Institut de Recerca en Economia Aplicada Regional i Pública Document de Treball 2015/02 1/28
Research Institute of Applied Economics Working Paper 2015/02 1/28

Grup de Recerca Anàlisi Quantitativa Regional Document de Treball 2015/02 1/28
Regional Quantitative Analysis Research Group Working Paper 2015/02 1/28

"Multiple-input multiple-output vs. single-input single-output neural network forecasting"

Oscar Claveria, Enric Monte and Salvador Torra

Research Institute of Applied Economics

Regional Quantitative Analysis Research Group

Abstract

This study attempts to improve the forecasting accuracy of tourism demand by using the existing common trends in tourist arrivals form all visitor markets to a specific destination in a multiple-input multiple-output (MIMO) structure. While most tourism forecasting research focuses on univariate methods, we compare the performance of three different Artificial Neural Networks in a multivariate setting that takes into account the correlations in the evolution of inbound international tourism demand to Catalonia (Spain). We find that the MIMO approach does not outperform the forecasting accuracy of the networks when applied country by country, but it significantly improves the forecasting performance for total tourist arrivals. When comparing the forecast accuracy of the different models, we find that radial basis function networks outperform multilayer-perceptron and Elman networks.

JEL classification: C22, C45, C63, L83, R11
Keywords: tourism demand, forecasting, multivariate, multiple-output, artificial neural networks

Oscar Claveria. AQR Research Group-IREA. Department of Econometrics. University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain. E-mail: oclaveria@ub.edu

Enric Monte. Department of Signal Theory and Communications, Polytechnic University of Catalunya. E-mail: enric.monte@upc.edu

Salvador Torra. Department of Econometrics & Riskcenter-IREA, Universitat de Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain. E-mail: storra@ub.edu

Acknowledgements

This work was supported by the *Ministerio de Economía y Competitividad* under the SpeechTech4All Grant (TEC2012-38939-C03-02).

1. Introduction

Tourism demand forecasting has become essential in one of today's fastest growing industries. Song and Li (2008) have acknowledged the importance of applying new approaches to tourism demand forecasting in order to improve the accuracy and the performance of the methods of analysis. Whilst most research efforts focus on conventional tourism forecasting methods (Gounopoulos, Petmezas, & Santamaria, 2012) or a combination of them (Chan, Witt, Lee, & Song, 2010), in recent years the availability of more advanced forecasting techniques and the requirement for more accurate forecasts of tourism demand have led to a growing interest in Artificial Intelligence (AI) techniques (Wu, Law, & Xu, 2012; Cang, 2013; Pai, Hung, & Lin, 2014). The suitability of AI models to handle nonlinear behaviour is one of the reasons why Artificial Neural Networks (ANNs) are increasingly used for forecasting purposes (Haviluddin & Rayner, 2014; Claveria, Monte, & Torra, 2014; 2016).

In spite of the increasing interest in AI methods for time series forecasting (Uysal, 2004), very few studies compare the accuracy of different ANN architectures for tourism demand forecasting. This study seeks to break new ground by comparing the performance of three different ANN models in a multivariate setting that takes into account the common trends in inbound international tourism demand shared by all visitor markets to a specific destination. We use three ANNs: the multi-layer perceptron (MLP) network, the radial basis function (RBF) network and the Elman network. ANNs are able to learn from experience. There are two major learning paradigms: supervised learning and non-supervised learning. MLP networks are supervised learning models, while RBF networks, combine both learning methods (hybrid learning). Each ANN architecture handles information in a different manner, so by comparing the different models we can evaluate the impact of alternative ways of processing data on forecast accuracy.

The present study deals with tourist arrivals to Catalonia, which is a region of Spain. Barcelona is the capital of Catalonia, and the most important destination in Spain. After France and the United States, Spain is the third most important destination of the world with 60 million tourist arrivals in 2013. Catalonia received 15,5 million tourists in 2013, up 8% over the previous year. Tourist spending grew by 14% in 2013, and it accounted for 25% of tourism revenues in Spain. In relation to 2012, the expenditure per tourist raised by 7.2%, while the expenditure per day by 4.6%. It follows that tourism is one of the fastest growing industries in Catalonia, accounting for 12% of GDP and providing employment for 15% of the working population. These figures show the importance of accurate forecasts of tourism volume at the destination level for policy makers and professionals in the tourism industry. Capó, Riera, and Rosselló (2007) and Balaguer and Cantavella-Jordá (2002) have shown the important role of tourism in the Spanish long-run economic development.

The main objective of this study is to improve forecasts of tourism demand with ANN models by using the common trends in inbound international tourism demand form all visitor markets to Catalonia. With this aim, we undertake an out-of-sample forecasting competition and compare the performance of three different ANN models in a multiple-input multiple-output (MIMO) structure to those of a single-input single-output (SISO) structure, in which forecasts are obtained country by country. Given that univariate specifications are limited and unable to capture dynamic interrelationships between different countries of origin, we analyze whether a multivariate approach, in which information about tourist arrivals from all origin countries is simultaneously used, provides useful for forecasting purposes. To our knowledge, this is the first study to analyze the forecasting performance of ANNs in a MIMO setting, using the correlated growth rates between all visitor markets to a specific destination.

We obtain forecasts of tourism demand in all countries of origin for different forecast horizons (1, 3 and 6 months). In addition, we compute a measure of forecast accuracy to compare the forecasting performance of the three NN architectures. Finally, we run the Diebold-Mariano test for significant differences between each two competing series. Another major contribution of this study is to assess the effects of expanding the memory on forecast accuracy. In order to do so, we repeat the experiment assuming different topologies with respect to the number of lags used for concatenation.

The article proceeds as follows. The next section reviews the literature on tourism demand forecasting with ANNs. Then, the different NN architectures used in the analysis are presented. Data is described in the fourth section. In the following section, results of the out-of-sample forecasting competition are discussed. Finally, the last section provides a summary, a discussion of the implications, and potential lines for future research.

2. Literature review

A growing body of literature has focused on tourism demand forecasting, but most research efforts apply conventional forecasting methods, either casual econometric models (Cortés-Jiménez & Blake, 2011; Page, Song, & Wu, 2012, Nordström, 2004), time series models (Chu, 2008, 2011; Assaf, Barros, & Gil-Alana, 2011; Gounopoulos, Petmezas, & Santamaria, 2012), or a combination of them (Shen, Li, & Song, 2008; Coshall & Charlesworth 2010). See Li, Song and Witt (2005), Song and Li (2008) and Peng, Song, and Crouch (2014) for a thorough review of tourism demand forecasting studies. Nevertheless, the need for more accurate forecasts has led to an increasing use of AI techniques, such as fuzzy time series models and support vector machines (SVMs), or a mix of them (Hadavandi, Shahanaghi, & Abbasian, 2011; Shahrabi, Hadavandi, & Asabi 2013; Pai, Hung, & Lin 2014; Cang & Yu 2014), in order to obtain more refined predictions of tourist arrivals at the destination level.

Yu and Schwartz (2006) and Tsaur and Kuo (2011) use fuzzy time series models in predicting annual U.S. tourist arrivals and monthly tourism demand in Taiwan respectively. Goh, Law, and Mok (2008) apply a rough sets algorithm to forecast U.S. and U.K. tourism demand for Hong Kong. The use of genetic algorithms for parameter selection has led to increased use of support vector machines (SVMs) (Pai & Hong, 2005) and their regression version (Chen & Wang, 2007; Chen, 2011; Hong, Dong, Chen, & Wei, 2011). Wu, Law, and Xu (2012) use a sparse GP regression (GPR) model for tourism demand forecasting in Hong Kong and find that its forecasting capability outperforms those of the ARMA and SVM models. Bloom (2005) implements a self-organizing (SOM) neural network for segmenting the international tourist market to Cape Town. In a recent meta-analysis of published tourism forecasting studies, Kim and Schwartz (2013) find that forecast accuracy is closely associated with data characteristics. The fact that ANNs are data-driven procedures that learn from past experience explain the growing interest in ANNs for tourism demand forecasting (Lin, Chen, & Lee, 2011; Teixeira & Fernandes, 2012; Claveria & Torra, 2014).

ANNs can be classified into two major types of architectures: feed-forward networks and recurrent networks. MLP networks are the most widely used feed-forward topology in tourism demand forecasting (Pattie & Snyder, 1996; Uysal & El Roubi, 1999; Law, 2000, 2001; Tsaur, Chiu, & Huang 2002; Kon & Turner, 2005; Zhang & Qi, 2005). A class of multi-layer feed-forward architecture with two layers of processing is the radial basis function (Broomhead & Lowe, 1988). RBF networks have the advantage of not suffering from local minima in the same way as MLP networks, which explains their increasing use in many fields. Cang (2013) uses RBF, MLP and SVM ANN forecasts in non-linear combination models. Recurrent networks are models with bidirectional data flow which allow for a temporal feedback from the outer layers to the lower layers. This feature is specially suitable for time series modelling. A special case of recurrent network is the Elman network (Elman, 1990). Whilst MLP networks are increasingly used with forecasting purposes, Elman neural networks have been scarcely used with forecasting purposes. The only previous study that uses Elman ANNs for tourism demand forecasting is that of Cho (2003), who applies the Elman architecture to predict the number of arrivals from different countries to Hong Kong.

Multivariate approaches to tourist demand forecasting are also few and have yielded mixed results. Athanasopoulos and Silva (2012) compare the forecasting accuracy of exponential smoothing methods in a multivariate setting against univariate alternatives. They use international tourist arrivals to Australia and New Zealand and find that multivariate models improve on forecast accuracy over the univariate alternatives. Contrary to what could be expected, du Preez and Witt (2003) find that multivariate time series models did not generate more accurate forecasts than univariate time series models.

With regard to studies on tourism in Spain at regional level, there have been several articles published in recent years (Aguiló & Rosselló, 2005; Roselló, Aguiló, & Riera, 2005; Garín-Muñoz & Montero-Marín, 2007; Bardolet & Sheldon, 2008; Santana-Jiménez & Hernández, 2011; Nawijn & Mitas, 2012; Andrades-Caldito, Sánchez-Rivero, & Pulido-Fernández, 2013; Cirer-Costa, 2014). Concerning tourism demand forecasting, Palmer, Montaño, and Sesé (2006) design a MLP neural network to forecast tourism expenditure in the Balearic Islands. Medeiros, McAleer, Slottje, Ramos, and Rey-Maquieira. (2008) develop a NN-GARCH model to estimate demand for international tourism also in the Balearic Islands. Bermúdez, Corberán-Vallet, and Vercher (2009) calculate prediction intervals for hotel occupancy in three provinces in Spain by means of a multivariate exponential smoothing. Claveria and Datzira (2009, 2010) use consumer expectations derived from tendency surveys to forecast tourism demand in Catalonia. Guizzardi and Stacchini (2015) also make use of business sentiment indicators form tendency surveys for real-time forecasting of hotel arrivals at a regional level, improving the forecasting accuracy of structural time series models.

3. Methodology

ANNs emulate the processing of human neurological system to identify related spatial and temporal patterns from historical data. ANNs learn from experience and are able to capture functional relationships among the data when the underlying process is unknown. The data generating process of tourist arrivals is too complex to be specified by a single linear algorithm, which explains the great interest that ANNs have aroused for tourism demand forecasting. As opposed to the traditional model-based methods, ANNs do not depend on a set of a priori assumptions, so to obtain a reliable network the parameters of the model are iteratively estimated by means of different algorithms.

Most of the algorithms used in training artificial neural networks employ some form of gradient descent. Therefore, each network is suited to a combination of a learning paradigm and a learning algorithm (forward-propagation, back-propagation, etc.). The main learning paradigms are supervised learning and non-supervised learning. In supervised learning weights are adjusted to approximate the network output to a target value for each pattern of entry, while in non-supervised learning the subjacent structure of data patterns is explored so as to organize such patterns according to their distances. The combination of both learning methods implies that part of the weights is determined by a supervised process while the rest are determined by non-supervised learning. This is known as hybrid learning. An example of hybrid model is the RBF network.

ANNs are composed of interconnected processing units called neurons and can also be classified into feed-forward networks and recurrent networks depending on the connecting patterns of the different layers of neurons. In feed-forward networks the information runs only in one direction, whilst in recurrent networks there are feedback connections from outer layers of neurons to lower layers of neurons. Feed-forward networks were the first ANNs devised. The MLP network is the most widely used feed-forward topology in tourism demand forecasting.

3.1. Multi-layer perceptron (MLP) neural network

MLP networks consist of multiple layers of computational units interconnected in a feed-forward way. MLP networks are supervised neural networks that use as a building block a simple perceptron model. The topology consists of layers of parallel perceptrons, with connections between layers that include optimal connections. The number of neurons in the hidden layer determines the MLP network's capacity to approximate a given function. In order to solve the problem of overfitting, the number of neurons was estimated by cross-validation. In this work we used the MLP specification suggested by Bishop (1995) with a single hidden layer and an optimum number of neurons derived from a range between 5 and 25:

$$
\begin{aligned}
& y_t = \beta_0 + \sum_{j=1}^{q} \beta_j g\left(\sum_{i=1}^{p} \varphi_{ij} x_{t-i} + \varphi_{0j} \right) \\
& \left\{ x_{t-i} = \left(1, x_{t-1}, x_{t-2}, \cdots, x_{t-p}\right)', i = 1, \ldots, p \right\} \qquad (1) \\
& \left\{ \varphi_{ij}, i = 1, \ldots, p, j = 1, \cdots, q \right\} \\
& \left\{ \beta_j, j = 1, \ldots, q \right\}
\end{aligned}
$$

Where y_t is the output vector of the MLP at time t; g is the nonlinear function of the neurons in the hidden layer; x_{t-i} is the input value at time $t-i$ where i stands for the memory (the number of lags that are used to introduce the context of the actual observation.); q is the number of neurons in the hidden layer; φ_{ij} are the weights of neuron j connecting the input with the hidden layer; and β_j are the weights connecting the output of the neuron j at the hidden layer with the output neuron. Note that the output y_t in our study is the estimate of the value of the time series at time $t+1$, while the input vector to the neural network will have a dimensionality of $p+1$.

We have considered a MLP$(p;q)$ architecture that represents the possible nonlinear relationship between the input vector and the output vector. Once the topology of the neural network is decided (i.e. the number of layers, etc.), the parameters of the network are estimated. The estimation can be done by means of different algorithms, which are either based on gradient search or line search. A summary of the different algorithms can be found in Bishop (1995). Another aspect to be taken into account, is the fact that the training is done by iteratively estimating the value of the parameters by local improvements of the cost function. To avoid the possibility that the search for the optimum value of the parameters finishes in a local minimum, we have used a multistarting technique that initializes the neural network several times for different initial random values and returns the best result.

3.2. *Radial basis function (RBF) neural network*

RBF networks consist of a linear combination of radial basis functions centered at a set of centroids with a given spread that controls the volume of the input space represented by a neuron (Bishop, 1995). RBF networks typically include three layers: an input layer; a hidden layer, which consists of a set of neurons, each of them computing a symmetric radial function; and an output layer that consists of a set of neurons, one for each given output, linearly combining the outputs of the hidden layer. The input can be modeled as a feature vector of real numbers, and the hidden layer is formed by a set of radial functions centered each at a centroid μ_j. The output of the network is a scalar function of the output vector of the hidden layer. The equations that describe the input/output relationship of the RBF are:

$$
\begin{aligned}
& y_t = \beta_0 + \sum_{j=1}^{q} \beta_j g_j(x_{t-i}) \\
& g_j(x_{t-i}) = \exp\left(- \frac{\sum_{j=1}^{p} (x_{t-i} - \mu_j)^2}{2\sigma_j^2} \right) \\
& \left\{ x_{t-i} = (1, x_{t-1}, x_{t-2}, \ldots, x_{t-p})', i = 1, \ldots, p \right\} \\
& \left\{ \beta_j, j = 1, \ldots, q \right\}
\end{aligned}
\tag{2}
$$

Where y_t is the output vector of the RBF at time t; β_j are the weights connecting the output of the neuron j at the hidden layer with the output neuron; q is the number of neurons in the hidden layer; g_j is the activation function, which usually has a Gaussian shape; x_{t-i} is the input value at time $t-i$ where i stands for the memory (the number of lags that are used to introduce the context of the actual observation); μ_j is the centroid vector for neuron j; and the spread σ_j is a scalar that measures the width over the input space of the Gaussian function and it can be defined as the area of influence of neuron j in the space of the inputs. Note that the output y_t in our study is the estimate of the value of the time series at time $t+1$, while the input vector to the neural network will have a dimensionality of $p+1$.

In order to assure a correct performance, before the training phase the number of centroids and the spread of each centroids have to be selected. There are different methods for the estimation of the number of centroids and the spread of the network. A complete summary can be found in Haykin (1999). In this study the training was done by adding the centroids iteratively with the spread parameter fixed. Then a regularized linear regression was estimated to compute the connections between the hidden and the output layer. Finally, the performance of the network was computed on the validation data set. This process was repeated until the performance on the validation database ceased to decrease.

3.3. Elman neural network

An Elman network is a special architecture of the class of recurrent neural networks. The architecture is based on a three-layer network with the addition of a set of context units that allow feedback on the internal activation of the network. There are connections from the hidden layer to these context units fixed with a weight of one. At each time step, the input is propagated in a standard feed-forward fashion, and then a back-propagation type of learning rule is applied. The fixed back connections result in the context units always maintaining a copy of the previous values of the hidden units. Thus the network can maintain a sort of state of the past decisions made by the hidden units, allowing it to perform such tasks as sequence-prediction that are beyond the power of a standard multilayer perceptron. The output of the network is a scalar function of the output vector of the hidden layer:

$$
\begin{aligned}
& y_t = \beta_0 + \sum_{j=1}^{q} \beta_j z_{j,t} \\
& z_{j,t} = g\left(\sum_{i=1}^{p} \varphi_{ij} x_{t-i} + \varphi_{0j} + \delta_{ij} z_{j,t-1} \right) \\
& \left\{ x_{t-i} = \left(1, x_{t-1}, x_{t-2}, \cdots, x_{t-p}\right)', i = 1, \ldots, p \right\} \\
& \left\{ \varphi_{ij}, i = 1, \ldots, p, j = 1, \ldots, q \right\} \\
& \left\{ \beta_j, j = 1, \ldots, q \right\} \\
& \left\{ \delta_{ij}, i = 1, \ldots, p, j = 1, \ldots, q \right\}
\end{aligned}
\qquad (3)
$$

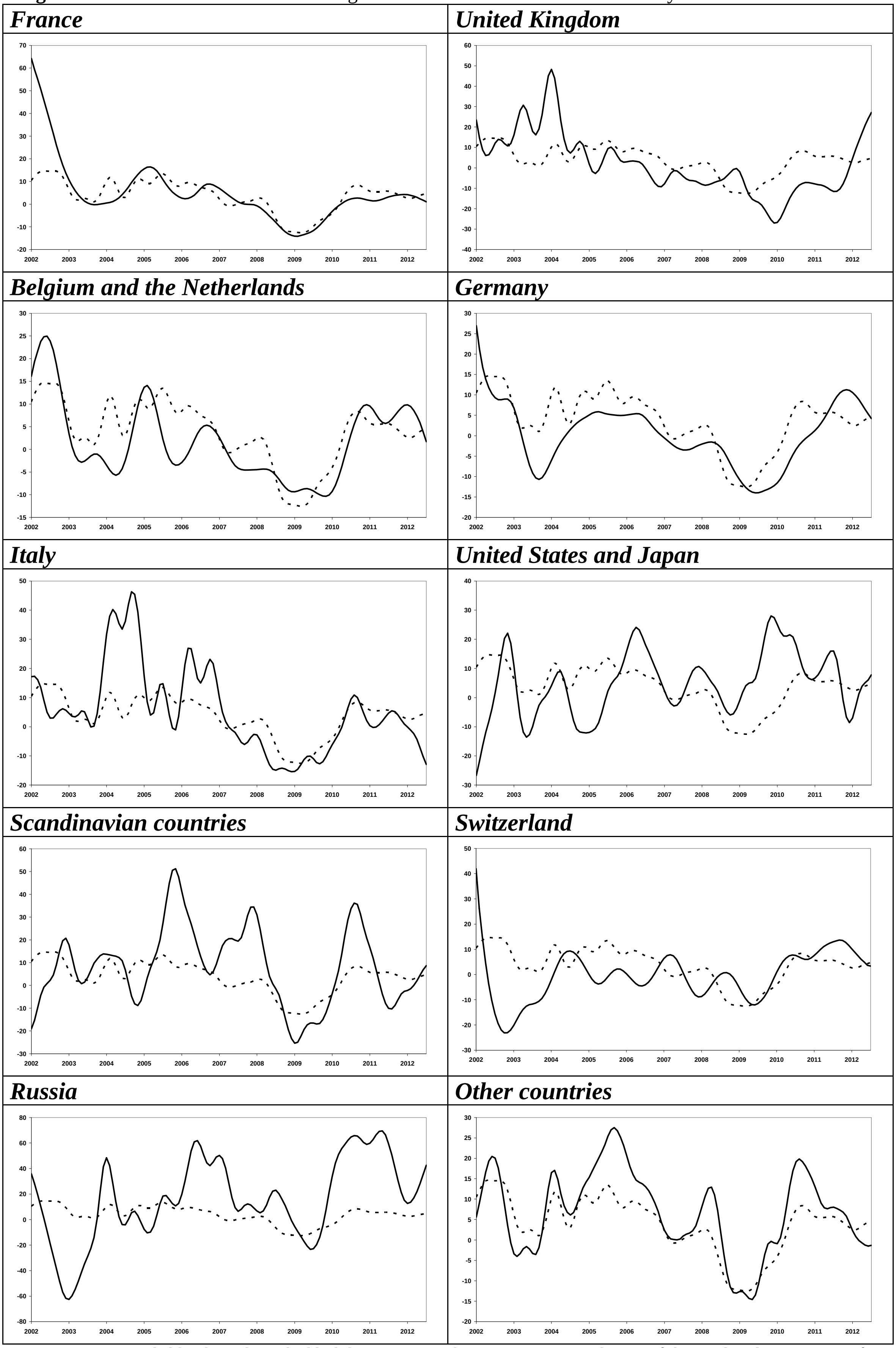
Where y_t is the output vector of the Elman network at time t; $z_{j,t}$ is the output of the hidden layer neuron j at the moment t; g is the nonlinear function of the neurons in the hidden layer; x_{t-i} is the input value at time $t-i$ where i stands for the memory (the number of lags that are used to introduce the context of the actual observation); φ_{ij} are the weights of neuron j connecting the input with the hidden layer; q is the number of neurons in the hidden layer; β_j are the weights of neuron j that link the hidden layer with the output; and δ_{ij} are the weights that correspond to the output layer and connect the activation at moment t. Note that the output y_t in our study is the estimate of the value of the time series at time $t+1$, while the input vector to the neural network will have a dimensionality of $p+1$.

There are different strategies for estimating the parameters of the Elman neural network. In this study, the training of the network was done by back-propagation through time, which is a generalization of back-propagation for feed-forward networks. The parameters of the Elman neural network are estimated by minimizing an error cost function. In order to minimize total error, we use gradient descent. A potential problem with gradient descent for standard recurrent architectures is that error gradients vanish exponentially quickly with the size of the time lag. Therefore recurrent NN cannot be easily trained for large numbers of neuron units.

4. Data

Data on tourists arrivals (first destinations) are provided by the Institute of Tourism Studies (IET) and are available at the Statistical Institute of Catalonia (IDESCAT). Data include the monthly number of tourists arriving from each visitor market over the time period 2001:01 to 2012:07. Table 1 shows a descriptive analysis of the data. It can be seen that the first four visitor markets (France, the United Kingdom, Belgium and the Netherlands and Germany) account for more than half of the total number of tourist arrivals to Catalonia. Nevertheless, when comparing the growth rates (Fig. 1), Russia and the Scandinavian countries experienced the highest growth in tourist arrivals during this period. Russia is also the country that presents the highest relative dispersion and the highest levels of Skewness and Kurtosis, while the United Kingdom shows the lowest levels of Skewness and Kurtosis.

Figure 1. Growth rates of tourists coming to Catalonia: from each visitor country vs. total arrivals



1. Source: Compiled by the author. The black line represents the year-on-year growth rates of the trend-cycle component of tourist arrivals to Catalonia from each visitor country. The dotted line represents the year-on-year growth rates of the trend-cycle component of total inbound tourism demand to Catalonia.

Table 1. Descriptive analysis of tourist arrivals (levels)

Country	Minimum	Maximum	Mean	Standard deviation	Skewness	Kurtosis
France	59,886	869,535	300,137	161,364	1.22	4.35
United Kingdom	34,128	293,005	152,223	70,762	0.10	1.78
Belgium and NL	23,818	467,505	118,974	100,198	1.74	5.61
Germany	26,588	258,600	112,126	53,834	0.37	2.26
Italy	24,077	271,975	83,805	42,335	1.96	7.76
US and Japan	20,984	131,089	60,795	22,869	0.80	3.53
Scandinavian countries	7,439	99,879	38,155	19,790	0.74	3.27
Switzerland	8,867	98,924	28,120	14,173	1.42	6.83
Russia	1,687	162,505	23,486	27,998	2.38	9.64
Other countries	101,894	442,597	246,241	76,311	0.36	2.38
Total	360,281	2,302,855	1,164,061	496,152	0.55	2.45

We use the year-on-year rates of the seasonally adjusted series to eliminate both linear trends as well as seasonality. These series are obtained using a Census X12 filter. In Fig.1 we compare the growth rate of the seasonally adjusted series of tourists coming to Catalonia from each visitor country to the growth rate of total inbound international tourism demand. Given the common patterns displayed by most countries we test for multicointegration using Johansen's (1988, 1991) maximum eigenvalue test. The maximum eigenvalue test tests the null hypothesis of r cointegrating vectors against the alternative hypothesis of $r+1$ cointegrating vectors.

In Table 2 we present the results of the unrestricted cointegration eigenvalue test. It can be seen that all different markets present correlated accelerations. The fact that the evolution of tourist arrivals is multicointegrated led us to use the correlations in the evolution of tourist arrivals between all different visitor markets. To forecast tourism demand, we design a MIMO setting in which forecasts of tourist arrivals for all countries are obtained simultaneously, and we compare the results to those of a SISO approach, in which models are estimated country by country.

Table 2. Cointegration test results. Unrestricted Cointegration Rank Test – Maximum eigenvalue

Hypothesized number of CE(s)	Type of test			
	Allow for linear deterministic trend in data			
	Intercept in CE		Intercept in CE	
	Test VAR		No trend in VAR	
	Maximum Eigenvalue	Critical value	Maximum Eigenvalue	Critical value
$H_0 : r = 0$ *	227.2916	64.50472	227.4935	68.81206
$H_0 : r \leq 1$ *	152.9724	58.43354	181.3408	62.75215
$H_0 : r \leq 2$ *	133.6029	52.36261	134.5977	56.70519
$H_0 : r \leq 3$ *	105.6646	46.23142	129.6588	50.59985
$H_0 : r \leq 4$ *	86.6518	40.07757	97.79509	44.4972
$H_0 : r \leq 5$ *	77.79057	33.87687	86.65054	38.33101
$H_0 : r \leq 6$ *	65.28306	27.58434	77.78193	32.11832
$H_0 : r \leq 7$ *	49.773	21.13162	64.52919	25.82321
$H_0 : r \leq 8$ *	36.80542	14.2646	49.7264	19.38704
$H_0 : r \leq 9$ *	10.98843	3.841466	35.64879	12.51798

1. Estimation period 2001:01-2012:07.
2. * Denotes rejection of the hypothesis at the 0.05 level **MacKinnon-Haug-Michelis (1999) p-values.

5.Empirical results

We carry out an out-of-sample forecasting competition between three different ANN architectures (MLP, RBF and Elman) using both a SIO and a MIMO setting. While a multiple-output approach allows to simultaneously obtain forecasts for each visitor market, a single-output approach requires to implement the experiment for each visitor country. The single-output approach is implemented by reconfiguring the architecture of the multiple-output ANNs into an array of single-output networks for each country. A multivariate approach seems especially suited for this specific data set in which growth rates of tourist arrivals from all the different countries of origin share a common stochastic trend (Table 2).

Following Bishop (1995) and Ripley (1996), we divided the collected data into three sets: training, validation and test sets. This division is done in order to asses the performance of the network on unseen data. Based on these considerations, the first sixty monthly observations (from January 2001 to January 2006) are selected as the initial training set, the next thirty-six (from January 2007 to January 2009) as the validation set and the last 20% as the test set.

Due to the large number of possible networks' configurations, the validation set is used for determining the following aspects of the neural networks:

a. The topology of the networks.

b. The number of epocs for the training of the MLP neural networks. The iterations in the gradient search are stopped when the error on the validation set increases.

c. The number of neurons in the hidden layer for the RBF. The sequential increase in the number of neurons at the hidden layer is stopped when the error on the validation increases.

d- The value of the spread σ_j in the RBF NN.

To make the system robust to local minima, we apply the multistartings technique, which consists on repeating each training phase several times. We repeat the training three times so as to obtain a low value of the performance error. The selection criterion for the topology and the parameters is the performance on the validation set. The results that are presented correspond to the selection of the best topology, the best spread in the case of the RBF neural networks, and the best training strategy in the case of the Elman neural networks.

To summarize the results of the out-of-sample competition and rank the methods according to their forecasting performance for different forecast horizons (1, 3 and 6 months) we compute the Mean Absolute Error (MAE) statistic for forecast accuracy. The results of our forecasting out-of-sample competition are shown in Tables 3 and 4. We also apply the Diebold-Mariano test (Table 5) for significant differences between each two competing series (single vs. multiple-output) for each forecast horizons in order to assess the value of the different models and settings.

We repeat the experiment assuming different topologies regarding the memory values. These values represent the number of lags introduced when running the models, denoting the number of previous months used for concatenation. The number of lags used in the different experiments ranged from one to three months for all the networks architectures. Therefore, when the memory is zero, the forecast is done using only the current value of the time series, without any additional temporal context.

Table 3. MAE (2010:04-2012:02). Memory (0) – No additional lags

	SISO ANN models			MIMO ANN models		
France	MLP	RBF	Elman	MLP	RBF	Elman
1 month	0.42	*0.38**	19.49	4.31	4.60	20.32
3 months	2.72	1.26	16.00	7.89	2.04	30.99
6 months	5.40	2.92	12.66	6.44	*1.48**	22.11
United Kingdom						
1 month	*2.77*	5.15	17.40	8.60	*4.58*	24.66
3 months	8.48	7.48	15.59	22.59	9.85	33.27
6 months	17.22	8.54	13.38	16.77	11.68	23.41
Belgium and the NL						
1 month	7.96	5.86	15.52	4.19	*3.89*	14.43
3 months	5.46	*3.29*	13.72	6.96	6.63	15.37
6 months	9.86	4.02	11.91	10.49	8.05	12.39
Germany						
1 month	7.95	7.48	15.03	*2.85*	4.77	10.43
3 months	5.07	4.12	16.96	5.34	5.81	13.82
6 months	5.68	*3.36*	9.25	7.71	6.41	11.34
Italy						
1 month	*1.45*	1.60	10.12	15.49	4.33	20.37
3 months	4.11	4.31	14.12	19.79	4.48	25.01
6 months	7.80	8.88	13.53	25.27	*3.96*	32.49
US and Japan						
1 month	5.12	*4.09*	12.94	8.45	*6.78*	17.50
3 months	8.28	7.62	20.39	14.02	10.00	19.29
6 months	10.01	9.78	13.79	15.90	10.06	19.40
Scandinavian countries						
1 month	4.10	*3.90*	18.84	16.36	*6.26*	26.34
3 months	9.85	8.99	16.70	26.95	14.15	30.29
6 months	13.38	12.75	23.33	30.20	14.42	34.02
Switzerland						
1 month	11.49	10.63	21.44	*5.38*	6.00	17.03
3 months	6.81	5.27	11.94	9.21	9.92	13.56
6 months	7.26	*5.05*	22.77	12.20	10.00	16.82
Russia						
1 month	29.74	*26.96*	34.59	23.39	*13.45*	41.46
3 months	34.47	29.33	32.57	39.12	35.81	48.67
6 months	35.39	33.68	49.63	50.01	43.01	60.74
Other countries						
1 month	2.64	*2.44*	11.11	9.73	*4.09*	13.51
3 months	5.88	4.59	13.24	14.40	4.98	17.93
6 months	8.02	6.92	12.03	17.27	5.90	18.35
Total						
1 month	*3.27*	3.41	15.64	6.64	*2.52*	8.25
3 months	5.98	3.75	13.37	10.49	2.53	11.07
6 months	14.72	3.45	10.88	8.68	2.67	9.78

1. *Italics:* best model for each country
2. * Best model

Table 4. MAE (2010:04-2012:02). Memory (3) – 3 additional lags

	SISO ANN models			MIMO ANN models		
France	MLP	RBF	Elman	MLP	RBF	Elman
1 month	*0.08**	0.21	14.99	8.19	3.43	19.67
3 months	1.27	1.12	16.40	6.06	*1.79**	19.31
6 months	4.56	4.32	10.90	5.82	1.83	14.60
United Kingdom						
1 month	*3.75*	4.53	10.20	16.28	*8.83*	23.00
3 months	5.86	6.81	9.46	23.42	11.84	21.20
6 months	9.28	11.94	14.83	15.40	11.96	35.57
Belgium and the NL						
1 month	8.38	8.21	13.87	6.25	*5.69*	9.97
3 months	6.92	7.04	11.89	7.83	7.43	14.19
6 months	12.10	*4.95*	10.10	11.31	7.40	13.91
Germany						
1 month	9.24	8.59	14.56	*3.64*	5.76	9.63
3 months	7.59	7.81	11.18	7.94	6.20	13.25
6 months	8.28	*7.31*	12.59	7.91	6.23	10.90
Italy						
1 month	*0.83*	1.56	11.69	18.46	5.78	18.03
3 months	5.44	3.63	13.41	23.55	*5.37*	18.93
6 months	11.09	8.70	11.74	19.07	6.69	23.19
US and Japan						
1 month	*4.71*	5.00	16.06	10.81	*9.28*	15.51
3 months	6.78	9.33	17.65	13.82	9.48	18.12
6 months	8.55	9.55	9.67	20.80	10.42	16.88
Scandinavian countries						
1 month	*3.08*	3.12	14.30	19.22	*12.45*	25.23
3 months	3.78	6.41	14.08	26.29	15.12	23.34
6 months	10.15	8.98	23.52	32.94	15.51	42.83
Switzerland						
1 month	14.58	11.00	9.95	*7.26*	8.01	12.75
3 months	14.97	9.84	14.94	11.37	10.43	11.73
6 months	8.55	*5.90*	12.09	7.58	10.91	19.44
Russia						
1 month	24.53	26.51	33.46	*25.02*	33.61	45.65
3 months	*23.18*	25.56	28.87	32.28	41.04	45.80
6 months	33.21	37.87	51.99	58.46	41.17	47.07
Other countries						
1 month	2.60	2.64	12.52	11.03	6.52	11.56
3 months	2.75	*2.34*	13.94	15.18	*6.13*	16.47
6 months	5.57	5.54	16.97	13.44	6.33	16.41
Total						
1 month	3.57	*3.44*	12.19	4.99	2.65	10.51
3 months	4.47	3.99	12.25	6.10	*2.38*	10.27
6 months	8.71	9.46	11.11	7.35	2.41	12.16

1. *Italics:* best model for each country
2. * Best model

When comparing the forecasting performance of the different neural architectures, RBF networks show lower MAE values than MLP and Elman networks, specially when no additional lags are introduced (Table 3). When the forecasts are obtained incorporating additional lags of the time series (Table 4), the forecasting performance of MLP networks improves for shorter horizons in the SISO approach. This result indicates that the number of previous months used for concatenation, conditions the forecasting performance of the different networks, although not in a significant way. An explanation for the better forecasting performance of RBF networks has to do with the fact that in this type of architecture, data are clusterized. On the other extreme, Elman networks systematically obtain the highest MAE values. This result suggests that the feedback topology of the Elman network could not capture the specificities of the time series. The fact that the number of training epocs had to be low in order to maintain the stability of the network suggests that the Elman architecture requires longer time series.

When analyzing the differences between countries, the lowest MAE value is always obtained for France, while Russia displays the highest MAE values for all models and scenarios. These results can be explained by the fact that France is the main visitor market, while Russian visitors only account for a small percentage of total arrivals and present high levels of dispersion. Countries could be grouped regarding the evolution of the forecasting performance as the forecasting horizon increases: while France, Germany and Switzerland show low MAE values for 6 months forecasts, forecasts for Scandinavian countries, Italy, UK, US and Japan worsen as the forecasting horizon increases. These clusters can be explained by the common patterns observed in the evolution of tourism demand for certain groups of countries (Fig. 1).

When testing for significant differences between a MIMO and a SISO approach for each two competing series (Table 5), we find that the multivariate analysis does not outperform the approach country by country. On the contrary, 83% of the cases in which there is a significant difference between single and multiple-output approaches (half of the 198 cases), the sign is negative, indicating that the MIMO structures present higher forecasting errors. Nevertheless for short horizons, we find that for Germany, Switzerland, Russia and Belgium and the Netherlands the MIMO approach presents significantly better results. For total arrivals, MAE values are lower for RBF and Elman networks with the multivariate approach, but the differences are not statistically significant.

Table 5. Diebold-Mariano loss-differential test statistic for predictive accuracy

	Memory (0) – no additional lags			Memory (3) – 3 additional lags		
	MLP	RBF	Elman	MLP	RBF	Elman
	Single vs. Multiple-output	Single vs. Multiple-output	Single vs. Multiple-output	Single vs. Multiple-output	Single vs. Multiple-output	Single vs. Multiple-output
France						
1 month	*-6.58*	*-5.79*	-0.14	*-6.64*	*-7.10*	-1.16
3 months	*-4.03*	*-2.95*	*-2.46*	*-4.50*	*-3.54*	-1.04
6 months	-0.47	*2.82*	-1.87	-0.74	*3.95*	-1.17
United Kingdom						
1 month	*-5.45*	0.62	*-2.04*	*-5.04*	*-3.07*	*-7.31*
3 months	*-2.99*	*-2.97*	*-2.16*	*-6.41*	*-4.34*	*-3.66*
6 months	0.08	*-2.50*	*-3.08*	*-2.77*	-0.02	*-3.56*
Belgium and the NL						
1 month	*2.41*	*2.12*	*2.12*	1.91	0.36	*2.10*
3 months	-0.86	*-2.92*	*-2.92*	-0.62	-0.50	-1.09
6 months	-0.19	*-2.97*	*-2.97*	0.18	-0.13	-1.55
Germany						
1 month	*3.46*	*2.10*	1.64	*3.92*	*2.45*	1.89
3 months	-0.23	*-2.63*	0.73	-0.21	1.28	-1.01
6 months	-0.99	-1.71	-0.93	0.28	0.73	0.45
Italy						
1 month	*-5.03*	*-4.86*	*-2.50*	*-6.46*	*-3.99*	-1.66
3 months	*-4.85*	-0.18	*-2.88*	*-4.64*	-1.44	-1.88
6 months	*-3.96*	*2.56*	*-3.54*	*-2.10*	0.86	*-2.75*
US and Japan						
1 month	*-2.77*	-1.33	-1.71	*-3.03*	*-3.27*	0.32
3 months	-1.55	-0.93	0.36	*-2.90*	-0.10	-0.13
6 months	*-2.88*	-0.13	-1.12	*-2.15*	-0.53	*-4.05*
Scandinavian countries						
1 month	*-4.07*	-1.57	-1.28	*-3.01*	*-4.66*	*-2.88*
3 months	*-2.35*	*-2.58*	*-4.27*	*-6.12*	*-2.78*	-1.95
6 months	*-3.74*	-0.74	*-2.09*	*-3.52*	*-2.55*	*-3.48*
Switzerland						
1 month	*4.06*	*3.52*	0.81	*4.82*	1.42	-1.45
3 months	-1.27	*-3.02*	-0.29	1.92	-0.40	0.94
6 months	-1.99	*-9.98*	1.53	0.57	*-8.26*	*-2.54*
Russia						
1 month	1.08	*3.13*	-0.91	-0.08	-1.34	-1.46
3 months	-0.64	*-2.17*	-1.65	-1.62	*-5.20*	*-3.34*
6 months	-1.59	*-2.31*	-1.04	-1.63	-1.04	0.40
Other countries						
1 month	*-3.81*	*-2.42*	-0.85	*-5.96*	*-3.99*	0.37
3 months	*-4.10*	-0.28	*-2.22*	*-5.96*	*-2.29*	-0.91
6 months	*-3.86*	1.08	*-2.21*	*-2.24*	-0.80	0.12
Total						
1 month	*-3.84*	1.05	*4.38*	-1.50	1.15	0.87
3 months	*-2.50*	1.42	0.88	-1.27	1.63	0.88
6 months	1.46	0.95	0.67	0.47	*2.62*	-0.51

1. Diebold-Mariano test statistic with NW estimator. Null hypothesis: the difference between the two competing series is non-significant. A negative sign of the statistic implies that the second model has bigger forecasting errors.
2. *Italics:* Significant at the 5% level (2.028 critical value).

6. Summary and Conclusions

The increasing importance of the tourism sector worldwide has led to a growing interest in new approaches to tourism demand forecasting. New methods provide more accurate estimations of anticipated tourist arrivals for effective policy planning. Artificial intelligence techniques such as Artificial Neural Networks have attracted increasing interest to refine the predictions of tourist arrivals at the destination level. From the wide array of neural network models, we have focused on three different architectures that represent three alternative ways of handling information: the multi-layer perceptron neural network, the radial basis function neural network and the Elman recursive neural network.

The main purpose of this study is to assess how forecasts of tourism demand can be improved by incorporating the existing common trends in tourist arrivals form all visitor markets to a specific destination. Given that the evolution of tourist arrivals form original countries to Catalonia presents a significant cross-correlation structure, we have analyzed whether a multivariate approach that takes into account the correlations in the evolution of tourist arrivals from different countries of origin has a significant effect on forecast accuracy. With this aim we have compared the performance of three different ANN topologies in a multiple-input multiple-output setting to that obtained estimating the models country by country.

When comparing the forecasting accuracy of univariate versus multivariate models country by country, we obtain better forecasting results with an univariate approach, with the exception of forecasts for short forecasting horizons in very few countries (Germany, Switzerland, Russia and Belgium and the Netherlands). Nevertheless, for total tourist arrivals we obtain lower forecasting errors with a multivariate approach. This result shows that a multiple-input multiple-output structure proves useful to forecast the inbound international demand to a destination when the evolution of tourist arrivals form all visitor markets share a common trend.

When comparing the forecasting accuracy of the different techniques, we find that radial basis function neural networks outperform multi-layer perceptron and Elman neural networks, being the Elman model the one showing the poorest forecasting performance. This result suggests that issues related with the divergence of the Elman neural network may arise when using dynamic networks with forecasting purposes.

Recurrent neural networks are not easy to train for large numbers of input units and may present scaling issues. These results reveal the suitability of hybrid models such as radial basis functions for tourism demand forecasting.

When analyzing the differences between countries, France displays the best forecasting results. On the other hand, we obtain the worst forecasting results with all models and in all scenarios for the predictions about the evolution of Russian tourists. These results can partly be explained by the fact that Russian visitors only account for a small percentage of total arrivals and show high levels of dispersion. As it could be expected, forecasts for Scandinavian countries, Italy, UK, US and Japan worsen as the forecasting horizon increases, while in France, Germany and Switzerland we obtain low forecasting errors for longer term forecasts. The distance to the destination could be explaining the differences between each group of countries.

In order to evaluate the effect of the memory on the forecasting results, we repeated the experiment considering different topologies regarding the number of lags used for concatenation. We find no significant differences when additional lags are incorporated in the feature vector. The fact that increasing the dimensionality of the input does not have a significant effect on forecast accuracy is indicative that the increase in the weight matrix is not compensated by the more complex specification, leading to overparametrization. This issue could be solved by increasing the length of the time series of tourist arrivals. Longer time series would also favor the learning process of the neural networks.

This study contributes to the tourism forecasting literature and to the tourism industry by highlighting the relevance of using the common trends in tourist arrivals from different visitor markets and the suitability of applying radial basis function neural networks to improve the forecasting accuracy of international inbound tourism demand. The proposed forecasting approach may prove useful for planning purposes, providing managers with a new and practical forecasting approach. Nevertheless, this study is not without its limitations. First, a comparison between different tourist destinations would allow to analyze whether regional differences have a significant influence on forecasting accuracy. Another question to be considered in further research is whether the implementation of supervised learning models such as support vector regressions, or the combination of the forecasts of different topologies and different time aggregations, may improve the forecasting performance of practical neural network-based tourism demand forecasting.

Acknowledgements

This work was supported by the *Ministerio de Economía y Competitividad* under the SpeechTech4All Grant (TEC2012-38939-C03-02).

References

Aguiló, E., & Rosselló, J. (2005). Host Community Perceptions. A Cluster analysis. *Annals of Tourism Research*, 32, 925-941. doi: 10.1016/j.annals.2004.11.004

Akkemik, K. A. (2012). Assessing the Importance of International Tourism for the Turkish Economy: A Social Accounting Matrix Analysis. *Tourism Management*, 33, 790-801. doi:10.1016/j.tourman.2011.09.002

Alon, I., Qi, M., & Sadowski, R. J. (2001). Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, 8, 147-156. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.198.7140&rep=rep1&type=pdf

Andrades-Caldito, L., Sánchez-Rivero, M. & Pulido-Fernández, J. I. (2013). Differentiating Competitiveness through Tourism Image Assessment. An Application to Andalusia (Spain). *Journal of Travel Research*, 52, 68-81. doi: 10.1177/0047287512451135

Assaf, A. G., Barros, C. P. & Gil-Alana, L. A. (2011). Persistence in the Short- and Long-Term Tourist Arrivals to Australia. *Journal of Travel Research*, 50, 213-229. doi: 10.1177/0047287510362787

Athanasopoulos, G., & de Silva, A. (2012). Multivariate Exponential Smoothing for Forecasting Tourist Arrivals. *Journal of Travel Research*, 51, 640-652. doi:10.1177/0047287511434115

Balaguer, J., & Cantavella-Jordá, M. (2002). Tourism as a Long-Run Economic Growth Factor: The Spanish Case. *Applied Economics*, 34, 877-884. doi:10.1080/00036840110058923

Bardolet, E., & Sheldon, P. J. (2008). Tourism in Archipelagos. Hawai'i and the Balearics". *Annals of Tourism Research*, 35, 900-923. doi:10.1016/j.annals.2008.07.005

Bermúdez, J. D., Corberán-Vallet, A. & Vercher, E. (2009). Multivariate Exponential Smoothing: A Bayesian Forecast Approach Based on Simulation. *Mathematics and Computers in Simulation*, 79, 1761-1769. doi:10.1016/j.matcom.2008.09.004

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Bloom, J. (2005). Market Segmentation. A Neural Network Application". *Annals of Tourism Research*, 32, 93-111. doi:10.1016/j.annals.2004.05.001

Broomhead, D. S., & Lowe, D. (1988). Multi-Variable Functional Interpolation and Adaptive Networks. *Complex Syst.*, 2, 321-355. Retrieved from https://www.complex-systems.com/pdf/02-3-5.pdf

Cang, S. (2013). A Comparative Analysis of Three Types of Tourism Demand Forecasting Models: Individual, Linear Combination and Non-Linear Combination, *International Journal of Tourism Research*, 15. doi:10.1002/jtr.1953

Cang, S., & Yu, H. (2014). A Combination Selection Algorithm on Forecasting, *European Journal of Operational Research*, 234, 127-139. doi:10.1016/j.ejor.2013.08.045

Capó, J., Riera, A. & Rosselló, J. (2007). Tourism and long-term growth. A Spanish perspective. *Annals of Tourism Research*, 34, 709-726. doi:10.1016/j.annals.2007.02.003

Chan, C. K., Witt, S. F., Lee, Y. C. E., & Song, H. (2010). Tourism Forecast Combination Using the CUSUM Technique. *Tourism Management*, 31, 891-897. doi:10.1016/j.tourman.2009.10.004

Chen, K. (2011). Combining Linear and Nonlinear Model in Forecasting Tourism Demand". *Expert Systems with Applications*, 38, 10368-10376. doi:10.1016/j.eswa.2011.02.049

Chen, K. Y., & Wang, C. H. (2007). Support Vector Regression with Genetic Algorithms in Forecasting Tourism Demand. *Tourism Management*, 28, 215-226. doi:10.1016/j.tourman.2005.12.018

Cho, V. (2003). A Comparison of Three Different Approaches to Tourist Arrival Forecasting. *Tourism Management*, 24, 323-330. doi:10.1016/S0261-5177(02)00068-7

Chu, F. (2008). A Fractionally Integrated Autoregressive Moving Average Approach to Forecasting Tourism Demand. *Tourism Management*, 29, 79-88. doi:10.1016/j.tourman.2007.04.003

Chu, F. (2011). A Piecewise Linear Approach to Modeling and Forecasting Demand for Macau Tourism. *Tourism Management*, 32, 1414-1420. doi:10.1016/j.tourman.2011.01.018

Claveria, O., & Datzira, J. (2009). Tourism Demand in Catalonia: Detecting External Economic Factors. *Tourismos*, 4, 13-28. Retrieved from http://www.chios.aegean.gr/tourism/vol4no1.pdf

Claveria, O., & Datzira, J. (2010). Forecasting Tourism Demand Using Consumer Expectations. *Tourism Review*, 65, 18-36. doi:10.1108/16605371011040889

Claveria, O., Monte, E., & Torra, S. (2014). Tourism Demand Forecasting with Neural Network Models: Different Ways of Treating Information. *International Journal of Tourism Research*. doi:10.1002/jtr.2016

Claveria, O., Monte, E., & Torra, S. (2016). A New Forecasting Approach for the Hospitality Industry. *International Journal of Contemporary Hospitality Management*, 28 (2) (forthcoming)

Claveria, O., & Torra, S. (2014). Forecasting Tourism Demand to Catalonia: Neural Networks vs. Time Series Models. *Economic Modelling*, 36, 220-228. doi:10.1016/j.econmod.2013.09.024

Cortés-Jiménez, I., & Blake, A. (2011). Tourism Demand Modeling by Purpose of Visit and Nationality. *Journal of Travel Research*, 50, 4408-416. doi:10.1177/0047287510363615

Coshall, J. T., & Charlesworth, R. (2010). A Management Orientated Approach to Combination Forecasting of Tourism Demand. *Tourism Management*, 32, 759-769. doi:10.1016/j.tourman.2010.06.011

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-263. doi:10.1080/07350015.1995.10524599

Du Preez, J., & Witt, S. F. (2003). Univariate versus Multivariate Time Series Forecasting: An Application to International Tourism Demand. *International Journal of Forecasting*, 19, 435-451. doi:10.1016/S0169-2070(02)00057-2

Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, 14, 179-211.

Garín-Muñoz, T., & Montero-Martín, L. F. (2007). Tourism in the Balearic Islands: A dynamic model for international demand using panel data. *Tourism Management*, 28, 1224-1235. doi:10.1016/j.tourman.2006.09.024

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

Goh, C., Law, R., & Mok, H. M. (2008). Analyzing and Forecasting Tourism Demand: A Rough Sets Approach. *Journal of Travel Research*, 46, 327-338. doi:10.1177/0047287506304047

Gounopoulos, D., Petmezas, D. & Santamaria, D. (2012). Forecasting tourist arrivals in Greece and the impact of macroeconomic shocks from the countries of tourists' origin. *Annals of Tourism Research*, 39, 641-666. doi:10.1016/j.annals.2011.09.001

Hadavandi, E., Ghanbari, A., Shahanaghi, K., & Abbasian, S. (2011). Tourist arrival forecasting by evolutionary fuzzy systems. *Tourism Management*, 32, 1196-1203. doi:10.1016/j.tourman.2010.09.015

Haviluddin, H., & Rayner, A. (2014). Daily Network Traffic Prediction Based on Backpropagation Neural Network . *Australian Journal of Basic and Applied Sciences*, 8, 164-169. Retrieved from http://www.ajbasweb.com/old/ajbas/2014/IPN%20JAKARTA/164-169.pdf

Haykin, S. (1999). *Neural Networks. A Comprehensive Foundation*. New Jersey: Prentice Hall.

Hong, W., Dong, Y., Chen, L. & Wei, S. (2011). SVR with Hybrid Chaotic Genetic Algorithms for Tourism Demand Forecasting. *Applied Soft Computing*, 11, 1881-1890. doi:10.1016/j.asoc.2010.06.003

Johansen, S. (1988). Statistical Analysis of Co-Integration Vectors. *Journal of Economic Dynamics and Control*, 12, 231-254. doi:10.1016/0165-1889(88)90041-3

Johansen, S. (1991). Estimation and Hypothesis Testing of Co-Integration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, 59, 1551-1580. doi: 10.2307/2938278

Kon, S., & Turner, L. L. (2005). Neural Network Forecasting of Tourism Demand. *Tourism Economics*, 11, 301-328. doi:10.5367/000000005774353006

Law, R. (2000). Back-Propagation Learning in Improving the Accuracy of Neural Network-Based Tourism Demand Forecasting. *Tourism Management*, 21, 331-340. doi:10.1016/S0261-5177(99)00067-9

Law, R. (2001). The Impact of the Asian Financial Crisis on Japanese Demand for Travel to Hong Kong: A Study of Various Forecasting Techniques. *Journal of Travel & Tourism Marketing*, 10, 47-66. doi:10.1080/10548400109511558

Li, G., Song, H., & Witt, S. F. (2005). Recent Developments in Econometric Modelling and Forecasting. *Journal of Travel Research*, 44, 83-99. doi: 10.1177/0047287505276594

Lin, C., Chen, H., & Lee, T. (2011). Forecasting Tourism Demand Using Time Series, Artificial Neural Networks and Multivariate Adaptive Regression Splines: Evidence from Taiwan. *International Journal of Business Administration*, 2, 14-24. doi:10.5430/ijba.v2n2p14

MacKinnon, J. G., Haug, A., & Michelis, L. (1999). Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. *Journal of Applied Econometrics*, 14, 563-577. doi:10.1002/(SICI)1099-1255(199909

Medeiros, M. C., McAleer, M., Slottje, D., Ramos, V., & Rey-Maquieira, J. (2008). An Alternative Approach to Estimating Demand: Neural Network Regression with Conditional Volatility for High Frequency Air Passenger Arrivals. *Journal of Econometrics*, 147, 372-383. doi:10.1016/j.jeconom.2008.09.018

Nawijn, J., & Mitas, O. (2012). Resident Attitudes to Tourism and Their Effect on Subjective Well-Being: The Case of Palma de Mallorca. *Journal of Travel Research*, 51, 531-541. doi:10.1177/004728751142648

Nordström, J. (2004). Estimating and Predicting International Tourism Demand in Sweden. *Scandinavian Journal of Hospitality and Tourism*, 4, 59-76. doi: 10.1080/15022250410006408

Page, S., Song, H., & Wu, D. C. (2012). Assessing the Impacts of the Global Economic Crisis and Swine Flu on Inbound Tourism Demand in the United Kingdom. *Journal of Travel Research*, 51, 142-153. doi: 10.1177/0047287511400754

Pai, P. & Hong, W. (2005). An Improved Neural Network Model in Forecasting Arrivals. *Annals of Tourism Research*, 32, 1138-1141.

Pai, P., Hung, K., & Lin, K. (2014). Tourism Demand Forecasting Using Novel Hybrid System. *Expert Systems with Applications*, 41, 3691-3702. doi:10.1016/j.eswa.2013.12.007

Palmer, A., Montaño, J. J., & Sesé, A. (2006). Designing an Artificial Neural Network for Forecasting Tourism Time-Series. *Tourism Management*, 27, 781-790. doi:10.1016/j.tourman.2005.05.006

Pattie, D. C., & Snyder, J. (1996). Using a Neural Network to Forecast Visitor Behavior. *Annals of Tourism Research*, 23, 151-164. doi:10.1016/0160-7383(95)00052-6

Peng, B., Song, H. & Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tourism Management*, 45, 181-193. doi:10.1016/j.tourman.2014.04.005

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.

Rosselló, J., Aguiló, E., & Riera, A. (2005). Modelling Tourism Demand Dynamics. *Journal of Travel Research,* 44, 111-116. doi: 10.1177/0047287505276602

Santana-Jiménez, Y., & Hernández, J. M. (2011). Estimating the Effect of Overcrowding on Tourist Attraction: The Case of Canary Islands. *Tourism Management*, 32, 415-425. doi:10.1016/j.tourman.2010.03.013

Shahrabi, J., Hadavandi, E., & Asadi, S. (2013). Developing a Hybrid Intelligent Model for Forecasting Problems: Case Study of Tourism Demand Time Series. *Knowledge-Based Systems*, 43, 112-122. doi:10.1016/j.knosys.2013.01.014

Shen, S., Li, G., & Song, H. (2008). An Assessment of Combining Tourism Demand Forecasts over Different Time Horizons. *Journal of Travel Research*, 47, 197-207. doi:10.1177/0047287508321199

Song, H., & Li, G. (2008). Tourism Demand Modelling and Forecasting – A Review of Recent Research". *Tourism Management*, 29, 203-220. doi:10.1016/j.tourman.2007.07.016

Teixeira, J. P., & Fernandes, P. O. (2012). Tourism Time Series Forecast – Different ANN Architectures with Time Index Input. *Procedia Technology*, 5, 445-454. doi:10.1016/j.protcy.2012.09.049

Tsaur, R., & Kuo, T. (2011). The Adaptive Fuzzy Time Series Model with an Application to Taiwan's Tourism Demand. *Expert Systems with Applications*, 38, 9164-9171. doi:10.1016/S0261-5177(01)00097-8

Tsaur, S., Y. Chiu, & C. Huang (2002). Determinants of Guest Loyalty to International Tourist Hotels: A Neural Network Approach. *Tourism Management*, 23, 397-405. doi:10.1016/S0261-5177(01)00097-8

Uysal, M., & El Roubi, M. S. (1999). Artificial Neural Networks versus Multiple Regression in Tourism Demand Analysis. *Journal of Travel Research*, 38, 111-118. doi:10.1177/004728759903800203

Uysal, M. (2004). Advancement in Computing: Implications for Tourism and Hospitality. *Scandinavian Journal of Hospitality and Tourism*, 38, 111-118. doi:10.1080/15022250410003834

Wu, Q., Law, R., & Xu, X. (2012). A Spare Gaussian Process Regression Model for Tourism Demand Forecasting in Hong Kong. *Expert Systems with Applications*, 39, 4769-4774. doi:10.1016/j.eswa.2011.09.159

Wu, Q., Li, R., & Song, X. (2012). Economic Analysis of Tourism Consumption Dynamics: A Time-varying Parameter Demand System Approach. *Annals of Tourism Research*, 39, 667-685. doi:10.1016/j.annals.2011.09.003

Yu, G., & Schwartz, Z. (2006). Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models. *Journal of Travel Research*, 45, 194-203. doi:10.1177/0047287506291594

Zhang, G. P., & M. Qi (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, 160, 501-514. doi:10.1016/j.ejor.2003.08.037

Institut de Recerca en Economia Aplicada Regional i Públic
Research Institute of Applied Economics

WEBSITE: www.ub-irea.com • **CONTACT:** irea@ub.edu

Grup de Recerca Anàlisi Quantitativa Regional
Regional Quantitative Analysis Research Group

WEBSITE: www.ub.edu/aqr/ • **CONTACT:** aqr@ub.edu

Universitat de Barcelona
Av. Diagonal, 690 • 08034 Barcelona