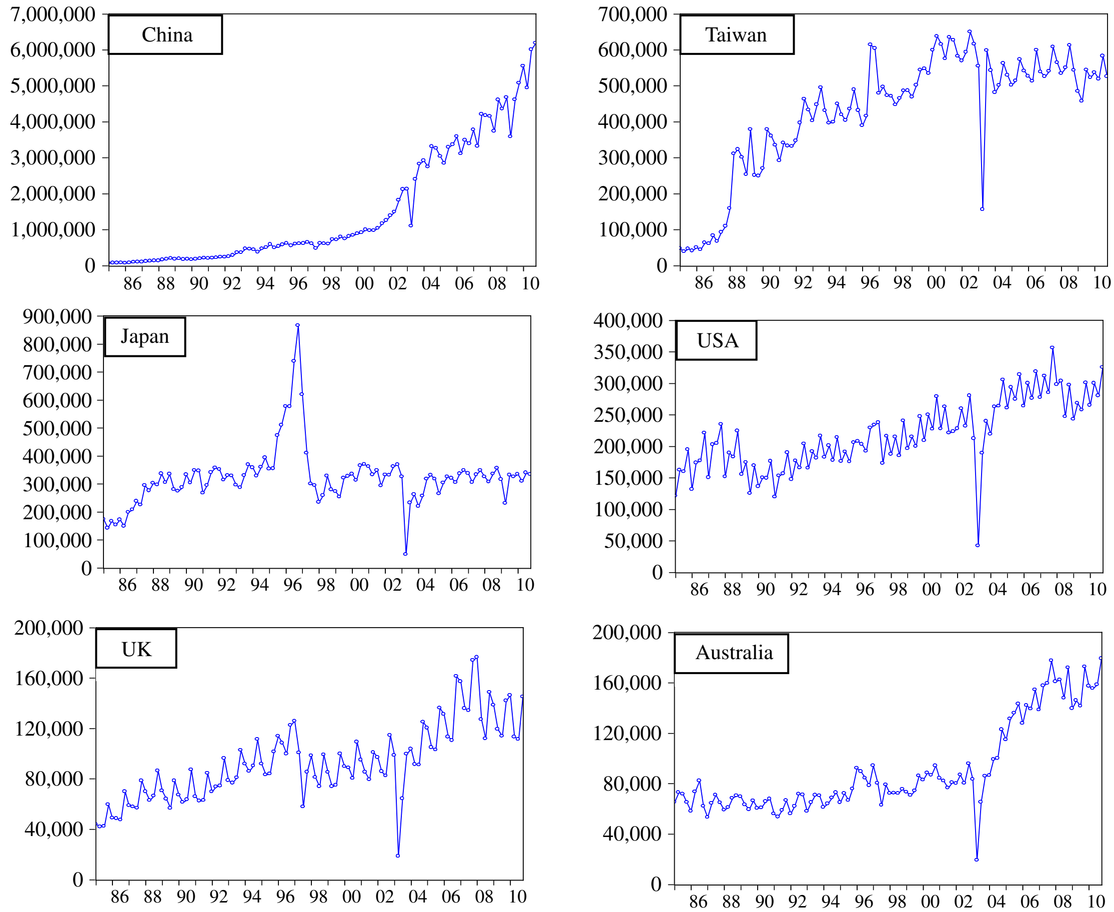


Fig. 12. Historical trend of tourist arrivals.

0.091). This was further confirmed to be the case by the RMSPE value ($t(12) = -1.737, p = 0.054$). Thus, the group performance improved significantly with the use of the Delphi approach. This finding is consistent with the results of previous studies, reporting that incorporating forecasters' technical and contextual knowledge into the statistical forecasts helps to improve the forecast accuracy (Sanders & Ritzman, 1995). The participants in this study were postgraduate (mainly Ph.D.) students and research assistants in tourism and hospitality management with a certain degree of domain knowledge about the development and prospects of the Hong Kong tourism industry. The effects of several special events that occurred during the forecast period were not considered in the statistical models but were incorporated as judgmental inputs, including the recovery from the 2008 financial crisis, the floods in Australia, the earthquake in Japan, the construction of the high-speed railway between Hong Kong and mainland China, and the London Olympic Games slated for 2012. In addition, the majority of the participants (76%) were well-trained in quantitative methods, which helped them to understand the statistical forecasting procedure.

4.3. Feedback survey

A feedback survey was distributed to all of the participants to examine how they perceived the effectiveness of the TDFS. The questionnaire was structured carefully using a 5-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (5) for all statements. Among the 12 positive responses, over half (58%) of the participants agreed that the "forecasting system is easy to use", whereas the rest had no strong opinion. 25% of the participants strongly agreed that they had a clear understanding of what they were expected to do in the forecasting tasks after reading the instructions. 42% "agreed" (albeit not strongly), and only 8% felt that they were unclear about the tasks even when they were provided with a step-by-step video demonstration in addition to the written instructions. Regarding the time needed to complete the survey, 67% of the participants thought that the time allotted (ranging from 20 to 40 min) was appropriate. When asked to evaluate the statistical feedback from the first round of the survey, half of the participants agreed that it was useful for assisting their adjustments in the second round. The participants indicated that the tabular and graphical data summaries were useful: 17% "strongly agreed" and 75% "agreed" that the graphical presentation was useful, whereas 8% "strongly agreed" and 67%

Table 2
Forecast performance evaluated by APE.

Country	Quarter	APE_{SF} (a)	APE_{GF1} (b)	APE_{GF2} (c)	$\%_{(b-a<0)}$	$\%_{(c-a<0)}$	$\%_{(c-b<0)}$
China	2010Q1	11.38	7.58	7.60	100	100	83
	2010Q2	21.96	17.73	17.32			
	2010Q3	15.22	11.34	10.87			
	2010Q4	6.94	2.29	1.85			
	2011Q1	13.44	8.98	8.48			
	2011Q2	28.18	23.51	21.62			
Taiwan	2010Q1	14.54	12.19	13.53	83	100	50
	2010Q2	14.45	11.79	11.16			
	2010Q3	7.36	4.94	5.83			
	2010Q4	0.63	1.60	0.54			
	2011Q1	9.18	6.69	8.08			
	2011Q2	18.30	15.58	14.98			
Japan	2010Q1	5.71	6.05	5.38	50	83	50
	2010Q2	24.33	21.79	21.95			
	2010Q3	3.43	3.64	3.16			
	2010Q4	2.18	1.83	1.97			
	2011Q1	2.32	3.02	2.53			
	2011Q2	12.74	11.12	12.22			
Australia	2010Q1	0.59	1.96	3.93	33	33	17
	2010Q2	0.45	2.11	4.05			
	2010Q3	10.51	8.28	6.98			
	2010Q4	2.23	0.34	1.92			
	2011Q1	3.58	5.37	8.04			
	2011Q2	0.29	1.72	4.37			
UK	2010Q1	0.43	4.23	3.09	0	0	100
	2010Q2	11.13	15.41	14.23			
	2010Q3	6.41	10.44	9.16			
	2010Q4	2.79	6.90	5.66			
	2011Q1	11.42	15.24	14.23			
	2011Q2	13.33	16.99	16.42			
US	2010Q1	6.47	4.21	4.40	100	100	67
	2010Q2	8.56	5.70	4.96			
	2010Q3	5.80	3.18	2.60			
	2010Q4	5.07	2.37	2.30			
	2011Q1	8.52	6.30	6.53			
	2011Q2	9.09	6.36	5.52			
Mean		**8.86**	**8.02**	**7.99**	**61**	**69**	**61**

Note: % denotes the frequency of a smaller APE (absolute percentage error) between any two forecasts among the SF (statistical forecast), GF1 (Round 1 forecast) and GF2 (Round 2 forecast).

"agreed" that the tabular information was useful. The majority of the respondents (84%) "agreed" that the "graphs on the website are more informative than the tables". Regarding the amount of historical data, approximately 60% of the participants agreed that the current system provided sufficient data to assist them with the adjustments. The participants were also asked to provide suggestions on the amount of historical data needed. 42% suggested data covering the previous five years, 42% suggested 10 years, and the remaining 16% suggested periods ranging from less than 5 years to more than 10 years.

5. Conclusion and discussion

The web-based TDFS is proposed as an innovative online platform for tourism demand forecasting that takes full advantage of web technologies and advanced tourism demand forecasting techniques. Like other web-based systems, the TDFS has four main attributes: wide accessibility, flexibility, reusability and user-friendliness. In addition, various new features distinguish the current TDFS from other forecasting support systems. They include (i) integrating statistical and judgmental forecasts through a dynamic online Delphi survey, (ii) creating different scenarios based on user-customized specifications, and (iii) applying JSP, which provides a connection to the R Engine. One significant benefit to tourism practitioners is that the TDFS allows the integration of quantitative and judgmental forecasts in a web-based forecasting system.

Overall, the experimental study showed that a greater forecast accuracy was achieved with the judgmentally adjusted statistical forecasts than with the statistical forecasts alone. In addition, the combined statistical and judgmental forecasts improved the forecast accuracy for four of the six source markets of interest (namely China, Taiwan, Japan and the US). The long-haul markets tended

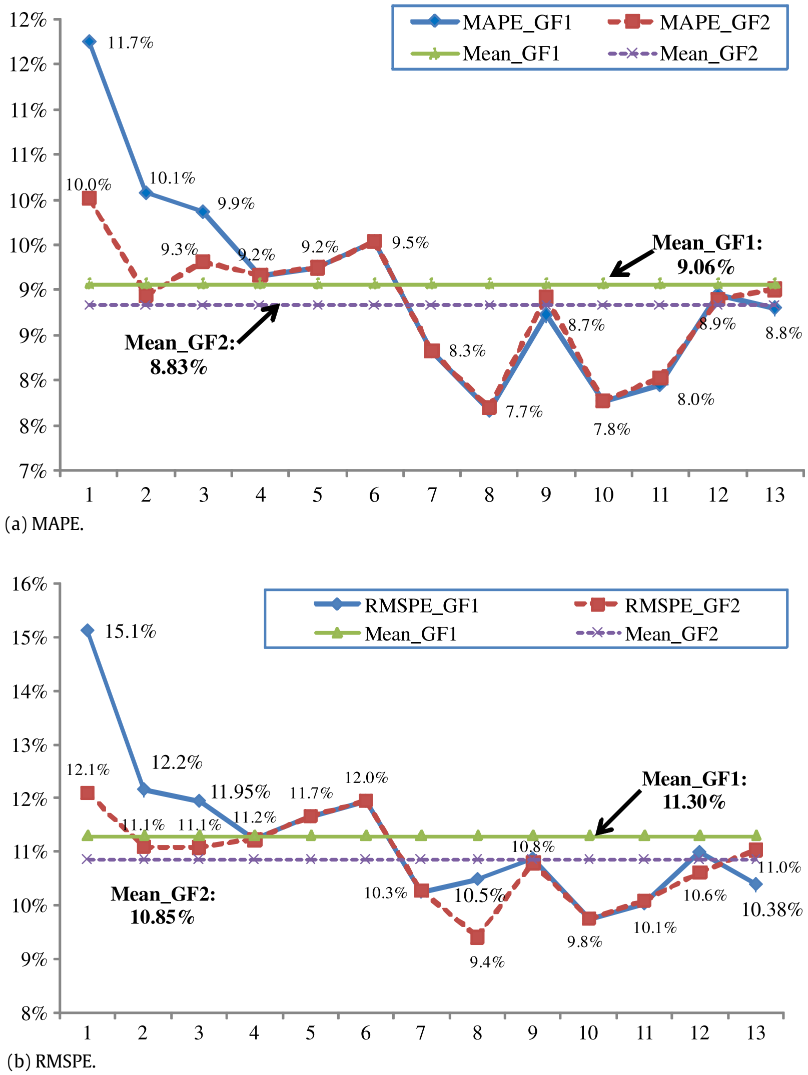


Fig. 13. Individual participants' forecasting performances over rounds.

to produce more accurate forecasts than the short-haul markets; however, more remarkable improvements were found for the three short-haul markets. This is probably due to the relatively more stable data patterns for the arrivals data of the long-haul markets, and the high capacity of the econometric models, which produce very accurate forecasts. Thus, including judgmental inputs for such series did not significantly improve the forecast accuracy; on the contrary, it harmed the forecast accuracy. The benefits of including judgmental inputs in quantitative forecasts depend on the characteristics of the data series being examined. These results suggest that the accuracy of the judgmental forecasts increased in the second round relative to the first round, with a reduction of the MAPE from 8.02% to 7.99%. The paired t-test results confirmed that there is a significant reduction in the MAPE and RMSPE over two rounds.

The TDFS's forecasting performance is achieved through the following factors. First, an advanced econometric modelling method (i.e., the ADLM) is used to estimate the demand model for each source market. Second, the TDFS provides flexible adjustment options for the forecasters to adjust their forecasts by either the year or the quarter over different forecasting periods. Third, the system provides useful feedback about the summary forecasts generated by all experts in the early rounds, with both high-resolution graphs and tables, so that the Delphi experts are well informed for subsequent adjustments. Fourth, the use of a web-based platform allows users to access the system anytime and anywhere, and allows collaboration between individuals in different geographic locations and representing different knowledge domains. Finally, participants who have a high level of technical knowledge of tourism demand forecasting, as well as some degree of contextual knowledge of the Hong Kong tourism industry, may contribute to the improvement in accuracy. This is supported by Sanders and Ritzman's (1995) finding that the combination of statistical forecasts and judgmental forecasts based on contextual knowledge could lead to significantly more accurate forecasts. It is worth noting that this study did not distinguish the factors that influence the accuracy of judgmental inputs into the econometric model or the forecasting system. A number of factors, such as the form of data presentation (with tables and graphs), the provision of feedback over different rounds of the Delphi survey, the capability of good functional forecasting modules, clear instructions for the Delphi survey, and the inclusion of historical data series, can lead to a good forecasting performance. Further research is thus needed to investigate which attributes contribute most to the improvement in accuracy.

Suggestions of further ways to improve the TDFS include the following. First, additional quantitative models, such as the naïve, exponential smoothing and Box–Jenkins models, should be included as benchmarks. The rationale for this is that combining two or more quantitative methods normally leads to an improved forecast accuracy (Wong, Song, Witt, & Wu, 2007). Second, more advanced econometric forecasting techniques, such as the time-varying parameter (TVP) model, as well as its variants, could be used to produce the baseline forecasts. Unlike the ADLM approach, this method relaxes the assumption of parameter constancy, and the behavioural change of tourists over time is traced using a statistical estimator known as a Kalman filter. The appropriateness of the TVP approach in tourism demand modelling has been tested by Song and Wong (2003), based on Hong Kong tourism demand data from six major source markets. Li, Song, and Witt (2006) developed a time-varying parameter linear almost ideal demand system (TVP-LAIDS) model and concluded that such a model improved the forecasting performance remarkably compared to the original static version of its fixed-parameter error correction counterparts in modelling and forecasting the demand for tourism in Western European destinations by UK residents. Third, a more user-friendly guidance system could be developed to help forecasters from industry – either with or without relevant technical knowledge – to interpret the diagnostic test results easily, and thus facilitate the development of a better decision-making process. The current TDFS presents the diagnostic test results but provides no further explanations. This is likely to challenge the usability of the system, particularly for those who lack such technical knowledge of tourism demand modelling and forecasting. Finally, the inclusion of an online forum connected to the Statistical Adjustment module may help experts to improve their forecasting performances, as it would provide rapid feedback for Delphi panel members. Recent research suggests that the accuracy of judgmental forecasts can be improved greatly if timely feedback is supplied to the panel experts (Rowe & Wright, 1996). One additional benefit of such an online forum would be that, in addition to its anonymity, it would provide a more cost-efficient means of communication among the participants. This enhanced communication might also help the experts to focus more on the tasks at hand, ensure equal participation and stimulate the generation of ideas (Dalal, Khodyakov, Srinivasan, Straus, & Adams, 2011).

References

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers.

Armstrong, J. S. (2006). Findings from evidence-based forecasting: methods for reducing forecast error. *International Journal of Forecasting*, *22*(3), 583–598.

Armstrong, J. S., & Collopy, F. (1998). Integration of statistical methods and judgment for time series forecasting: principles from empirical research. In G. Wright, & P. Goodwin (Eds.), *Forecasting with judgment* (pp. 269–293). New York: John Wiley & Sons.

Benson, P. G., & Önkal, D. (1992). The effects of feedback and training on the performance of probability forecasters. *International Journal of Forecasting*, *8*(4), 559–573.

Blattberg, R. C., & Hoch, S. J. (1990). Database models and managerial intuition: 50% model + 50% manager. *Management Science*, *36*(8), 887–899.

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: an integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, *101*(2), 127–151.

Bunn, D., & Wright, G. (1991). Interaction of judgemental and statistical forecasting methods: issues and analysis. *Management Science*, *37*(5), 501–518.

Chon, K. K. S., Li, G., Lin, S., & Gao, Z. (2010). Recovery of tourism demand in Hong Kong from the global financial and economic crisis. *Journal of China Tourism Research*, *6*(3), 259–278.

Clements, M. P., & Hendry, D. F. (1998). *Forecasting economic time series*. Cambridge: Cambridge University Press.

Collopy, F., & Armstrong, J. S. (1992). Rule-based forecasting: development and validation of an expert systems approach to combining time series extrapolations. *Management Science*, *38*(10), 1394–1414.

Croce, V., & Wober, K. W. (2011). Judgmental forecasting support systems in tourism. *Tourism Economics*, *17*(4), 709–724.

Dalal, S., Khodyakov, D., Srinivasan, R., Straus, S., & Adams, J. (2011). ExpertLens: a system for eliciting opinions from a large pool of non-collocated experts with diverse knowledge. *Technological Forecasting and Social Change*, *78*(8), 1426–1444.

Delen, D., Sharda, R., & Kumar, P. (2007). Movie forecast Guru: a web-based DSS for Hollywood managers. *Decision Support Systems*, *43*(4), 1151–1170.

Donihue, M. R. (1993). Evaluating the role judgment plays in forecast accuracy. *Journal of Forecasting*, *12*(2), 81–92.

Fildes, R., Goodwin, P., & Lawrence, M. (2006). The design features of forecasting support systems and their effectiveness. *Decision Support Systems*, *42*(1), 351–361.

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, *25*(1), 3–23.

Frechtling, D. C. (2001). *Forecasting tourism demand: methods and strategies*. Oxford: Butterworth-Heinemann.

Goodwin, P. (2000). Improving the voluntary integration of statistical forecasts and judgment. *International Journal of Forecasting*, *16*(1), 85–99.

Goodwin, P., & Fildes, R. (1999). Judgmental forecasts of time series affected by special events: does providing a statistical forecast improve accuracy? *Journal of Behavioral Decision Making*, *12*(1), 37–53.

Goodwin, P., Fildes, R., Lawrence, M., & Nikolopoulos, K. (2007). The process of using a forecasting support system. *International Journal of Forecasting*, *23*(3), 391–404.

Hannan, E. J., & Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika*, *69*(1), 81–94.

Harvey, N. (1995). Why are judgments less consistent in less predictable task situations? *Organizational Behavior and Human Decision Processes*, *63*(3), 247–263.

Hyndman, R. J. (2011). Forecast: forecasting functions for time series. R Package Version 3.16. http://CRAN.R-project.org/package=forecast.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2005). Prediction intervals for exponential smoothing using two new classes of state space models. *Journal of Forecasting*, *24*(1), 17–37.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Attadale, Australia: Springer.

Khandakar, Y., & Hyndman, R. J. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, *27*(3), 1–22.

Klassen, R. D., & Flores, B. E. (2001). Forecasting practices of Canadian firms: survey results and comparisons. *International Journal of Production Economics*, *70*(2), 163–174.

Kottemann, J. E., Davis, F. D., & Remus, W. E. (1994). Computer-assisted decision making: performance, beliefs, and the illusion of control. *Organizational Behavior and Human Decision Processes*, *57*(1), 26–37.

Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: a review of progress over the last 25 years. *International Journal of Forecasting*, *22*(3), 493–518.

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

Li, G., Song, H., & Witt, S. F. (2006). Time varying parameter and fixed parameter linear AIDS: an application to tourism demand forecasting. *International Journal of Forecasting*, *22*(1), 57–71.

Lim, J. S., & O'Connor, M. (1995). Judgemental adjustment of initial forecasts: its effectiveness and biases. *Journal of Behavioral Decision Making*, *8*(3), 149–168.

Lim, J. S., & O'Connor, M. (1996). Judgmental forecasting with time series and causal information. *International Journal of Forecasting*, *12*(1), 139–153.

Liu, L. M. (1989). Identification of seasonal ARIMA models using a filtering method. *Communications in Statistics—Theory and Methods*, *18*(6), 2279–2288.

McCarthy, T. M., Davis, D. F., Golicic, S. L., & Mentzer, J. T. (2006). The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices. *Journal of Forecasting*, *25*(5), 303–324.

McNees, S. K. (1990). The role of judgment in macroeconomic forecasting accuracy. *International Journal of Forecasting*, *6*(3), 287–299.

Mélard, G., & Pasteels, J. M. (2000). Automatic ARIMA modeling including interventions, using time series expert software. *International Journal of Forecasting*, *16*(4), 497–508.

O'Connor, M. (1989). Models of human behavior and confidence in judgment: a review. *International Journal of Forecasting*, *5*(2), 159–169.

O'Connor, M., Remus, W., & Griggs, K. (1993). Judgemental forecasting in times of change. *International Journal of Forecasting*, *9*(2), 163–172.

Petropoulos, C., Patelis, A., Metaxiotis, K., Nikolopoulos, K., & Assimakopoulos, V. (2003). SFTIS: a decision support system for tourism demand analysis and forecasting. *Journal of Computer Information Systems*, *44*(1), 21–32.

Rowe, G., & Wright, G. (1996). The impact of task characteristics on the performance of structured group forecasting techniques. *International Journal of Forecasting*, *12*(1), 73–89.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, *15*(4), 353–375.

Sanders, N. R., & Manrodt, K. B. (1994). Forecasting practices in US corporations: survey results. *Interfaces*, *24*(2), 92–100.

Sanders, N. R., & Ritzman, L. P. (1992). The need for contextual and technical knowledge in judgmental forecasting. *Journal of Behavioral Decision Making*, *5*(1), 39–52.

Sanders, N. R., & Ritzman, L. P. (1995). Bringing judgment into combination forecasts. *Journal of Operations Management*, *13*(4), 311–321.

Sanders, N. R., & Ritzman, L. P. (2004). Integrating judgmental and quantitative forecasts: methodologies for pooling marketing and operations information. *International Journal of Operations and Production Management*, *24*(5–6), 514–529.

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

Song, H., Lin, S., Witt, S. F., & Zhang, X. (2011). Impact of financial/economic crisis on demand for hotel rooms in Hong Kong. *Tourism Management*, *32*(1), 172–186.

Song, H., Lin, S., Zhang, X., & Gao, Z. (2010). Global financial/economic crisis and tourist arrival forecasts for Hong Kong. *Asia Pacific Journal of Tourism Research*, *15*(2), 223–242.

Song, H., & Witt, S. F. (2000). *Tourism demand modelling and forecasting: modern econometric approaches*. Cambridge: Pergamon Press.

Song, H., & Witt, S. F. (2003). Tourism forecasting: the general-to-specific approach. *Journal of Travel Research*, *42*(1), 65–74.

Song, H., Witt, S. F., & Zhang, X. (2008). Developing a web-based tourism demand forecasting system. *Tourism Economics*, *14*(3), 445–468.

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

Song, H., Wong, K. K. F., & Chon, K. K. S. (2003). Modelling and forecasting the demand for Hong Kong tourism. *International Journal of Hospitality Management*, *22*(4), 435–451.

Syntetos, A. A., Nikolopoulos, K., Boylan, J. E., Fildes, R., & Goodwin, P. (2009). The effects of integrating management judgement into intermittent demand forecasts. *International Journal of Production Economics*, *118*(1), 72–81.

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, *19*(4), 715–725.

Turner, D. S. (1990). The role of judgement in macroeconomic forecasting. *Journal of Forecasting*, *9*(4), 315–345.

Vokurka, R. J., Flores, B. E., & Pearce, S. L. (1996). Automatic feature identification and graphical support in rule-based forecasting: a comparison. *International Journal of Forecasting*, *12*(4), 495–512.

Witt, S. F., Song, H., & Wanhill, S. (2004). Forecasting tourism-generated employment: the case of Denmark. *Tourism Economics*, *10*(2), 167–176.

Wong, K. K., Song, H., Witt, S. F., & Wu, D. C. (2007). Tourism forecasting: to combine or not to combine? *Tourism Management*, *28*(4), 1068–1078.

Haiyan Song is Chair Professor of Tourism in the School of Hotel and Tourism Management, The Hong Kong Polytechnic University. His main research areas include tourism impact assessment, tourism demand forecasting, tourism supply chain management and forecasting. He has published widely in such journals as the *Journal of Applied Econometrics*, *International Journal of Forecasting* and *Journal of Forecasting*.

Bastian Z. Gao is a Ph.D. student in the School of Hotel and Tourism Management, The Hong Kong Polytechnic University. His Ph.D. thesis is studying the web-based tourism demand forecasting system.

Vera S. Lin is a Ph.D. student in the School of Hotel and Tourism Management, The Hong Kong Polytechnic University. Her Ph.D. study focuses on the combination of statistical and judgmental forecasts in the context of tourism.