Forecasting British Tourist Arrivals to Balearic Islands Using Meteorological Variables and Artificial Neural Networks

Marcos Álvarez-Díaz and Jaume Rosselló-Nadal
Centre de Recerca Econòmica (UIB-Sa Nostra)

Abstract

There is a clear understanding of the benefits of getting accurate predictions that allow diminishing the uncertainty inherent to the tourism activity. Managers, entrepreneurs, politicians and many other agents related to the tourism sector need good forecasts to plan an efficient use of tourism-related resources. In spite of the consensus on this need, tourism forecasters must make an even greater effort to satisfy the industry requirements. In this paper, the possibility of improving the predictive ability of a tourism demand model with meteorological explanatory variables is investigated using the case study of monthly British tourism demand to the Balearic Islands (Spain). For this purpose, a transfer function model and a causal artificial neural network are fitted. Meanwhile, the results are compared with those obtained by non-causal methods: an ARIMA model and an autoregressive neural network. The results seem to indicate that adding meteorological variable can increase the predictive power but, however, the most accurate prediction is obtained using a non-causal model, specifically an autoregressive neural network.

Keywords: Tourism forecasting, meteorological variables, artificial neural networks.
1. Introduction

Getting accurate forecasts of tourism demand is an essential requirement for different sectors of the tourism industry. Hotels, restaurants, supermarkets, ferry service firms, cruise ship lines as well as airline companies are only a few examples of business that require accurate short-term predictions to avoid shortages or surpluses in the provision of goods and services. Managers need accurate predictions to reduce the strong uncertainty inherent to the tourist activity and, therefore, improve the entrepreneurial decision making process. Apart from the private sector, policy makers are also interested in handling good predictions. Tourism has become one of the most important economic activities. It has significantly contributed to the economic development of a country through foreign exchange earnings, stimulation of infrastructure investment, employment generation and government revenues. These all reasons explain why politicians are so interested in analyzing the future evolution of the tourism activity. In general terms, tourism forecasting is of paramount importance for managers and politicians because it is a crucial source of information to plan an optimal use of tourism-related resources.

There are many different quantitative forecasting methods available in the literature on tourism demand forecasting. They can be broadly divided into non-causal models and causal econometric models (Song and Li, 2008). The non-causal models only use historical values of a time series to predict its future evolution. They have been extensively used in the last decades for tourism demand forecasting with the dominance of the integrated moving average models (ARIMAs) developed by Box and Jenkins (1976).

On the other hand, a causal econometric model attempts to model tourism demand using economic variables such as income, living expenditures and exchange
Among the causal models, multivariate regression using ordinary least squares has been one of the most widely used methods in tourism demand studies (Uysal and El Roubi, 1999). In this sense, researchers usually establish a linear relationship between tourism demand and a set of economic explanatory variables. Relatively recently, in order to avoid a spurious relationship which often appears in traditional regression analysis based on ordinary least squares, great effort has been made to further advance the econometric approach in the context of tourism modeling and forecasting. Modern econometric methods, such as the autoregressive distributed lag model, the error correction model, the vector autoregressive model, the time varying parameter model and the transfer function model have emerged as the main forecasting methods in the current tourism demand forecasting literature (Song and Li, 2008).

In spite of the econometric improvements, the conventional causal models suffer from some drawbacks. First, although the economic factor is unquestionable relevant to explain tourism demand, there are other many important factors that are not taken into account. It seems to be reasonable to think, for instance, that weather condition is a relevant explanatory factor that has a clear effect on the decision of traveling. The inclusion of weather conditions into a causal model should allow improving the predictive power. However, this possibility has not been deeply explored until now. A second drawback is due to the fact that tourism forecasters generally assume a traditional parametric point of view. According to this modeling approach, the functional form of the model is discretionally imposed by the researcher rather than observed in the data. It seems to be more suitable to consider a modeling perspective which does not assume any a priori and discretionary hypothesis on the functional form of the model and allows obtaining models in which “data speak for themselves”. This modeling procedure, called non-parametric or data-driven, allows discovering and
exploiting possible hidden nonlinear patterns. As a result of this, the forecasting accuracy could be greatly improved.

Remarkable developments of computer hardware and software have allowed an improvement and generalized use of sophisticated non-parametric quantitative forecasting methods. These techniques have been successfully applied in different fields of research, and forecasters have recently used them to model and predict tourism data. Some empirical tourism studies have employed artificial neural networks (Palmer et al., 2006), genetic programming (Alvarez-Diaz et al., 2007), support vector machine (Cheng and Wang, 2007) or fuzzy logic (Pai et al., 2006). At a first glance, the results published in the literature seem to indicate that they can be very useful techniques to predict tourism demand.

In this context, the main goal of this research is to investigate whether the inclusion of meteorological variables (days of sun, mean temperature or days of rain, among others) allows to significantly improving the predictive performance of different models, regardless of whether the forecaster assumes a parametric or non-parametric perspective. For this purpose, a transfer function model (a parametric model) and an artificial neural network (a non-parametric method) are estimated. The predictive performance is later compared with different non-causal methods. Specifically, the benchmark will be an ARIMA model and an autoregressive neural network. The variable to be predicted will be the monthly British tourism demand to the Balearic Islands (Spain). Nowadays, tourism constitutes the main and fundamental economic activity in this archipelago where more than 60% of the Balearic GDP is directly or indirectly related to the tourism. In consequence, given the obvious importance of tourism in this local economy, studies on tourism forecasting are considered of great relevance.
The remaining sections of this study are organized as follows. After this introductory section, the forecasting methods (ARIMA, transfer function and neural networks) are briefly explained. Section three describes the data and specifies the set-up of the forecasting exercise. In section four the predictive results are discussed and, finally, in section five the conclusions with a summary of the main findings are presented.

2. Forecasting methods

2.1. ARIMA model

A forecasting method that has remained very popular over the years and many times used as a benchmark approach is the ARIMA model. The specification of an ARIMA model can be expressed through the traditional formulation:

\[ \phi_p(L)Y_t = \theta_q(L)a_t, \]  

where \( Y_t \) is the tourism demand from a particular origin, \( a_t \) is the innovation or moving average term, and \( \phi_p(L) \) and \( \theta_q(L) \) are the lag operator polynomials for both \( Y_t \) and \( a_t \), respectively. Then, the forecasted values can be calculated from the expression,

\[ Y_{t+h} = \sum_{i=0}^{p} \rho_i Y_{t-i} + \sum_{j=0}^{q} \theta_j a_{t-j} + a_t, \]  

where \( \rho \) and \( \theta \) are parameters to be estimated. As usual in time series, the conventional Box-Jenkins methodology to identify the most suitable orders for the ARIMA model was followed.
2.2. Transfer functions

Additional information about the behavior of the time series to be predicted can be included in the specification of Equation [1]. This modeling procedure is called the transfer function method and presents the following structure

$$
\phi_p(L)Y_t = \theta_q(L)\alpha_t + \phi^k(L)d^k_t
$$

where \(d^k_t\) is a vector of \(k\) variables that determines tourism demand in the sense of the traditional econometric models used in the tourism demand estimation literature and described by Song and Li (2008), and \(\phi^k(L)\) are the lag operator polynomials (or transfer function) for each one of the determining \(d^k_t\) variables. In a similar way the forecasted values can be calculated from the expression:

$$
Y_t = \sum_{i=1}^{q_1} \rho_i Y_{t-i} + \sum_{i=1}^{q_2} \theta_i \varepsilon_{t-q_i} + \sum_{i=1}^{q_3} \pi_{31} d^1_{t-i} + \sum_{i=1}^{q_4} \pi_{32} d^2_{t-i} + \cdots + \sum_{i=1}^{q_k} \pi_{dk} d^k_{t-i} + \varepsilon_t
$$

where \(\rho, \theta\) are parameters to be estimated.

2.3. Artificial neural networks

Finally, artificial neural networks have been considered in order to include the most popular non-parametric technique used in modeling and forecasting tourism demand (Song and Li, 2008). This forecasting method attempts to mimic in a computer the functioning of the human brain and the nervous system. It has been widely applied to solve numerous classification and forecasting problems in diverse fields, including tourism forecasting. The main reason that explains the noteworthy use of neural networks is that they have a great capability to detect and exploit any non-linearity that might exist in the data, even under conditions of incomplete data or where the presence of noise is important. Indeed, it has been demonstrated that they can approximate any continuous function to any desired level of accuracy if they are correctly designed.
(Cybenko, 1989). Therefore, it seems that they can be more suitable than traditional methods to model and predict phenomena characterized by a complex behavior. In this sense, the majority of the empirical applications on tourism forecasting have showed that neural networks scored as good as, or significantly better than, the traditional parametric methods (Law and Au, 1999; Turner and Kon, 2005).

Although there are many types of networks, the most popular one for tourism demand forecasting is the feed-forward multi-layer network with a learning algorithm based on the back-propagation technique (Palmer et al., 2006; Kon and Turner, 2005). Other types of networks such as radial-basis function networks, recurrent neural networks or wavelets are also very useful, but much less used. Basically, a network is composed by an input layer, an output layer and one or more hidden layers. Each layer has a group of process units called neurons or nodes. These nodes are connected to nodes at adjacent layer. The connections, called synapses, are weighted by a series of coefficients. The goal will be to find the values of these weights that minimize the forecast errors.

In general terms, the predicted value of a causal neural network can be expressed as

$$\hat{Y}_t = \Phi \left( \sum_{h=1}^{H} \sum_{p} \beta_p \cdot \Psi \left( \alpha_0 + \sum_{k=1}^{K} \sum_{j=1}^{J} \alpha_{kj} \cdot x_{k,t-j} + \sum_{p=1}^{P} \sum_{j=1}^{J} \varphi_{pj} \cdot Y_{t-p} \right) \right)$$

where $\hat{Y}_t$ is the predicted value (output of the neural network model), the input vector contains delays of $K$ explanatory variables $(\{x_1, x_2, ..., x_k\})$ and $P$ lagged-values of the dependent variable $(\{Y_{t-1}, Y_{t-2}, ..., Y_{t-P}\})$, the functions $\Psi(\cdot)$ and $\Phi(\cdot)$ are denoted as transfer functions of the hidden and output levels, and $H$ is the number of process units in the hidden layer. The parameters of the model $\alpha_{kj}$, $\varphi_{pj}$ and $\beta_p$ are called weights, and
they are randomly determined within a given range of values. By means of an iterative learning process based on the back-propagation technique (Rumelhart et al., 1986), these weights are modified in order that some predictive error measure such as the sum of the squared errors (SSE) is minimized (Zhang et al., 1997). In the particular case of a non-causal forecasting problem, the inputs are only the past values of the time series under analysis \((\alpha_{kj} = 0, \forall k, h, j)\). Thus the network, called autoregressive neural network, is equivalent to a general nonlinear autoregressive model.

In addition to the complexity of the data, the predictive success of a neural network depends on the correct determination of its architecture. It is therefore necessary to accurately specify the vector of inputs and the number of process units in the hidden level \((H)\), as well as to select the right structure of the transfer functions, \((\Psi(\cdot)\) and \(\Phi(\cdot)\)). An excessive number of \(H\), for example, might create overfitting problems, thus eliminating any generalization. On the other hand, with an insufficient number of process units, the network could lose its forecasting capability because it would not fully exploit the non-linearity in the data. In the literature, one can find different rules for defining how many inputs and process units there must be in the hidden unit, but none of them is perfect nor have any of them ever been adopted as a general rule (Yao, Tan and Poh, 1999). A frequent recommendation is to determine the number of inputs and neurons through a process of trial-\textit{and}-error. Therefore, following this recommendation, it is necessary to consider different architectures and choose that one which produces the minimum predictive error in a sub-set reserved exclusively to select an optimal architecture (i.e., the \textit{selection set}).
3. Forecasting Design

3.1. Data and variables of the models

The variable to be predicted in this study is the number of tourist air arrivals to Balearic Islands from United Kingdom. Air arrivals constitute 95% of tourist arrivals to the Balearic Islands; in consequence sea arrivals were disregarded. The time series comes from the CITTIB (Balearic Tourist Statistical Office, office belonging to the Balearic government). The periodicity of the data is monthly and the sample period goes from December 1981 to December 2006. Therefore, it contains a total of 301 observations. The choice of tourist arrivals as variable of interest is justified because is still the most popular measure of tourism demand over the past few years (Song and Li, 2008).

Previous studies have suggested that the economic variables constitute the main explanatory factor in forecasting models for tourism demand (Crouch, 1996; Lim 1999). The justification for selecting economic variables is well established in the literature (Witt and Witt, 1992) and, among them, income, exchange rates and prices are considered to play a central role (Crouch, 1994). For this reason, in this study the nominal exchange rate of the British Pound against the Euro (EX), the British Industrial Production Index (IPI) and the relative consumer price index between Spain and Great Britain (CPI) are considered as explanatory variables. The data were obtained from different statistical sources: the Spanish Central Bank in the case of the nominal exchange rate, the Spanish National Institute of Statistics for the Spanish consumer price index, and the British Office for National Statistics for the industrial production index and the consumer price index in Great Britain.
Apart from the economic factors, tourism demand is also influenced by many other factors. The goal of this study is to check whether the incorporation of meteorological variables in a model with only lagged-values of the dependent variable and economic variables would significantly improve the predictive ability. Certainly, the weather conditions in an origin country could have a clear impact on the number of tourist arrivals to a specific place, and it could be adequate to include these variables as explanatory factors. For this reason, in addition to the economic variables, meteorological data from the origin is included in the forecasting models.

Data for the monthly climatic variables for the UK were collected from the British Public Weather Service and they include mean maximum daily temperature, mean minimum temperature, days of air frost (AF), total rainfall (RAIN), total sunshine duration (SUN). The variables were available for 21 stations along the UK territory (Figure 1) for the period January 1980 December 2006, then, means were calculated in order to get country indicators. On the other hand, different transformations of the climatic variables were calculated in order to get more reasonable indicators.

Then, monthly mean temperature (TMEAN) was calculated as the average of the maximum and minimum temperatures. However, it could be argued that the variation of tourism demand with temperature could be non-linear, increasing both for decreasing and for increasing temperatures. In this way, cold temperatures could persuade tourists to travel to warm regions, while hot temperatures could too in the sense that warmer regions are more prepared to face up heat waves. This non-linear response of the tourism demand is often applied in electricity demand estimation (Pardo, Meneu and
Valor, 2002). Specifically, it is suggested the use of two temperature-derived functions allowing the separation of the winter and summer data. Such a separation will help to obtain better results in the linear model estimations and can be achieved by using the degree–day functions defined as heating degree–days \( \text{HITEMP} = \max(T_{\text{ref}} - T_{\text{MEAN}}, 0) \), and cooling degree–days \( \text{LOWTEMP} = \max(T_{\text{MEAN}} - T_{\text{ref}}, 0) \), where \( T_{\text{ref}} \) is a reference temperature that should be adequately selected to separate the cold and heat branches of the demand–temperature relationship. Then, the degree–day functions measure the intensity and duration of cold or heat in winter and summer months, respectively.

3.2. Sub-samples: training, selection and out-of-sample

Usually, in the majority of the tourism forecasting applications, the total available data are divided into an in-sample set and a test set (out-of-sample or hold-out sample). In theory, the in-sample set is used for the construction of the model while the test set is employed for measuring the predictive ability of the method. Nevertheless, this methodological procedure could not be adequate employing computational forecasting methods and, specifically, neural networks. Researchers using neural networks can be tempted to try different architectures in the in-sample set and select that with the highest accuracy in the test set. The result would be a network with a strong predictive ability in the test set but, however, it would be not capable of generalizing and performing well with new data. The forecasting exercise would suffer from an overfitting problem and the forecasting utility would be practically scarce.

In order to avoid these problems and develop a useful and fair predictive exercise, the technical and practical recommendation proposed in the literature on Computer Science (Bishop, 1995), and more recently in tourism forecasting (Palmer et
al., 2006), is adopted here. These studies advise to divide the sample period into three sub-samples: training, selection and out-of-sample. In particular, in the application the first 211 observations are used in order to train the network. The next 60 observations are employed to select an optimal architecture (specifically, inputs and number of hidden neurons). Finally, the last 30 observations (the last two years and a half) are not used in the modeling process. These untouched data is the out-of-sample set. The value of the accuracy measure obtained in this sub-sample is employed to validate the predictive ability of the network.

This methodological procedure is necessary to detect possible spurious results and guaranty that the network is capable of generalizing and performing well with new data. The network must show a relatively high fitness in the out-of-sample set. If this condition was verified, it would be proved the ability of the neural network architecture to generalize new observations and, therefore, there would not be suspects of overfitting problems.

3.3. Forecasting Accuracy Measures

There is no consensus on which is the most appropriate measure to assess the performance of a forecasting method. Different accuracy measures are available for tourism demand forecasting evaluation but the predominant one in the majority of the studies is the mean absolute percentage error (MAPE) (Li, Song and Witt, 2005). MAPE is calculated from the expression

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \frac{|e_t|}{y_t}$$  \[6\]
where $e_t$ is the forecasting error, $Y_t$ is the time series to be predicted and $T$ is the total number of observations. Certainly, MAPE is a relative error magnitude measure and it has some characteristics considered appropriate for an accuracy measure such as being independent of scale, easy to interpret, reliable and valid (Amstrong, 2001). Nevertheless, many authors think that it can be more adequate to use a measure specified in terms of square errors than one in terms of absolute errors (Witt and Witt, 1991). The reason is that a measure based on square errors gives greater weight to large errors than to smaller ones because the errors are squared before being summed. Since the ability of a forecasting method to detect large errors is often regarded as one of the most important criteria, different goodness of fit based on the sum of the square errors have been popular for years. In this sense, besides MAPE, the root mean square percentage error (RMSPE) is also considered

$$RMSPE = \frac{1}{\sqrt{T}} \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left( \frac{e_t}{Y_t} \right)^2} \cdot 100$$  \[8\]

which is a measure more sensitive to one extremely bad forecast than MAPE. Indeed, it heavily penalizes those wrong predictions caused by external shocks. Therefore, a conservative or cautious forecaster could prefer the RMSPE criterion to choose the forecasting method that most avoids these large mistakes. In general, MAPE and RMSPE are considered good measures to use in evaluating the forecasting of tourism demand models (Witt and Witt, 1992; Song and Witt, 2000), and the choice of one or another will be purely a matter of personal preferences of the forecaster.
3.4. Statistical significance of the predictive improvement

The majority of empirical studies on tourism forecasting do not provide any test to support the statistical significance of the predictions. This deficiency supposes an important drawback in any predictive exercise. As Li, Song and Witt (2005) indicated, formal statistical tests need to be applied to examine if the difference in the accuracy of competing forecasts is statistically significant. According to this requirement, the Diebold-Mariano test (Diebold and Mariano, 1994), a test that allows to assert whether one forecasting method is significantly better than other, is applied. Formally, given two forecasting competing methods, this statistical test is defined as

\[ D - M \text{ Test } = \frac{\bar{d}}{2\pi \hat{f}_d(0)^{\frac{1}{2}}} \]  

where \( T \) is the sample size, \( \hat{f}_d(0) \) is a consistent estimate of the spectral density of the loss differential at frequency zero corrected for serial correlation and \( \bar{d} \) is the sample mean loss differential based on a quadratic cost function

\[ \bar{d} = \frac{\sum (error_{Method1}^2 - (error_{Method2}^2)^2)}{T} \]  

or an absolute cost function

\[ \bar{d} = \frac{\sum |error_{Method1} - error_{Method2}|}{T} \]  

Diebold and Mariano (1995) demonstrated that under the null hypothesis of equal forecasts ability between method 1 and method 2
the D-M test follows asymptotically a standard normal distribution. Therefore, a negative and statistically significant D-M test would imply to reject the null hypothesis and, in consequence, it could be asserted that the forecasting method 1 provides statistically better predictions than method 2.

4. Results

Once the forecasting design has been defined and some important aspects specified, in this section the results obtained using the different forecasting methods and explanatory variables considered in the study are presented. The results will be focused on the out-of-sample analysis because it is the forecasting capability that researchers, practitioners and politicians are most interested in.

4.1. Transfer Functions

Table 1 shows the out-of-sample forecasting results obtained using a transfer function model with only economic variables, and another one with the same variables but including meteorological information. The selection of variables is carried out by the backward stepwise technique. According to this procedure, several delays of all variables have been initially included. The less significant delays are discarded from the model and a new transfer function is estimated without the deleted delays. The process is repeated until finding a model where all variables are statistically significant at a 10 percent level of significance.

In the case of the transfer function with an autoregressive/moving average structure and economic variables, it can be observed how the variables IPI with six
months of delay and EX with twelve months of delay are the most significant economic
variables to explain the number of arrivals. Regarding the forecasting accuracy, the
MAPE reaches a percentage of 7.96%. The low MAPE indicates that the deviation
between the predicted and the actual values are very small. In this sense and following
the scale developed by Lewis (1982), these forecasts would be judged as highly
accurate. On the other hand, the forecasting accuracy in terms of RMSPE reflects a
percentage of 9.68%.

When economic and meteorological information are included into the analysis,
the structure of the transfer function basically does not change. The delays of the
economic variables and the structure of the autoregressive and moving average are the
same than before but, however, some delays of meteorological variables are selected by
the backward procedure. Specifically, the days of sun six months ago and the mean
temperature two months ago have a statistical significant effect on the number of
arrivals. However, in spite of being statistically significant, these variables do not
suppose a predictive improvement neither in terms of MAPE (8.08%) nor in terms of
RMSPE (10.75%). It seems, therefore, that the inclusion of meteorological information
is not useful from a predictive point of view. However, these results should not be
considered definitive or conclusive. It should not be forgeted that the modeling
approach followed until now is based on presupposing a restrictive hypothesis on the
functional form of the model. Therefore, It is necessary to go further on the analysis and
check if the use of an artificial neural network and the inclusion of meteorological
variables allows to significantly improve the predictive power.
4.2. Causal Neural Networks

Table 2 shows the out-of-sample forecasting results using a causal neural network with economic variables, and another one with economic and meteorological variables. In each case, the architecture of the network (inputs and number of neurons) was that one that achieved the highest fitness in the selection period. Additionally, Tables 4 contains the values of the D-M test and their associated p-values assuming an absolute cost function; and Table 5 presents the same information but considering a quadratic function. The interpretation of the tables is as follows: a positive and significant value of the D-M test implies that the forecasting method on the left side provides statistically better predictions than the method on the top.

A detailed analysis of Table 2 allows to observe how the network with economic variables gets a MAPE and a RMSPE equal to 7.11% and 8.85%, respectively. When meteorological information is added, the predictive accuracy is higher in terms of MAPE (6.27%), but it is worse in terms of RMSPE (9.05%). In this case, the apparent inconsistency between the ranking given by MAPE and RMSPE is due to different assumptions regarding the forms of the loss function. This fact does not support the belief in tourism forecasting that both measures indicate approximately the same ranking of models (Witt and Witt, 1992). Therefore, the choice of the best forecasting model will depend on the fitness criterion considered in the predictive exercise. Nevertheless, it must also be remarked that neither the improvement in terms of MAPE nor the worsening in terms of RMSPE are statistically significant (D-M test=0.7092 with a p-value=0.4782 and D-M test=-0.1161 with a p-value=0.9075, respectively).

Regarding the comparison with the transfer function models, the results obtained here support the convenience of using neural networks in forecasting tourism time series. Specifically, the causal neural networks outperform in all cases the transfer
functions models. However, this improvement is only statistically significant in the case of the neural network with economic and meteorological variables and when the MAPE is used as fitness criterion (D-M test=1.7279 with a p-value=0.084 and D-M test=2.5688 with a p-value=0.010).

4.3. Non-causal Methods

Table 3 contains the out-of-sample predictive results for the non-causal forecasting methods. The ARIMA model obtains a fitness value of equal to 6.77%, and 8.68% in terms of MAPE and RMSPE, respectively. On the other hand, the autoregressive network reaches a MAPE and RMSPE equal to 5.76% and 8.82%, respectively. At a first glance, some general comments can be made. First of all, the results corroborate the empirical evidence that non-causal models are more likely to outperform causal models in short term forecasts (Martin & Witt, 1989; Preez & Witt, 2003; Sheldon, 1993; Witt & Witt, 1992; Dharmaratne, 1995). The non-causal methods provide the most accurate predictions both in terms of MAPE as well as in terms of RMSPE. Second, the autoregressive network is the best forecasting technique in terms of MAPE. It outperforms the rest of the methods considered in this study, but the predictive improvement is only significantly more accurate when the comparison is done with the transfer function models. Finally, the comparison between the non-causal methods reflects that the neural network is better when the MAPE criterion is considered, but worse in terms of RMSPE. It seems that the neural network approach adjusts better the data and obtains better predictions but, however, it is more sensitive to outliers or external shocks than ARIMA.
5. Conclusion

For a long time it has been recognized the need of getting accurate forecasts for tourism demand. Increasing the accuracy of short-term forecasts is an essential requisite to improve the managerial, operational, and tactical decision making process. For this reason, tourism forecasters must make greater efforts to get more accurate forecasts and satisfy even better the predictive requirements of the tourism sector. In spite of the efforts made and the clear advances achieved in the field, there still exist several gaps that must be fulfilled. In this sense, this paper tries to make some contributions to the literature on tourism forecasting in different ways. First, it explores whether the inclusion of meteorological information into a tourism demand model allows a significant improvement of the predictive ability. Second, it analyzes whether the traditional parametric modeling procedure leads to a predictive loss. Third, the belief that non-causal models provide better predictions than causal models is checked. Fourth, it has been analyzed whether the selection of a specific fitness measure modifies the predictive ranking of models. Finally, the research supposes a complete and fair predictive exercise, especially when applying neural networks. In this sense, the use of three sub-samples (training, selection and out-of-sample) in this forecasting study implies a methodological improvement in front of the traditional procedure of applying only two sub-samples (in-sample and out-of-sample). Moreover, as it was recommended and demanded in the literature (Li et al., 2005), formal test to examine if the difference in the accuracy of competing forecasts is statistically significant is applied.

The results indicate that adding meteorological variable can increase the predictive power when an artificial neural network is used, and only when the MAPE is adopted as fitness criterion. Therefore, on one hand, the use of a non-parametric
perspective seems more adequate. The greater flexibility in the modeling procedure permits an improvement and, in some cases, this predictive gain is statistically significant. Moreover, the results obtained here support the idea that non-causal methods provide better predictions in the short term. In any case, the ranking of models according to their predictive power will depend on the fitting criterion chosen by the forecaster.

Future research should try to go in deep in the kind of relationships that connect tourism demand series and climatic variables, looking for non-linear interactions and investigating the pool of meteorological variables that can be related with tourism demand.

References


Amstrong J. S. (2001) Principles of Forecasting, a handbook for researchers and practitioners, Kluwer Publisher, USA.


Figure 1. Location of meteorological stations in the UK.

Source: Public Weather Service

Table 1. Out-of-sample forecasting results using transfer function models

<table>
<thead>
<tr>
<th>METHOD</th>
<th>VARIABLES</th>
<th>Out-of-Sample MAPE</th>
<th>Out-of-Sample RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer Function with Economic Variables</td>
<td>DLIPI(-6); DLEX(-12); AR(1); AR(2); AR(3); MA(11); MA(12)</td>
<td>0.0796</td>
<td>0.0968</td>
</tr>
<tr>
<td>Transfer Function with Economic and Meteorological Variables</td>
<td>DLIPI(-6); DLEX(-12); TMEAN(-2); SUN(-6); AR(1); AR(2); AR(3); MA(11); MA(12).</td>
<td>0.0808</td>
<td>0.1075</td>
</tr>
</tbody>
</table>
Table 2. Out-of-sample forecasting results using causal neural networks

<table>
<thead>
<tr>
<th>METHOD</th>
<th>VARIABLES</th>
<th>Out-of-Sample MAPE</th>
<th>Out-of-Sample RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Network with Economic Variables</td>
<td>AR(1); AR(2); AR(3); AR(4); AR(5); AR(6); AR(7); AR(8); AR(9); AR(10); AR(11); AR(12); AR(13); IPI(-6); EX(-2).</td>
<td>0.0711</td>
<td>0.0885</td>
</tr>
<tr>
<td>Neural Network with Economic and Meteorological Variables</td>
<td>AR(1); AR(2); AR(3); AR(4); AR(5); AR(6); AR(7); AR(8); AR(9); AR(10); AR(11); AR(12); AR(13); IPI(-6); EX(-2); TMEAN(-12); SUN(-1); LOTEM(-12).</td>
<td>0.0627</td>
<td>0.0905</td>
</tr>
</tbody>
</table>

Table 3. Out-of-sample forecasting results using causal neural networks

<table>
<thead>
<tr>
<th>METHOD</th>
<th>VARIABLES</th>
<th>Out-of-Sample MAPE</th>
<th>Out-of-Sample RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>AR(1); AR(2); AR(3); MA(11); MA(12)</td>
<td>0.0677</td>
<td>0.0868</td>
</tr>
<tr>
<td>Autoregressive Neural Network</td>
<td>AR(1); AR(2); AR(3); AR(4); AR(5); AR(6); AR(7); AR(8); AR(9); AR(10); AR(11); AR(12); AR(13).</td>
<td>0.0576</td>
<td>0.0882</td>
</tr>
</tbody>
</table>
Table 4. Values of the Diebold-Mariano Test based on a quadratic cost function (MAPE)

<table>
<thead>
<tr>
<th>CAUSAL MODELS</th>
<th>NON-CAUSAL MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRANSFER FUNCTION MODEL</td>
<td>ARTIFICIAL NEURAL NETWORKS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic &amp; Meteorological Variables</th>
<th>Economic &amp; Meteorological Variables</th>
<th>Economic &amp; Meteorological Variables</th>
<th>Economic &amp; Meteorological Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAUSAL MODELS</strong></td>
<td><strong>NON-CAUSAL MODELS</strong></td>
<td><strong>ARIMA</strong></td>
<td><strong>AUTOREGRESSIVE NEURAL NETWORK</strong></td>
</tr>
<tr>
<td><strong>TRANSFER FUNCTIONS</strong></td>
<td><strong>ARTIFICIAL NEURAL NETWORKS</strong></td>
<td><strong>ARIMA</strong></td>
<td><strong>AUTOREGRESSIVE NEURAL NETWORK</strong></td>
</tr>
<tr>
<td><strong>with Economic Variables</strong></td>
<td><strong>with Economic and Meteorological Variables</strong></td>
<td><strong>with Economic Variables</strong></td>
<td><strong>with Economic and Meteorological Variables</strong></td>
</tr>
<tr>
<td><strong>with Economic Variables</strong></td>
<td>-</td>
<td>0.1068 (0.9149)</td>
<td>-1.7279* (0.0840)</td>
</tr>
<tr>
<td><strong>with Economic and Meteorological Variables</strong></td>
<td>-</td>
<td>-0.7502 (0.4531)</td>
<td>-2.5688* (0.0102)</td>
</tr>
<tr>
<td><strong>ARTIFICIAL NEURAL NETWORKS</strong></td>
<td><strong>with Economic Variables</strong></td>
<td>-0.7092 (0.4782)</td>
<td>-0.3659 (0.7144)</td>
</tr>
<tr>
<td><strong>with Economic Variables</strong></td>
<td>0.7525 (0.4824)</td>
<td>-</td>
<td>-0.7092 (0.4782)</td>
</tr>
<tr>
<td><strong>with Economic and Meteorological Variables</strong></td>
<td>0.7502 (0.4531)</td>
<td>-</td>
<td>-0.7092 (0.4782)</td>
</tr>
<tr>
<td><strong>with Economic Variables</strong></td>
<td>1.7279* (0.0840)</td>
<td>2.5688* (0.0102)</td>
<td>0.7092 (0.4782)</td>
</tr>
<tr>
<td><strong>with Economic and Meteorological Variables</strong></td>
<td>1.7279* (0.0840)</td>
<td>2.5688* (0.0102)</td>
<td>0.7092 (0.4782)</td>
</tr>
<tr>
<td><strong>ARIMA</strong></td>
<td>1.2045 (0.2284)</td>
<td>1.0795 (0.2804)</td>
<td>0.3659 (0.7144)</td>
</tr>
<tr>
<td><strong>AUTOREGRESSIVE NEURAL NETWORK</strong></td>
<td>2.5077*** (0.0122)</td>
<td>2.0357** (0.0418)</td>
<td>1.4097 (0.1586)</td>
</tr>
</tbody>
</table>
Table 5. Values of the Diebold-Mariano Test based on an absolute cost function (RMSPE)

<table>
<thead>
<tr>
<th></th>
<th>CAUSAL MODELS</th>
<th>NON-CAUSAL MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TRANSFER FUNCTIONS</td>
<td>ARTIFICIAL NEURAL NETWORKS</td>
</tr>
<tr>
<td></td>
<td>with Economic Variables</td>
<td>with Economic and Meteorological Variables</td>
</tr>
<tr>
<td>CAUSAL MODELS</td>
<td>TRANSFER FUNCTIONS</td>
<td>with Economic Variables</td>
</tr>
<tr>
<td></td>
<td>with Economic and Meteorological Variables</td>
<td>-</td>
</tr>
<tr>
<td>ARTIFICIAL NEURAL NETWORKS</td>
<td>with Economic Variables</td>
<td>0.6280 (0.5300)</td>
</tr>
<tr>
<td></td>
<td>with Economic and Meteorological Variables</td>
<td>0.5595 (0.5758)</td>
</tr>
<tr>
<td>NON-CAUSAL MODELS</td>
<td>ARIMA</td>
<td>1.0556 (0.2911)</td>
</tr>
<tr>
<td></td>
<td>AUTOREGRESSIVE NEURAL NETWORK</td>
<td>0.5488 (0.5831)</td>
</tr>
</tbody>
</table>