Global Journal of Management and Business Research: F
Real Estate, Event & Tourism Management
Volume 16 Issue 1 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4588 & Print ISSN: 0975-5853

Business Confidence and International Tourism Demand: Evidence from a Global Panel of Experts

By Valeria Croce

Introduction- In 2003, the United Nations agency for tourism (UNWTO), established a Panel of Tourism Experts, to collect regular information on the short-term development of tourism. Experts' opinions are since used to estimate a confidence index, which offers fairly accurate information on the current and future development of the tourism sector worldwide and by macro-regions. The significance of this instrument became evident during the 2008/2009 economic and financial crisis, when indications about the impact and duration of the crisis were scarce, but particularly relevant to a sector having experience virtually uninterrupted growth until then. This piece of research intends to achieve a better understanding of confidence index's contribution to forecasting tourism demand. Results confirm the tourism confidence index as an effective method to improve the accuracy of forecasts.

GJMBR-F Classification: JEL Code: L99

Strictly as per the compliance and regulations of:

Business Confidence and International Tourism Demand: Evidence from a Global Panel of Experts

Valeria Croce

I. Introduction

In 2003, the United Nations agency for tourism (UNWTO), established a Panel of Tourism Experts, to collect regular information on the short-term development of tourism. Experts' opinions are since used to estimate a confidence index, which offers fairly accurate information on the current and future development of the tourism sector worldwide and by macro-regions. The significance of this instrument became evident during the 2008/2009 economic and financial crisis, when indications about the impact and duration of the crisis were scarce, but particularly relevant to a sector having experienced virtually uninterrupted growth until then. This piece of research intends to achieve a better understanding of confidence index's contribution to forecasting tourism demand. Results confirm the tourism confidence index as an effective method to improve the accuracy of forecasts.

This piece of research intends to achieve a better understanding of confidence index's contribution to forecasting tourism demand. Following the approach proposed by Guizzardi and Stacchini (2015), this study assesses the predictive power of the UNWTO Tourism Confidence Index-possibly the world's most widely used and influential forecasts for the tourism sector- by factoring the index in structural time series models. Forecasts are produced for the global scale, for advanced and emerging economies and five geographic macro-regions. Models' performance is evaluated on in- and out-of-sample basis, and benchmarked against frequently used univariate models. The index is first considered as stand-alone forecasting tool, with analyses focusing on its correlation with series of actual values and assessing the accurateness of forecasts derived from the Index. A second part of the analysis focuses on the Index value as explanatory variable in model-based forecasts, testing its usefulness as predictor by factoring the index in structural time-series models. The predictive accuracy of augmented models is eventually tested against simpler versions of structural models and auto-regressive models. Results show that information gathered through a simple and rather inexpensive tool can be effective in improving the accurateness of short-term forecasts of tourism demand.

Managerial implications of research findings are manifold. As demonstrated by Guizzardi and Stacchini (2015), tourism confidence indexes offer the possibility to obtain timely estimates of current and near-future levels of tourism demand. They hence represent a cost-effective solution to compensate for the lag in official statistics publication. Previous research also proved that confidence indexes in general, and the UNWTO index in particular, have a good predictive power in identifying discrete turning points in the business cycle (Taylor and McNabb 2007, Croce 2016), and can actively contribute to strengthen the resilience of a sector vulnerable to changes in the external environment. The widespread use of the Internet and ICT developments further offer unprecedented opportunities to leverage collective intelligence from large groups of individuals for forecasting purposes (Segaran 2007), helping small businesses and budget-constrained destinations to embrace forecasting practices, replicating processes which proved fruitful in other sectors (Wolfers and Zitzewitz 2004).

Furthermore, the data environment related to the tourism sector is characterised by lengthy statistical processes causing a dominating backward-looking approach (Vanhove, 2005). The suboptimal availability of data, both in terms of scope and timeliness, coupled with a complex demand and supply, make a case for the use of confidence indexes in tourism, as they represent a cost effective methodology to forecast future changes in the development of the sector.

II. Literature Review

The view that confidence measures can predict fluctuations in economic series, such as the level of economic activity or consumer spending, is a popular one in the economic literature. In the tourism-related literature, instead, the use of confidence measures in forecasting is still a largely unexplored area (Swarbrooke and Horner 2001, Njegovan 2005, Yap and Allen 2011, Mihalic, Kester et al. 2013, Guizzardi and Stacchini 2015, Croce 2016).

The existing literature on tourism forecasting is by far dominated by quantitative approaches, and only a

Author: e-mail: valeria.croce@gmail.com

Global Journal of Management and Business Research (F) Volume XVI Issue I Version I Year 2016

limited number of studies emphasize the suitability of qualitative approaches to forecast tourism demand (for reviews, see Witt and Witt 1995, Goh 2004, Song and Li 2008, Goh and Law 2011, Peng, Song et al. 2014). In general, the use of subjective information in forecasting -inter alia sentiment measures-is justified whenever quantitative indicators are missing, or their provision is not sufficiently timely. Sentiment indexes are frequently considered as a valid source only when they compensate for the lack of "hard" information. In reality, these measures feature a number of properties that make them useful indicators even in presence of quantitative information, such as the capability to factor non-measurable variables in estimates, and to provide a synthetic value of the impact these factors would have on tourism development, and to provide near real-time forecasts (Guizzardi and Stacchini 2015).

The vast literature about forecasting provides clear directions about environment-specific conditions determining which approach is the most suitable for a forecasting task. Wherever routine decisions are involved, the assumption of continuity is realistic and sufficient quantitative information is available, the most complex quantitative methods provide more accurate predictions than the simpler statistical methods (Makridakis, Wheelwright et al. 1998, Armstrong and Fildes 2006). In areas such as physics and engineering, for instance, the accuracy of causal and statistical forecasting models has achieved remarkable results in providing almost error free predictions (Makridakis and Taleb 2009). Whenever the conditions of the above-mentioned requirements are non-optimal, a combination of statistical methods and judgement proved to produce a more accurate forecast than each approach used separately (Goodwin 2002, Fildes, Goodwin et al. 2006, Fildes, Goodwin et al. 2009).

The kaleidoscopic and human-based nature of tourism poses serious threats to regular data collection and to the correct specifications of exogenous variables required by econometric models. Confidence indicators may offer a way out of this trade-off, as they incorporate in a single value the impact of many predictors, including 'those economic and market phenomena that are known but not quantified" (Caniato, Kalchschmidt et al. 2011).

The use of confidence indicators in forecasting is very popular in the general economic literature, but it is still a rather unexplored research area in tourism. Swarbrooke and Horner (2001) first introduced this approach in the tourism literature, using business travellers' expectations about a foreign country's economic development to forecast business travel flows. Njegovan (2005) used a probit model to examine whether leading indicators could be used for the purpose of predicting short-term shifts in demand for business travel by air, to and from the UK. The estimated probit model provided timely predictions of industry recessions and it was overall more accurate than benchmark models. Allen and Yap (2011) examined whether a consumer confidence index could help predicting Australian domestic tourism demand. Their findings suggest that the index has significant impacts in forecasting VFR, but not other forms of tourism. Guizzardi and Stacchini (2015) provide evidence that information retrieved from business surveys is effective in improving real-time forecasting of hotel arrival to regional level. Croce (2016) found that prospects of UNWTO Panel of Tourism Experts can provide meaningful indications about the future sign of international tourism demand growth, and that they can significantly contribute to improve forecasting accuracy under specific circumstances.

This study intends to escalate the approach proposed by Guizzardi and Stacchini (2015) to the global and macro-regional level, using data provided by the UN agency for tourism, UNWTO. Following their design, Tourism Confidence Index prospects are factored in structural time series models and compared with forecasts produced by the same models, without the index. Forecasts are produced for the total number of international tourist arrivals, as well as for international arrivals to advanced and emerging economies and to five geographic macro-regions (see Figure 1). Models' performance is evaluated on in- and out-of-sample basis, and benchmarked against a baseline ARIMA model.

The choice of quantitative models is grounded in evidence Peng, Song et al. (2014) provide in their latest attempt to generalise models' performance in forecasting tourism demand. Based on empirical findings of by-then published studies, their review confirms ARIMA processes as the most frequently used time-series approach, but also highlights the advantages of structural time series models, when exogenous variables can be included (Gonzalez and Moral 1995, Kulendran and King 1997, Kulendran and Witt 2001, Turner and Witt 2001). Structural time series models are based on the traditional decomposition of time series into trend, seasonal and cycle components (Harvey 1990). This characteristic is expected to well capture evolutions of tourism demand, especially in those macro-regions characterised by regular variations in their growth patterns. The strong correlation of index values with international tourist arrivals, and the good fit of regression models including the confidence index as predictor (Croce 2016), are also convincing about the positive contribution of the index in improving forecasts.

III. Data Characteristics

In this piece of research, UNWTO Tourism Confidence Index is used as input in regressive forecast models. International tourist arrivals (ITA) are selected as dependent variables, and reflect the cross-border flow of people for tourism purposes. The following section describes the data used and illustrates those properties of time series, which are a prerequisite for the selection of appropriate forecasting models.

a) International tourism development from 2003 to 2015, key indicators

Since the early 1990s, UNWTO secretariat has been collecting series of international tourist arrivals from destinations around the world, which are regularly published on the organisation's World Tourism Barometer. Despite methodological inconsistencies that characterize the production of tourism statistics, the information provided by the UN agency can be deemed as the most reliable dataset covering virtually all world destinations, not last due to a compensation effect that characterizes large aggregates, as the eight series (see Figure 1) here in exam.

Monthly growth rates in international tourist arrivals, compared to the same month in previous year, have been calculated and aggregated in 4-month periods, in order to create time series consistent with the period covered by confidence indexes. Each series consists of 38 observations, dating from 2003/2 to 2015/3 (see Figure 1). Of these, 26 observations have been used to estimate forecasting models' parameters for the first iteration, and the remaining 12 observations are devoted to out-of-sample accuracy tests (shadowed area in Figure 1).

During the period in exam, international tourism has enjoyed growth in all world regions. Overall, international tourist arrivals grew on average by 4,5% each year, with typically regular variations around the long-term trend (Standard Deviation(SD) = 3.87). Large variations were caused by external shocks of relevant magnitude, such as the breakout of the SARS syndrome in 2003, and the 2008-2009 economic and financial crisis. In this period, emerging economies typically grew faster (Means(M) = 5.61; SD = 5.02) than advanced ones (M = 3.64; SD = 3.64), but these latter followed a more regular growth pattern, as indicated by a lower standard deviation. Mature destinations such as Europe (M = 3.37; SD = 3.46) and the Americas (M = 4.22; SD = 4.64) are characterised by a comparatively lower growth, largely explained by the already large volumes of international tourists attracted by these regions, and partly also by the distance from the most rapidly expanding outbound travel markets, as China. Large volumes are also an indication that these regions host many established destinations, which is reflected in regular growth patterns. Tourism demand for emerging destinations in Asia and the Pacific (M = 7.24; SD = 10.00), the Middle East (M = 6.73; SD = 12.51) and Africa (M = 5.11; SD = 5.12) grew at a faster pace, but also proved more volatile and vulnerable to external shocks. For all series bar the Middle East, international tourism demand seems to grow more regularly during the out-of-sample period, although a pairwise test of means difference reveals no significant difference between the two sub-samples.

An augmented Dickey-Fuller test, performed on the in-sample subset of data, depicts all series as non-stationary processes, with a confidence level of 99%[1]. First-order and second-order differencing returns stationary series with a type-I error probability lower than 1%. The aggregation in 4-month periods removes seasonality from arrival series, but a few series (Americas, Asia Pacific and Europe) seem to feature a cyclical component. Growth rates pattern in these macro-regions could be fit by a sinusoidal curve with cycles lasting approximately 3 years for Americas and Asia Pacific, and 6 to 7 years for Europe. Given the heavy weight destinations in Asia Pacific have in the emerging economies aggregate, this series seems also to be characterised by a cyclical component with cycles of 3 years.

[1] The ADF test depicts the series Americas, Asia Pacific and Middle East as stationary at a confidence level of 95%.

Figure 1: International tourist arrivals and Tourism Confidence Index prospects, 4-month data series, 2003/1 to 2015/3 (source: UNWTO).

b) UNWTO Tourism Confidence Index

Since 2003, UNWTO Secretariat has been collecting retrospective and prospective evaluations of tourism performance from hundreds of tourism stakeholders recruited through the organisation's global network[2]. Evaluations and prospects capture the development of tourism demand compared to what experts could reasonably expect for the period of

[2] Panel members are conveniently recruited by the UNWTO Secretariat based on their professional background, without any formal assessment of the candidate's expertise with the forecasting task.

reference. Changes are measured on a 5-step Likert (1932) interval scale, and collected by the means of an e-mail questionnaire. Experts are asked to motivate their assessment mentioning the main determinants that underlie the expected evolution of tourism demand. Waves are conducted three times a year and consist of a single round survey. Once a year, evaluations and prospects for the full year are also collected.

The Tourism Confidence Index (TCI) is constructed using the un-weighted net balance of survey responses collected through this survey, with 100 being the threshold value between contraction and expansion (see UNWTO 2016). In order to combine intra-year prospects and international tourist arrivals, index values have been homologised by the means of a linear transformation as in (1):

$$\bar{P}_t = \alpha + \beta P_t \tag{1}$$

where P is the index value at time t. Intercept α and slope β are estimated through a linear regression between the index and the equivalent series of international tourist arrivals to return estimated growth values ($\bar{P}$) at each time, hereafter referred to as homologised TCI values. Series of homologised TCI values covering 4-month periods are used in the analysis that follows.

Since the second quarter of 2003, when the survey started, an average of 320 experts submit prospects and/or evaluations each wave. Such a large basis of participating experts allows breaking down the index into 18 different subsets, according to the level of economic development, the macro-region or the sector associated with the expert (see Table 1). For most subsets of data, the average number of experts participating in each round exceeds the threshold of 20 experts identified by Rowe and Wright (2001) as the optimal maximum number for Delphi surveys (the italic font in Table 1 denotes those subsets which don't meet this criterion). All subsets are also characterised by a relatively high deviation from the average index value, a positive factor in expert forecasting, where forecast accuracy tends to increase with diversity of inputs (Gordon and Helmer 1964, Woundenberg 1991, Williams and Webb 1994, Gupta and Clarke 1996, Rowe and Wright 1999).The average of all prospects collected throughout the past 13 years is reported as indicator of the level of optimism (or pessimism) that distinguishes each subset.

Table 1: Prospects of the UNWTO Panel of Tourism Experts (average value and standard deviation), average number of participants per round, and relationship with the correspondent series of international tourist arrivals (R^2 and Granger test).

	Prospects (4 month)		Experts / round	Diagnostics⁺			
	Average	StDev		R^2	Granger test (F)	DF	MAE
Total	120	42	282	0,59	14,897***	-2	1,65
State of economic development							
Advanced	113	39	153	0,51	8,974**	-2	1,77
Emerging	128	42	129	0,64	3,312	-2	1,95
Macro-region							
Africa	129	43	20	0,27	0,846	-1	3,47
Americas	125	42	73	0,59	7,208**	-2	2,19
Asia and the Pacific	126	44	38	0,32	4,425*	-2	4,53
Europe	115	38	122	0,51	6,925**	-2	1,83
Middle East	125	42	14	0,26	0,553	-1	8,88
Type of sector							
Public	125	40	119	0,53	5,188*	-2	1,70
Private	117	42	163	0,62	16,580***	-2	1,75
Sector							
Consultancy, Research & Media	118	40	88	0,46	15,726***	-2	1,86
Destinations (national)	126	40	75	0,64	2,254	-2	1,58
General Industry Bodies & Other	119	40	43	0,46	6,388**	-2	1,83
Accommodation & Catering	115	42	29	0,39	5,722*	-1	2,25
Tour Operators & Travel Agencies	117	43	29	0,48	1,094	-2	1,87
Destinations (regional & local)	121	37	23	0,39	7,553**	-2	2,25
Global Operators	109	39	15	0,35	7,166**	-2	2,13
Transport	121	45	12	0,28	1,635	-2	2,20
Business & meetings (MICE)	123	40	7	0,39	4,236*	-2	2,05

⁺: Diagnostics are related to the total number of ITA for sector subsets, and the correspondent series of ITA for all other subsets
: Significant at p<0.05; *: significant at p<0.001*

When considered as independent forecasting tool, prospects of the UNWTO Tourism Confidence Index show a good fit with correspondent series of international tourist arrivals, as synthesised by R^2 values in Table 1. Differences among subsets are largely explained by panel size, as the average number of experts participating in the survey varies greatly across series and is strongly, and positively, correlated with R^2

values (ρ = 0.72). Expectedly, R^2 values are also inversely related to the accurateness of forecasts based on homologised prospects (ρ = -0.60). Series with an R^2 of at least 0.50 are associated with a Mean Absolute Error (MAE) of 1.8 percentage points on average, which goes up to 3.0 percentage points for series with a lower R^2 value. Expectedly, the index fits regular series better than volatile ones.

This study argues that prospects can also be a valuable predictor of future tourism demand development, and they can significantly improve the accurateness of univariate forecasting models. On average, there is a positive elasticity between the index and international tourist arrivals growth. Based on (2), one percentage point change in arrivals growth corresponds to a change between 4 and 6 points in the index value. These values remain rather consistent for periods of positive and negative growth, and across geographical aggregates.

$$E_{ITA,t} = (I_t - 100) / \left[(ITA_t - ITA_{t-1}) / ITA_t \right] \times 100 \quad (2)$$

A Granger (1969) causality test has also been performed on series of prospects and correspondent series of arrivals, to test the null hypothesis that prospects are exogenous with respect to international tourist arrivals. If the hypothesis is rejected, it is likely that including past values of prospects in the forecasting model would provide statistically more information than past values of arrivals alone. Granger test values are also used to select those sub-sets of prospects (by the sector experts work for), which can be useful in predicting values of international tourist arrivals at total level. Only those series with a significant level of Granger-causality will be included in the analyses that follow. The test has been performed using the 'granger test' function available in 'lmtest' package R, the free software environment for statistical computing and graphics. This function compares the unrestricted model - in which the dependent variable (y) is explained by the lags (up to the specified order) of y and of the independent variable (x) -, and the restricted model - in which y is only explained by the lags of y - by the means of a simple Wald test [3]. Results are reported in Table 1.

In general, the test confirms prospects as a predictor that can significantly improve international tourist arrivals forecasts (the test is significant for 13 out of 19 series). The null hypothesis that the lagged series (lag = 2) of total prospects, as well as prospects provided by the Private Sector and Consultancy and Media experts, would not be useful in predicting total international tourist arrivals' growth rates, can be rejected with a Type-I error probability of 1%. The same hypothesis can be rejected for the lagged version (lag = 2) of prospects provided by General Industry Bodies & Other, Local and Regional Destinations and Global Operators, although with an error probability of 5%. Looking at aggregates by level of economic development and macro-regions, lagged series of prospects (lag = 2) can be considered Granger-causing the correspondent series of international tourist arrivals in Advanced economies, Americas and Europe, with a 95% confidence level.

It is appropriate to mention, that aggregates by level of economic development and regional series are marked by a higher conceptual correspondence between series of actual values and series of prospects, than aggregates by sector. While international tourist arrivals measure inbound flows to the country of destination, the survey asks experts to estimate future developments in the tourism sector in their region and/or the business sector they operate in. Prospects hence return estimates, which are not only conceptually broader than just the international demand component, but which may also be strongly biased by local patterns, when provided by experts operating in local businesses. Unsurprisingly, the stronger Granger-causality is observed in groups of experts operating in global businesses. As such a discrepancy is lower for aggregates by economic development and macro-region, all seven aggregates will be included in the forecasting exercise described below (bold font in Table 1).

IV. Forecasting Models

The assessment of UNWTO Tourism Confidence Index contribution to improve forecast accuracy is based on a comparison of forecasts obtained with structural time series models, in which the index is factored, and structural models without the index as predictor. Forecasts based on simple auto-regressive models are also computed, as baseline forecasts, as these models frequently outperformed more complex models in previous forecasting competitions (Li, Song et al. 2005, Athanasopoulos, Hyndman et al. 2011, Peng, Song et al. 2014, Akın 2015, Guizzardi and Stacchini 2015).

For each of the eight series in exam, forecasts have been obtained using packages in R. The 12 forecasts in the out-of-sample dataset allow for a robust test of significance in forecast accuracy between baseline and competing models.

a) *Baseline forecasts: auto-regressive moving-average models*

For each series, different ARIMA models have been estimated in order to select those who perform best in estimating one-step-ahead forecasts. The choice

[3] It must be noted, that the original definition of Granger causality does not account for latent confounding effects and does not capture non-linear causal relationships.

of models was updated each time a new observation was added, as in a real-life situation, where data arrive sequentially and in-sample dataset is updated accordingly, as new data become available. This process was routinised by the use of the 'auto.arima' function available in R "forecast" package. The function uses a variation of the Hyndman and Khandakar algorithm, which performs repeated KPSS tests to set the order of differencing (d), a stepwise search based on AICc values to set the number of auto-regressive terms (p) and lagged forecast errors (q), and maximum likelihood estimation (MLE) to identify the best fitting ARIMA model [4]. Model parameters and diagnostics are reported in Table 2.

In general, ARIMA forecasts show a low out-of-sample fit (R^2) with series of actual values, which is reflected in mean absolute error (MAE) values exceeding one percentage point (p.p.) for virtually all series. The only exception is the moving average model of order one, used to forecast international tourist arrivals to Africa, with an R^2 value of 0.77. This is reflected in comparatively accurate forecasts (MAE = 1.86), given the highly varying pattern that characterises this series. As predictable, ARIMA models perform best with regular series. The auto regressive model of order two, applied to total growth in international arrivals, stands out as the best performing model, with an average error and standard deviation both below the one percentage point threshold.

Table 2: ARIMA models, parameters estimates and accuracy measures by macro-regions.

ARIMA - one-step-ahead forecasts

	World	Africa	Americas	Asia Pacific	Europe	Middle East	Advanced ec.	Emerging ec.
Model°	(2,0,0)	(0,0,1)	(1,0,0)	(0,0,2)	(1,0,0)	(0,0,0)	(2,0,0)	(0,0,2)
ar1	1,07	-	0,71	-	0,82		1,13	-
ar2	-0,53	-	-	-	-		-0,50	-
ma1	-	0,39	-	0,88	-		-	0,97
ma2	-	-	-	0,64	-		-	0,54
intercept	4,20	6,79	-	-	-	8,89	2,82	5,98
Out-of-sample								
R^2	0,03	0,77	0,32	0,12	0,04	0,07	0,00	0,15
MAE	0,60	1,86	1,80	1,16	1,05	9,70	1,58	1,47
stdev MAE	0,58	1,70	1,50	1,18	0,78	7,04	0,83	1,02
MAPE	13%	121%	34%	19%	27%	475%	33%	35%

°: all values refer to the ARIMA model fitting the first 26 observations (2003/1 to 2011/3).

b) Structural time series models

Structural time series models are a flexible approach for time series analysis. They can be considered as state-space models for time series, based on a decomposition of the series into their unobserved, latent components, namely trend, cycle and seasonality. These models are frequently used not only to provide a description of the salient features of time series, but also to forecast their future values (Holden, Peel et al. 1990). The characteristic of modelling time series components well suits to forecast the series in exam, as the cyclical component, observed in some of them does not appear explicitly in the definition and selection of ARIMA models.

For each of the eight series in exam, forecasts have been estimated for a one-year horizon (three periods). As for the baseline forecasts, the process has been replicated for all 12 periods of the out-of-sample data. Structural time series models were estimated using R 'stsm', 'dlm' and 'KFKSDS' packages. The 'stsm' package offers the opportunity to apply five different models. Different models have been tested on each of the series, and the best performing model, in terms of in-sample MAE, has been retained for comparison with competing models.

The decomposition of each time series into its components generates small improvements in forecasts accuracy for five of the eight series in exam (Table 3). Accuracy gains are measured as percentage increments compared to corresponding baseline forecasts[5]. The magnitude of accuracy gains is rather limited, as it falls between a minimum of 15% and a maximum of 29% of the respective ARIMA Mean absolute error. Considerable accuracy gains are observed for series with large movements around the trend, such as Asia Pacific (accuracy gains of 24 p.p. compared to the MAE of baseline forecasts) and Middle East (gains of 15 p.p.), but also for series with stable growth patterns, such as the series advanced economies (29 p.p.) and total (17 p.p.). A common

[4] For more information see: https://www.otexts.org/fpp/8/7.

[5] The difference is calculated between the MAE of STS models and the MAE of ARIMA models, relative to this latter measure. A negative value indicates a lower MAE value for STS models.

characteristic of these series is the regularity of their behaviour, being they regular in their cyclical growth pattern[6], or being they near-stationary series.

Table 3: Structural time series models, parameters estimates and accuracy measures by macro-regions.

	World	Africa	Americas	Asia Pacific	Europe	Middle East	Advanced ec.	Emerging ec.
Model	Llm+seas	Trend	BSM	Level	Llm+seas	Level	Llm+seas	BSM
Parameters coeff.°								
μ	13,15	2,92	21,39	125,67	9,04	9,52	8,91	23,82
ε	0,00	15,19	0,00	0,00	0,00	144,16	0,00	0,00
δ	-	-	0,00	-	-	-	-	
ψ	-	-	-	-	-	-	-	0,00
ϕ	-	-	0,00	-	-	-	-	0,00
out-of-sample								
R^2	0,14	0,29	0,26	0,16	0,11	0,12	0,03	0,27
MAE	0,50	2,83	2,09	0,88	1,17	8,26	1,12	1,24
stdev MAE	0,25	1,62	1,00	0,86	0,82	6,10	0,64	0,84
MAPE	11%	183%	39%	14%	30%	405%	23%	30%
Gains % ARIMA	-17%	52%	17%	-24%	11%	-15%	-29%	-16%

°: all values refer to the first iteration of the model with 26 observations.
: Significant at p<0.05; *: significant at p<0.01*

Structural time series models tend to reduce the variance of absolute errors; hence they contribute to make predictions more robust. This improvement is particularly remarkable for the series Middle East, Americas, World and Asia Pacific. This is reflected in the improved fit between forecasts and actual values, as expressed by R^2 values. Compared to baseline forecasts, structural models return smoother cyclical fluctuations and a positive trend, which better fits the actual development of tourism in these regions. Minor improvements in R^2 can be indeed observed for all series, with the exception of Africa and Americas. A closer comparison of these two approaches reveals that structural time series models are more effective in capturing turning points, while moving average and auto-regressive models more nicely match the pace of growth before and after these points (see Figure 2).

c) *Augmented structural time series models*

Homologised values of Tourism Confidence Index prospects (see section 3.b) were eventually factored in the structural models used in the previous analysis. Both original series of homologised prospects and lagged versions thereof have been added to the models, for a comparison of the best performing models. Seasonally differenced series (lag = 3) of homologised prospects stand out as the regressors, which provide the most valuable inputs to compound forecasts across all series. The only exception is the series emerging economies, for which original series of homologised prospects have been used. Parameter estimates and accuracy measures are reported in Table 4.

[6] Based on the longer in-sample dataset, tourist arrivals growth to the Middle East are not characterised by a cyclical component, but the 12 observations in the out-of-sample data a synusoydal type of pattern.

Table 4: Augmented structural time series models, parameters estimates and accuracy measures by macro-regions.

	World	Africa	Americas	Asia Pacific	Europe	Middle East	Advanced ec.	Emerging ec.
Model	Llm + seas[1]	Trend[1]	BSM[2]	Level[2]	Llm + seas[1]	Level[2]	Llm + seas[1]	BSM[1]
Parameters°								
σ^2_ε	0,14	15,09	2,62	1,79	1,83	105.14	0,29	0,38
σ^2_ξ	2,95	-	8,46	11,31	3,85	10,94	3,18	5,20
σ^2_ω	-	0,01	-	-	-	-	-	0,03
x reg	0,70	0,47	0,25	0,20	0,52	0,29	0,69	0,56
out-of-sample								
R^2	0,01	0,77	0,50	0,23	0,11	0,46	0,02	0,03
MAE	0,46	2,03*	1,38*	0,71	0,77	6,20*	1,13	2,01
stdev MAE	0,36	1,33	1,14	0,71	0,00	4,96	0,56	1,38
MAPE	0,03	0,14	0,10	0,05	0,05	0,44	0,08	0,14
Gains % ARIMA	-23%	9%	-23%	-39%	-27%	-36%	-29%	37%
Gains % STS	-7%	-28%	-34%	-19%	-35%	-25%	1%	62%

°: all values refer to the first iteration of the model with 23 observations (strating 2004, 2).
1: first order 1; 2: second-order differencing.
**: DM test significant at $p<0.05$; **: significant at $p<0.01$.*

Accuracy gains brought by the use of the Index, if any, are first assessed in terms of percentage increments compared to the corresponding non-augmented version and baseline forecasts. The augmented version of structural time series models returns more accurate forecasts than the corresponding benchmarks for all series, with the exception of the two aggregates by stage of economic development. For advanced economies, the use of the index doesn't bring any substantial change in forecast accuracy, although the Granger test (performed on the subset of data) pointed to a possible contribution. This is likely explained by the uncertainty that lingered among tourism experts during the period 2012/2 until 2014/2, especially those in Europe – where most advanced economies are- Crises of various natures, primarily the European debt crisis, the risk of Grexit and the Crimea crisis, dragged experts' prospects down. In fact, tourism demand proved to be more resilient than experts expected and negative factors were more than offset by the rebound of important outbound travel markets, aggressive price policies as well as a favorable economic environment and plummeting oil prices. When applied to emerging economies, the Index instead slightly deteriorates the quality of forecasts (0.77 p.p. less accurate). This can be explained by the diversity in tourism development that can be observed within emerging economies worldwide.

For all other aggregates, accuracy gains brought by the Index range between negligible values (e.g. for the series total the MAE is only 7%, or 0.03 p.p., below the non-augmented version) to a maximum of one-third of the benchmark forecast error (or -2.1 p.p. for the Middle East, and -0.7 p.p. for Americas). Data plotted in Figure 1 provide explanation for these differences. In general, information contained in the index helps improving turning points and smoothening the amplitude of cyclical variations. While the non-augmented version models, more or less efficiently, regular variations in time series, the index introduces those changes that enhance the match with observed values. This confirms Guizzardi and Stacchini's (2015: 219) conclusions that a confidence index can be usefully exploited to explain deviations from trend-cycle, due to short-term shocks. This is symptomatic of experts' capability to infer behaviour from the past, but also factor current inputs in their estimates, and deliver a realistic version of future tourism developments. The use of the Index further reduces deviations from mean errors over the 12 out -of-sample observations.

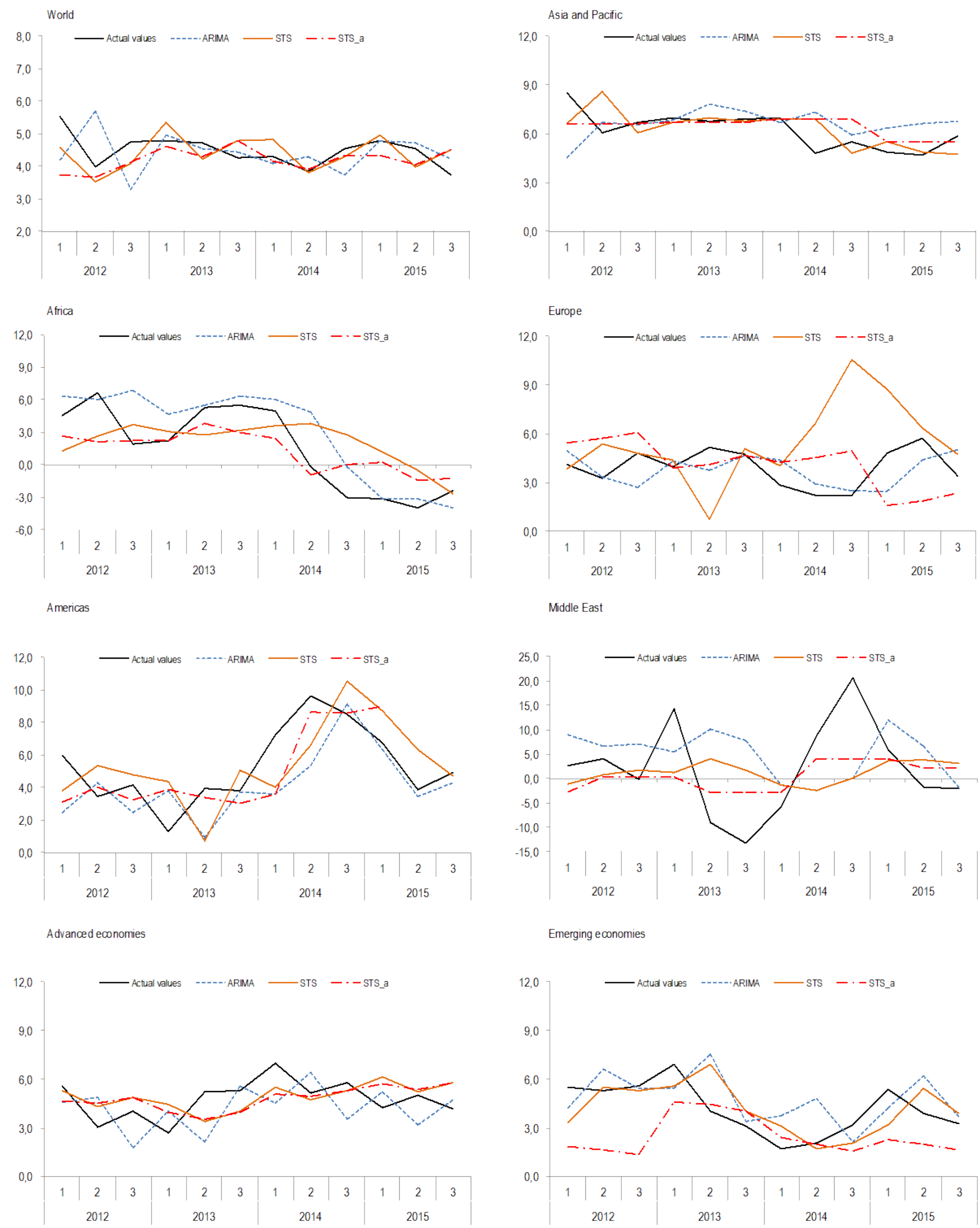

Figure 2: International tourist arrivals by region or economic development, actual values and forecasts, 4-month data series, 2012/1 to 2015/3 (source: UNWTO).

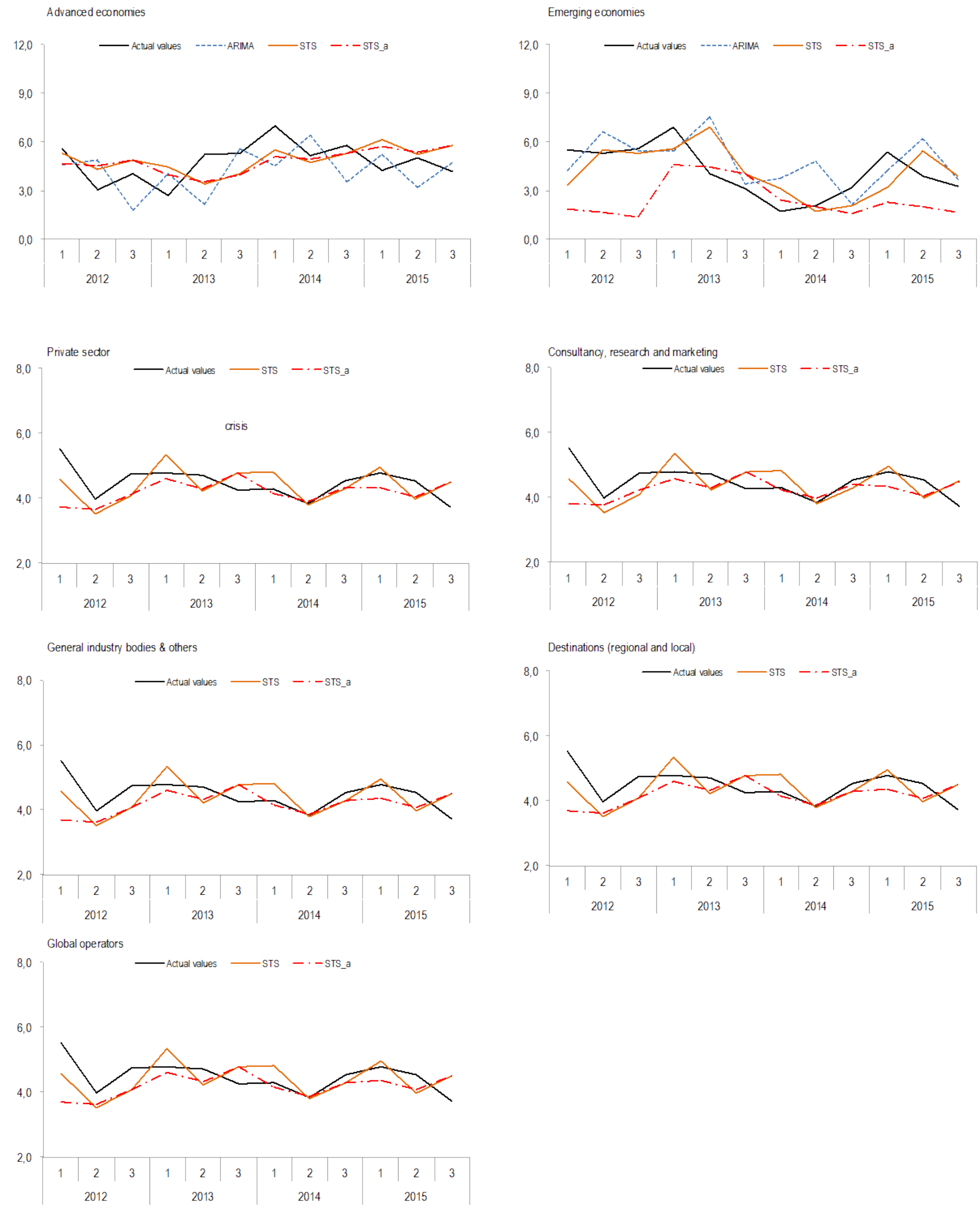

Figure 3: International tourist arrivals by sector, actual values and forecasts, 4-month data series, 2012/1 to 2015/3 (source: UNWTO).

The large size of UNWTO Expert group allows the analysis of different subsets of the index by sector. Prospects, provided by experts from five different tourism-related sectors, have been used as predictors of growth in international tourist arrivals worldwide. Parameter estimates and accuracy measures are reported in Table 5 and forecasts plotted in Figure 3. The analysis doesn't point to professional experience in a specific sector as relevant to improve forecasts for the overall tourism sector.

Global Journal of Management and Business Research (F) Volume XVI Issue I Version I Year 2016

Table 5: Augmented structural time series models, parameters estimates and accuracy measures by sector.

	Private sector	Consultancy, R+M	General ind.	Destinations (r+l)	Global operators
Model	Llm + seas[1]	Llm + seas[1]	Llm + seas[1]	Llm + seas[1]	Llm + seas[1]
Parameters*					
σ^2_ε	0,16	0,72	-	-	-
σ^2_ξ	3,30	3,23	4,65	5,88	5,30
σ^2_ω	-	-	-	-	-
x reg	0,72	0,67	0,49	0,46	0,62
out-of-sample					
R^2	0,01	0,02	0,02	0,02	0,02
MAE	0,50	0,47	0,50	0,50	0,50
stdev MAE	0,46	0,44	0,47	0,47	0,47
MAPE	11%	11%	11%	11%	11%

[0]: all values refer to the first iteration of the model with 23 observations (strating 2004, 2).
1: order 1 differencing; 2: order 2 differencing.

V. Comparing Predictive Accuracy

Accuracy gains are eventually tested on out-of-sample values using the Diebold-Mariano (1995)[7] statistics computed on out-of-sample absolute errors. The use of this test is a standard practice to compare the predictive accuracy of independent forecasts, and is particularly suited to compare the accuracy of model-free forecasts, as for instance survey-based forecasts. Furthermore, the Diebold-Mariano test accommodates for a number of series characteristics, among which the presence of serially correlated forecast errors (see Diebold & Mariano, 1995: 10). When used with short forecast horizons, as it is the case here, forecast errors correlation can lead to particularly conservative results of the DM test, with the null hypothesis being rejected too often. This may explain the limited number of statistically significant results, and encourages the interpretation of significant errors as solid recommend-dations about the validity of the correspondent forecasting approach.

The predictive accuracy of augmented models is reported in Table 4. The test suggests that the use of the index returns significantly more accurate forecasts for three series, namely Africa, Americas and Middle East. This is a noteworthy result from an operational viewpoint, as forecasts for series with very high variations, such as Africa and Middle East, are frequently exposed to large errors. Differences in forecast accuracy between the simpler version of structural models and ARIMA models appear not to be significant, based on the Diebold-Mariano test. STS models for the series Middle East and advanced economies, augmented with the index, become statistically more accurate than their ARIMA counterpart, with a type-I error probability of 5%.

For all other series, results are encouraging, but not significant from a statistical viewpoint. In particular, the use of the index from different sectors as predictor in STS methods seems to reduce the predictive power of models. This is explained by the conceptual discrepancy between the prospects and actual series. While international tourist arrivals measure inbound travel flows, UNWTO prospects capture changes in the overall tourism sector, including domestic demand. The related index returns a measurement of the overall business climate rather than just its demand component, which may partly explain the lack of significant results for some of the series.

Another caveat, strictly intertwined with this aspect, is that international tourism demand is typically proxied by visitor arrivals, or overnights, at a destination. In spite of a remarkable progress, statistics on international flows are still largely not comparable as based on a host of different methodologies, hence falling short of the rigorous international comparability that would be required. Depending on the methodology, international tourist arrivals may be close to the official definition of a visitor, namely a person, who travels to a place outside her usual environment for personal or business/professional purposes, or may just monitor a subset of it, as for instance visitors staying at hotels and similar establishments. This lack of homogeneity is also believed to introduce a bias in results.

Another caveat stems from the use of a nominal group technique approach to collect prospects. This approach, appreciable for its simplicity, overlooks some of the techniques introduced to limit bias. Among them, it is worth mentioning the lack of multiple rounds, neglecting experts the opportunity to retrace their evaluation process based on group feedback, and the lack of feedback on their performance, which further limits their chances to learn from past experiences.

[7] The DM test has been chosen as measure of significance due to the non-zero mean and serially correlated nature of forecast error series. Empirical applications of the test suggest that on small samples the test can have the wrong size and reject the null hypothesis too often. For this purpose, confidence levels start at 0.1.

VI. Conclusions

For the varied nature of its demand, and the composite nature of supply, the analysis of tourism requires a complex system of indicators (Candela and Figini 2012), for which structured and comparable data are seldom available. Monitoring such a complex system of determinants is difficult and expensive (Vanhove 2011), and it is often not a priority of the many governments, who don't consider tourism a strategic asset for their economies. The tourism sector is therefore one of those domains, where the use of qualitative forecasts is most promising, (Vanhove 2011).

The general forecasting literature provides evidence that qualitative forecasting methods are particularly valuable in generating accurate indications on the future values of phenomena, which are otherwise impossible or difficult to measure (Helmer and Rescher 1959). Among these methods, confidence surveys are a convenient and simple method to implement, and can effectively contribute to provide fairly accurate estimates of key statistical indicators, Internet and ICT developments further offer unprecedented opportunities to spread the use of this technique, and involve large numbers of individuals in collaborative tasks at reasonable costs (Segaran 2007), which is already a popular approach in finance (see for instance Wolfers and Zitzewitz 2004).

The analysis presented in this report adds strong evidence of the informative power of confidence surveys in producing short-term forecasts of tourism demand, across different tourism regions. The confidence index proves particularly efficient to improve forecast accuracy for destinations, whose irregular growth patterns can be hardly fit by purely statistical models. The index also contributes to increase the accuracy of forecasts for destinations with stable growth patterns. The use of the index also contributes to reduce error variance, thus making the outcome of forecasting exercises more robust.

These results have important managerial implications. The main challenges faced by tourism analysts and forecasters stem from problems related to data availability, such as the incidence of missing information, the lack of consistent data series, the need of a rather complex set of indicators explaining tourists' behaviour (Frechtling 2001). Volatility of demand, and the sensitivity of demand to external events such as war, terrorism and catastrophes, further complicates the matter and limits the performance of quantitative methods commonly used in other sectors of the economy. Due to the lack of adequate information to model and forecast tourism demand, the proposed approach certainly proves a convenient alternative to fill this gap.

The tourism confidence index is here confirmed as an effective method to obtain reasonably accurate forecasts. A simple method, bearing limited costs for organisations, the use of confidence indexes in the tourism sector should be strongly supported. A broader use of this approach indeed would not only lead to a substantial improvement of insights intelligence available to policy makers and private operators, at any level, but it would also lead to a higher acceptance of forecasts and strategic analysis by practitioners, and encourage the adoption of forward looking attitudes in atypical backward-looking sector.

References Références Referencias

1. Akın, M. (2015). "A novel approach to model selection in tourism demand modeling." Tourism Management48: 64-72.
2. Armstrong, J. S. and R. Fildes (2006). "Making progress in forecasting." International journal of forecasting 22(3): 433-441.
3. Athanasopoulos, G., R. J. Hyndman, H. Song and D. C. Wu (2011). "The tourism forecasting competition." International Journal of Forecasting 27(3): 822-844.
4. Candela, G. and P. Figini (2012). The economics of tourism destinations. The Economics of Tourism Destinations, Springer: 73-130.
5. Caniato, F., M. Kalchschmidt and S. Ronchi (2011). "Integrating quantitative and qualitative forecasting approaches: organizational learning in an action research case." Journal of the Operational Research Society: 413-424.
6. Croce, V. (2016). "Can tourism confidence index improve tourism demand forecasts?" Journal of Tourism Futures 2(1): 6-21.
7. Croce, V., K. Wöber and J. Kester (2015). "Expert identification and calibration for collective forecasting tasks." Tourism Economics (In print).
8. Diebold, F. and R. Mariano (1995). Comparing predictive accuracy, 3HMJG; EH@, MKCG.
9. Fildes, R., P. Goodwin and M. Lawrence (2006). "The design features of forecasting support systems and their effectiveness." Decis. Support Syst. 42 (1): 351-361.
10. Fildes, R., P. Goodwin, M. Lawrence and K. Nikolopoulos (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.
11. Frechtling, D. C. (2001). Forecasting Tourism Demand: Methods and Strategies. Oxford, Butterworth-Heinemann.
12. Goh, C. and R. Law (2011). "The methodological progress of tourism demand forecasting: a review of related literature." Journal of Travel & Tourism Marketing 28(3): 296-317.

13. Goh, K. L. C. (2004). A decision rule-based forecasting model for tourism demand: An application and comparison Ph.D., Hong Kong Polytechnic University (People's Republic of China).
14. Gonzalez, P. and P. Moral (1995). "An analysis of the international tourism demand in Spain." International Journal of Forecasting 11(2): 233-251.
15. Goodwin, P. (2002). "Integrating management judgment and statistical methods to improve short-term forecasts." Omega 30(2): 127-135.
16. Gordon, T. J. and O. Helmer (1964). Report on a Long-Range Forecasting Study. RAND.
17. Granger, C. W. (1969). "Investigating causal relations by econometric models and cross-spectral methods." Econometrica: Journal of the Econometric Society: 424-438.
18. Guizzardi, A. and A. Stacchini (2015). "Real-time forecasting regional tourism with business sentiment surveys." Tourism Management47: 213-223.
19. Gupta, U. G. and R. E. Clarke (1996). "Theory and applications of the Delphi technique: A bibliography (1975-1994)." Technological Forecasting and Social Change 53(2): 185-211.
20. Harvey, A. C. (1990). Forecasting, structural time series models and the Kalman filter, Cambridge university press.
21. Helmer, O. and N. Rescher (1959). "On the epistemology of the inexact sciences." Management science 6(1): 25-52.
22. Holden, K., D. A. Peel and J. L. Thompson (1990). Economic forecasting: an introduction, Cambridge University Press.
23. Kester, J. and V. Croce (2011). Tourism development in advanced and emerging economies: What does the Travel & Tourism Competitiveness Index tell us.
24. Kulendran, N. and M. L. King (1997). "Forecasting international quarterly tourist flows using error-correction and time-series models." International Journal of Forecasting 13(3): 319-327.
25. Kulendran, N. and S. F. Witt (2001). "Cointegration versus least squares regression." Annals of Tourism Research 28(2): 291-311.
26. Li, G., H. Song and S. F. Witt (2005). "Recent Developments in Econometric Modeling and Forecasting." Journal of Travel Research 44(1): 82-99.
27. Likert, R. (1932). "A technique for the measurement of attitudes." Archive of Psychology 22(140): 55.
28. Makridakis, S. and N. Taleb (2009). "Decision making and planning under low levels of predictability." International Journal of Forecasting 25(4): 716-733.
29. Makridakis, S., S. C. Wheelwright and R. J. Hyndman (1998). Forecasting: Methods and Applications. New York, Wiley.
30. Mihalic, T., J. Kester and L. Dwyer (2013). "Impacts of the global financial crisis on African tourism: a tourism confidence index analysis." Tourism and Crisis, London: Routledge: 94-112.
31. Njegovan, N. (2005). "A leading indicator approach to predicting short-term shifts in demand for business travel by air to and from the UK." Journal of Forecasting 24(6): 421-432.
32. Peng, B., H. Song and G. I. Crouch (2014). "A meta-analysis of international tourism demand forecasting and implications for practice." Tourism Management 45: 181-193.
33. Rowe, G. and G. Wright (1999). "The Delphi technique as a forecasting tool: issues and analysis." International journal of forecasting 15(4): 353-375.
34. Rowe, G. and G. Wright (2001). Expert opinions in forecasting: Role of the Delphi technique. Principles of Forecasting: A Handbook for Researchers and Practitioners. J. S. Armstrong. Norwell, Kluwer Academic Publishers.
35. Segaran, T. (2007). Programming Collective Intelligence. Sebastopol, O'Reilly.
36. Song, H. and G. Li (2008). "Tourism demand modelling and forecasting - A review of recent research." Tourism Management 29(2): 203-220.
37. Swarbrooke, J. and S. Horner (2001). Business travel and tourism, Routledge.
38. Taylor, K. and R. McNabb (2007). "Business Cycles and the Role of Confidence: Evidence for Europe*." Oxford Bulletin of Economics and Statistics 69(2): 185-208.
39. Turner, L. W. and S. F. Witt (2001). "Factors influencing demand for international tourism: Tourism demand analysis using structural equation modelling, revisited." Tourism Economics 7(1): 21-38.
40. UNWTO (2016). UNWTO World Tourism Barometer. UNWTO. Madrid, UNWTO. 14.
41. Vanhove, N. (2011). The economics of tourism destinations, Routledge.
42. Williams, P. L. and C. Webb (1994). "The Delphi technique: a methodological discussion." Journal of Advanced Nursing 19(1): 180-186.
43. Witt, S. F. and C. Witt (1995). "Forecasting tourism demand: A review of empirical research." International journal of forecasting11: 447-475.
44. Wolfers, J. and E. Zitzewitz (2004). Prediction markets, National Bureau of Economic Research.
45. Woundenberg, F. (1991). "An Evaluation of Delphi." Technological Forecasting and Social Change40: 131-150.
46. Yap, G. and D. Allen (2011). "Investigating other leading indicators influencing Australian domestic tourism demand." Mathematics and Computers in Simulation 81(7): 1365-1374.