DATA PRE-PROCESSING FOR NEURAL NETWORK-BASED FORECASTING: DOES IT REALLY MATTER?

Oscar CLAVERIA[a], Enric MONTE[b], Salvador TORRA[c]

[a]AQR-IREA, Department of Econometrics and Statistics,
University of Barcelona, Barcelona, Spain
[b]Department of Signal Theory and Communications,
Polytechnic University of Catalunya, Barcelona, Spain
[c]Riskcenter-IREA, Department of Econometrics and Statistics,
University of Barcelona, Barcelona, Spain

Abstract. This study aims to analyze the effects of data pre-processing on the forecasting performance of neural network models. We use three different Artificial Neural Networks techniques to predict tourist demand: multi-layer perceptron, radial basis function and the Elman neural networks. The structure of the networks is based on a multiple-input multiple-output (MIMO) approach. We use official statistical data of inbound international tourism demand to Catalonia (Spain) and compare the forecasting accuracy of four processing methods for the input vector of the networks: levels, growth rates, seasonally adjusted levels and seasonally adjusted growth rates. When comparing the forecasting accuracy of the different inputs for each visitor market and for different forecasting horizons, we obtain significantly better forecasts with levels than with growth rates. We also find that seasonally adjusted series significantly improve the forecasting performance of the networks, which hints at the significance of deseasonalizing the time series when using neural networks with forecasting purposes. These results reveal that, when using seasonal data, neural networks performance can be significantly improved by working directly with seasonally adjusted levels.

Keywords: artificial neural networks, forecasting, multiple-input multiple-output (MIMO), seasonality, detrending, tourism demand, multilayer perceptron, radial basis function, Elman.

JEL Classification: L83, C53, C45, R11.

Introduction

International tourism has become one of today's fastest growing industries. Tourism accounts for almost 10% of total international trade and plays a fundamental role in the long-run economic development of many regions (Akkemik 2012; Sigala, Chalkiti 2014). To

---

Corresponding author Oscar Claveria
E-mail: *oclaveria@ub.edu*

achieve a sustainable tourism model, policy makers and professionals need more accurate predictions of the number of tourist arrivals at the destination level. Many authors have acknowledged the importance of applying new approaches to tourism demand forecasting in order to improve the accuracy of the methods of analysis (Song, Li 2008). The availability of more advanced forecasting techniques has led to a growing interest in Artificial Intelligence (AI) models (Yu, Schwartz 2006; Goh *et al.* 2008; Lin *et al.* 2011; Chen 2011; Celotto *et al.* 2012; Wu *et al.* 2012; Cang, Yu 2014) to the detriment of time series models (Chu 2008, 2011; Assaf *et al.* 2011) and causal econometric models (Page *et al.* 2012). Some of the new AI based techniques are fuzzy time series models (Tsaur, Kuo 2011), genetic algorithms (Hadavandi *et al.* 2011), expert systems (Shahrabi *et al.* 2013; Pai *et al.* 2014) and Support Vector Machines (SVMs) (Chen, Wang 2007; Hong *et al.* 2011). Recent research has shown the suitability of Artificial Neural Networks (ANNs) for dealing with tourism demand forecasting (Teixeira, Fernandes 2012; Claveria *et al.* 2015).

In spite of the successful use of ANNs for time series forecasting, very few studies compare the accuracy of different NN architectures for tourism demand forecasting at a regional level. The present study deals with tourist arrivals to Catalonia. 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 more than 15 million tourists in 2013, an 8% raise over the previous year. 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 tourism planning.

The fact that tourism data are characterised by strong seasonal patterns and volatility, make it a particularly interesting field in which to apply different types of ANN architectures. The raw time series of tourism data usually require significant pre-processing in order to be used for forecasting purposes. While the effects of data pre- processing on forecast accuracy have been widely studied in the context of time series analysis, there are very few studies of tourism demand with neural networks (Zhang, Qi 2005). Empirically, ANNs have shown to be suited to forecast nonlinear time series. Nevertheless, studies reach different conclusions on how to deal with seasonal time series (Hamzaçebi 2008). While Nelson *et al.* (1999) and Zhang and Kline (2007) concluded that in order to obtain a better ANN forecasting, the seasonal effect should be removed from the raw data, Franses and Draisma (1997) and Alon *et al.* (2001) found that ANNs are capable of modelling the seasonal and trend effects in data structure without removing the seasonal effects.

The objective of this study is to investigate the effects of data pre-processing in the forecast performance of ANNs when using seasonal time series, extending to tourist demand forecasting the results of previous research on economics. Given that univariate specifications are limited and unable to capture dynamic interrelationships between variables, we analyze whether a multivariate setting provides useful for forecasting purposes. With this aim, we implement a multiple-input multiple-output (MIMO) approach (i.e. each output corresponds to a given country) to predict international tourism demand to Catalonia (Spain) from all countries of origin. We compare the forecasting performance of three different ANN architectures: multi-layer perceptron (MLP), radial basis function

(RBF) and the Elman networks. To analyze the effects of data pre-processing on forecast accuracy for the different ANN architectures, we design the experiment using alternative approaches for data pre-processing: levels, growth rates, seasonally adjusted levels and seasonally adjusted growth rates. To assess the value of the different models we compute the Diebold-Mariano (DM) test for significant differences between each two competing series. To our knowledge, this is the first study to analyze the forecasting performance of MIMO ANNs for tourism data.

The structure of the paper proceeds as follows. Section 1 briefly reviews the literature on tourism demand forecasting with ANNs. In section 2, we present the different NN architectures used in the analysis. In the following section we explain how to design the experiment and implement the models. Data is presented in the 4th section, where the results of the out-of-sample forecasting competition are discussed. Finally, a summary and a discussion of the implications are given in the last Section.

1. Artificial Neural Networks in tourism demand forecasting

ANNs are models capable of identifying temporal patterns from historical data, capturing functional relationships among the data when the underlying process is unknown. The data generating process of tourist arrivals is too rich to be specified by a single linear algorithm, which might not be able to take into account saturation or exponential effects, interactions between different time series, etc. This explains the great interest that ANNs have aroused for tourism demand forecasting. Each type of network is suited to a combination of a learning paradigm, a learning rule and a learning algorithm (back-propagation, etc.). The main learning paradigms are supervised learning and non-supervised learning. In supervised learning, the 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 a distance criterion. MLP networks are supervised learning models, while RBF networks, combine both learning methods. 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.

ANNs can also be classified into feed-forward and recurrent networks regarding the connecting patterns of the different layers. The most widely used feed-forward topology in tourism demand forecasting is the MLP network (Law 2000; Tsaur *et al.* 2002; Kon, Turner 2005; Palmer *et al.* 2006). RBF networks are a special class of multi-layer feed-forward architecture with two layers of processing. In contrast to MLP networks, RBF networks are based in local approximations of the functions by means of centroids. Unlike feed-forward networks, recurrent neural networks are models that allow for a feedback of the past states of the network. While a feed-forward network propagates data linearly from input to output, recurrent networks also propagate data from later processing stages to earlier stages. A special class of recurrent network is the Elman network. Whilst MLP neural networks are increasingly used with forecasting purposes, RBF and the Elman neural networks have been scarcely used in tourism demand forecasting. Cang (2014) has recently combined MLP, RBF and SVM forecasts to predict UK inbound tourist arrivals. Cho (2003) used the Elman architecture to predict the number of arrivals from different countries to Hong Kong.

In recent years, several studies have been published on tourism in Spain at a regional level (Nawijn, Mitas 2012; Andrades-Caldito *et al.* 2013), but very few concerning tourism demand forecasting. Palmer *et al.* (2006) designed a MLP neural network to forecast tourism expenditure in the Balearic Islands. Medeiros *et al.* (2008) developed a NN-GARCH model to estimate demand for international tourism also in the Balearic Islands. Bermúdez *et al.* (2009) applied a multivariate exponential smoothing model, and by means of a Bayesian approach calculate prediction intervals for hotel occupancy in three provinces in Spain. Claveria and Torra (2014) compared the forecasting accuracy of time series models to that of MLP networks in Catalonia. In this study we analyze the effects of data pre-processing on the forecasting accuracy of three alternative ANN architectures.

2. Artificial Neural Network models for the forecasting competition

We use three ANN models: MLP, RBF and Elman networks. Equations (1), (2) and (3) respectively describe the input/output relationship of the three architectures:

MLP:

$$y_t = \beta_0 + \sum_{j=1}^{q} \beta_j g\left(\sum_{i=1}^{p} \varphi_{ij} x_{t-i} + \varphi_{0j} \right);$$ (1)

$$\left\{ x_{t-i} = \left(1, x_{t-1}, x_{t-2}, \cdots, x_{t-p} \right)', i = 1, \ldots, p \right\};$$

$$\left\{ \varphi_{ij}, i = 1, \ldots, p, j = 1, \cdots, q \right\};$$

$$\left\{ \beta_j, j = 1, \ldots, q \right\}.$$

RBF:

$$y_t = \beta_0 + \sum_{j=1}^{q} \beta_j g_j \left(x_{t-i} \right);$$

$$g_j \left(x_{t-i} \right) = \exp\left(-\frac{\sum_{j=1}^{p} \left(x_{t-i} - \mu_j \right)^2}{2\sigma_j^2} \right);$$ (2)

$$\left\{ x_{t-i} = \left(1, x_{t-1}, x_{t-2}, \ldots, x_{t-p} \right)', i = 1, \ldots, p \right\};$$

$$\left\{ \beta_j, j = 1, \ldots, q \right\}.$$

Elman:

$$y_t = \beta_0 + \sum_{j=1}^{q} \beta_j z_{j,t};$$

$$z_{j,t} = g\left(\sum_{i=1}^{p} \varphi_{ij} x_{t-i} + \varphi_{0j} + \delta_{ij} z_{j,t-1} \right);$$ (3)

$$\left\{ x_{t-i} = \left(1, x_{t-1}, x_{t-2}, \cdots, x_{t-p}\right)', i = 1, \ldots, p \right\};$$

$$\left\{ \varphi_{ij}, i = 1, \ldots, p, j = 1, \ldots, q \right\};$$

$$\left\{ \beta_j, j = 1, \ldots, q \right\};$$

$$\left\{ \delta_{ij}, i = 1, \ldots, p, j = 1, \ldots, q \right\}.$$

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 vector at time $t - i$, where i stands for the number of lags that are used to introduce the context of the actual observation (which is set to one in this study); 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. In the RBF specification g_j is the activation function, which usually has a Gaussian shape; 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. In the Elman network, $z_{j,t}$ is the output of the hidden layer neuron j at the moment t and δ_{ij} are the weights that correspond to the output layer and connect the activation at moment t. Further information about these three ANN architectures can be found in Bishop (1995) and Haykin (1999).

The models used for the forecasting comparison have two different kinds of parameters: the weights and the hyperparameters related to each topology (in the case of the RBF, the spread of each radial basis). The hyperparameters are determined by the performance of the networks on cross validation. The estimation of the parameters can be done by means of different algorithms, which are either based on gradient search, line search or quasi Newton search. In this paper we use a variant of the quasi Newton search called Leven-berg-Marquardt.

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 use a multi-starting technique that initializes the neural network several times for different initial random values and returns the best result on a validation database. The values in for the number of neurons in the hidden layer range from 5 to 25, and the margin for the spread of the radial basis from 0.1 to 2 with increments of 0.2. Note that the complexity of the search space is low, so we chose an enumeration strategy which finds the best combination. The specific values of these parameters depend on the forecasting horizon and the algorithm. As the forecasting horizon increases, the number of neurons needed in the hidden layer raises and varies between 10 and 20.

In order to assure a correct performance of RBF networks, the number of centroids and the spread of each centroid have to be selected before the training phase. In this study, the training is done by adding the centroids iteratively with the spread parameter fixed. Then a regularized linear regression is estimated to compute the connections between the hidden

and the output layers. Finally, the performance of the network is assessed on the validation data set. This process is repeated until the performance on the validation database ceases to decrease. The hyperparameter sigma is selected before determining the topology of the network and is tuned outside the training phase. The optimal value depends on the Euclidean distance that is computed inside each neuron. Values of the sigma parameter vary from 0.8 to 1.3 depending on the experiment, and are also dependent on the horizon of the forecast, possibly due to the uncertainty that arises when the forecast horizon increases to 6 months.

In the case of the Elman networks, the training is 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, gradient descent is used to change each weight in proportion to its derivative with respect to the error, provided the nonlinear activation functions are differentiable. A major problem with gradient descent for standard recurrent architectures is that error gradients vanish exponentially quickly with the size of the time lag.

3. Design of the experiment

By means of the Johansen test (see Table 1), we find a correlated evolution between all different visitor markets, which leads us to apply a MIMO approach to obtain forecasts of tourism demand for different forecast horizons. Given that univariate specifications are limited and unable to capture dynamic interrelationships between different countries of origin, we use a multivariate approach, in which information about all visitor markets to a destination is simultaneously used. This is the first study to analyze the forecasting performance of ANNs in a multivariate setting that allows incorporating cross-correlations between the evolutions of tourist arrivals from different countries to a specific destination.

We carry out an out-of-sample forecasting comparison between three different ANN architectures (MLP, RBF and Elman) using a MIMO setting. While a single-input single-output (SISO) approach requires implementing the experiment for each visitor country, a MIMO approach allows to simultaneously obtaining forecasts for each visitor market. A MIMO approach seems especially suited for this particular data set in which seasonal

Table 1. Number of cointegrating relations by model

	Type of model				
	Assume no deterministic trend in data		Allow for linear deterministic trend in data		Allow for quadratic deterministic trend in data
Test type	No intercept in CE	Intercept in CE	Intercept in CE	Intercept in CE	Intercept and trend in CE
	No test VAR	No intercept in VAR	Test VAR	No trend in VAR	Linear trend in VAR
Trace	9	10	9	10	10
Maximum Eigenvalue	9	10	9	10	10

Notes: Estimation period 2001:01–2012:07. Critical values based on MacKinnon *et al.* (1999).

adjusted levels of tourist arrivals from all the different countries of origin share a common stochastic trend.

Multivariate approaches to tourist demand forecasting are few and have yielded mixed results. Athanasopoulos and Silva (2012) compared the forecasting accuracy of exponential smoothing methods in a multivariate setting against univariate alternatives. They used international tourist arrivals to Australia and New Zealand and found that multivariate models improved on forecast accuracy over the univariate alternatives. Tsui *et al.* (2014) estimated multivariate ARIMA models (ARIMAX) with explanatory variables to forecast airport passenger traffic for Hong Kong and found similar forecasting results to SARIMA models. Contrary to what could be expected, Du Preez and Witt (2003) found that multivariate time series models did not generate more accurate forecasts than univariate time series models.

Following Bishop (1995) and Ripley (1996), we divide 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. The partition between train and test sets is done sequentially in order to maximize the size of the training database. As the prediction advances, forecasts are incorporated to the training database, successively increasing its size. The effect of this strategy is to improve the training of the network as the prediction advances, thus refining the performance at the end of the test phase.

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/Elman 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 multistarting technique, which consists on repeating each training phase several times. The benefits of the multistarting technique derive from the fact that it is a technique for solving the problem that arises from using a gradient search for optimizing the parameters of the neural network. The use of gradient search yields solutions that might be a local minimum of the cost function. These local minima consist of flat zones of the cost function where the gradient is near zero, and therefore the optimization algorithm stops the search. As the geometry of the cost functions is difficult to analyze, one strategy to overcome the problem is to try different initial values, selecting the result with the best performance on the validation database. In our case, 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. Forecasts for 1, 3 and 6 months ahead are computed in a recursive way. All neural networks are implemented using Matlab™ and its Neural Networks toolbox.

4. Results of the out-of-sample forecasting competition

In this section we implement a MIMO approach to predict arrivals to Catalonia from the different visitor countries. We use the number of tourists (first destinations) provided by 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. In Figure 1 we present the evolution of tourist arrivals to Catalonia. Tabulation of the data is presented in Table 2. It can be seen that the first four visitor markets (France, Belgium and the Netherlands, the United Kingdom and Germany) account for more than half of the total number of tourist arrivals to Catalonia.

We compare the forecasting performance of three different MIMO ANN architectures: MLP, RBF and the Elman recursive neural network. We repeat the experiment using alternative approaches for data pre-processing: levels, growth rates, seasonally adjusted levels and seasonally adjusted growth rates. Seasonally adjusted series are obtained using a Census X12 filter with a multiplicative decomposition. Forecasts for 1, 3 and 6 months ahead are generated in a recursive way (Pérez-Rodríguez *et al.* 2005). To summarize this information we compute the Mean Absolute Percentage Error (MAPE) statistic of forecast accuracy. The results of our multiple-step-ahead forecasting experiment are shown in Tables 3, 4 and 5. We have also used the DM test (Table 6) to obtain significant differences between each two competing series: levels vs. seasonally adjusted levels, rates vs. seasonally adjusted rates and seasonally adjusted levels vs. seasonally adjusted rates.
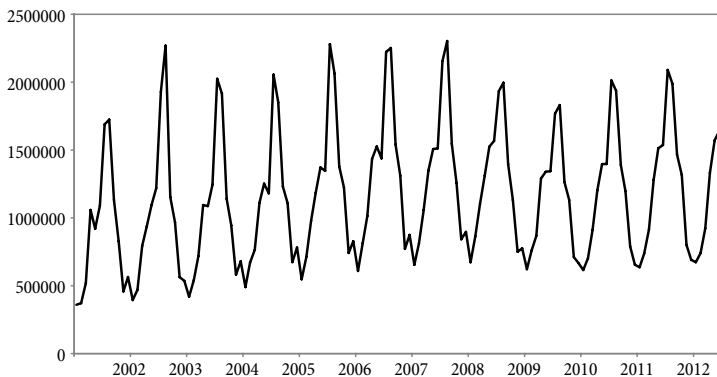

Fig. 1. Level of tourists coming to Catalonia
Source: compiled by the author.

Table 2. Distribution of the frequency of tourist arrivals

Year 2011	Tourist arrivals	%	% cumulated
France	3,614	24.1%	24.1%
Belgium and the Netherlands	1,382	9.2%	33.4%
United Kingdom	1,343	9.0%	42.4%
Germany	1,257	8.4%	50.7%
Italy	1,050	7.0%	57.8%
US and Japan	988	6.6%	64.4%
Russia	602	4.0%	68.4%
Northern countries	576	3.9%	72.2%
Switzerland	350	2.3%	74.6%
Other countries	3,806	25.4%	100%

Source: Compiled by the author, using data from the Statistical Institute of Catalonia (IDESCAT). Tourist arrivals expressed in thousands.

When analysing the forecast accuracy of the different ANN models, MLP and RBF networks show lower MAPE values than the Elman networks. MLP and RBF networks outperform the Elman networks in most countries and forecasting horizons. A possible explanation for this result is the length of the time series used in the analysis. The fact that the number of training epocs had to be low in order to maintain the stability of the network suggests that this architecture requires longer time series. For long training phases, the gradient sometimes diverged. The worse forecasting performance of the Elman neural networks compared to that of MLP and RBF architectures indicates that the feedback topology of the Elman network could not capture the specificities of the time series. Conversely, RBF networks display the lowest MAPE values in most countries, especially for growth rates.

When comparing the different pre-processing options, we obtain the best forecasting performance with raw data, especially for seasonally adjusted levels. In this context all three architectures show low MAPE values. The relative performance of MLP and Elman networks compared to RBF worsens when using growth rates, although the lowest MAPE is obtained with RBF networks in most cases. When using levels we also observe that the forecasting accuracy for longer time horizons improves in some countries, showing that working with levels is more indicated for longer term predictions.

When testing for significant differences between each two competing series (Table 6), we obtain better forecasts with seasonally adjusted levels than with levels, although not always significant. The only exception is Russia, the UK and Other countries for 6-month ahead forecasts. We also obtain significantly better forecasts with seasonally adjusted rates than with growth rates. The only exception being the predictions for longer forecast horizons in Russia and Other countries. Finally, when comparing seasonally adjusted levels and seasonally adjusted growth rates, with the exception of Russia, we always obtain significantly better forecasts with seasonally adjusted levels. These results show the suitability of deseasonalizing but not detrending for neural network forecasting with time series, especially for the longest forecasting horizons.

Table 3. MAPE (2010:04–2012:02) Levels vs. SA Levels

	Levels (original series)			Levels (seasonally adjusted)		
	MLP	RBF	Elman	MLP	RBF	MLP
France						
1 month	0.69	0.36	1.03	0.10	*0.04*	0.12
3 months	0.87	0.36	0.91	0.09	0.05	0.12
6 months	0.91	0.24	1.08	0.15	0.05	0.10
United Kingdom						
1 month	0.70	0.79	1.06	0.36	0.36	0.44
3 months	1.10	0.91	1.37	0.43	0.51	*0.33*
6 months	0.73	0.42	1.00	0.47	0.48	0.41
Belgium and the NL						
1 month	1.44	0.68	2.22	0.12	*0.05*	0.11
3 months	1.68	0.74	2.29	0.11	0.06	0.12
6 months	1.44	0.35	2.49	0.18	0.07	0.12
Germany						
1 month	0.72	0.55	0.83	0.08	0.12	*0.10*
3 months	0.84	0.63	0.98	0.12	0.15	0.16
6 months	0.71	0.20	1.02	0.17	0.15	0.13
Italy						
1 month	0.45	0.17	0.91	0.32	0.07	0.21
3 months	0.68	0.16	0.97	0.33	0.06	0.18
6 months	0.94	0.19	0.74	0.41	*0.05*	0.26
US and Japan						
1 month	0.43	0.30	0.30	*0.15*	0.25	0.22
3 months	0.52	0.35	0.72	0.25	0.30	0.16
6 months	0.40	0.31	0.53	0.27	0.30	0.17
Northern countries						
1 month	0.67	0.34	0.62	0.43	*0.16*	0.30
3 months	0.56	0.44	0.83	0.27	0.20	0.30
6 months	0.65	0.31	0.46	0.31	0.23	0.24
Switzerland						
1 month	0.54	0.35	0.67	0.11	*0.05*	0.13
3 months	0.87	0.45	0.86	0.14	0.07	0.09
6 months	0.81	0.25	0.80	0.17	0.07	0.08
Russia						
1 month	0.66	0.60	0.93	*0.22*	0.46	0.27
3 months	1.02	0.60	1.35	0.37	0.55	0.38
6 months	1.16	0.51	0.99	0.52	0.57	0.38
Other countries						
1 month	0.32	0.23	0.27	0.24	*0.16*	0.24
3 months	0.39	0.26	0.38	0.25	0.18	0.24
6 months	0.32	0.18	0.37	0.26	0.21	0.18
Total						
1 month	0.45	0.31	0.67	0.11	0.04	0.15
3 months	0.53	0.38	0.60	0.13	0.04	0.13
6 months	0.46	0.14	0.56	0.16	*0.03**	0.11

*Note: Italics: best model for each country. * Best model.*

Table 4. MAPE (2010:04–2012:02) Rates vs. SA Rates

	Raw Rates			SA Rates		
	MLP	RBF	Elman	MLP	RBF	MLP
France						
1 month	9.52	2.79	28.58	2.13	2.86	12.70
3 months	10.00	2.53	16.39	4.83	1.64	15.86
6 months	20.94	2.67	37.07	3.91	*1.33*	16.81
United Kingdom						
1 month	1.78	1.18	4.42	1.50	*0.77*	4.80
3 months	2.68	1.17	2.96	6.19	1.36	6.10
6 months	3.22	1.11	3.25	6.73	1.91	10.94
Belgium and the NL						
1 month	9.44	1.03	5.87	0.76	*0.65*	2.52
3 months	3.20	1.06	5.18	1.86	0.96	3.06
6 months	5.65	1.06	7.68	2.19	1.15	2.88
Germany						
1 month	4.32	1.11	5.09	3.20	2.10	8.56
3 months	3.42	*1.03*	6.13	4.24	1.66	13.53
6 months	4.54	1.17	3.61	2.91	2.30	14.87
Italy						
1 month	11.65	10.01	38.88	29.84	8.97	48.33
3 months	28.29	9.27	18.31	23.98	11.66	51.24
6 months	33.59	*5.19*	88.64	41.43	11.29	45.96
US and Japan						
1 month	1.82	1.07	2.96	1.46	*0.69*	2.13
3 months	3.00	1.00	3.16	1.37	1.14	3.46
6 months	2.30	1.15	4.47	2.40	1.08	4.15
Northern countries						
1 month	3.06	1.88	3.33	2.52	*1.49*	4.32
3 months	2.47	2.03	4.42	2.87	2.09	4.47
6 months	4.38	1.88	5.05	3.72	2.55	5.23
Switzerland						
1 month	2.37	1.13	5.39	*0.59*	0.63	2.18
3 months	3.49	0.96	4.10	1.04	1.06	1.36
6 months	4.04	1.04	4.25	1.38	1.06	1.96
Russia						
1 month	0.86	0.79	1.67	0.53	*0.31**	0.95
3 months	0.94	0.80	1.47	1.01	0.62	1.13
6 months	1.25	0.75	1.75	1.44	0.74	1.34
Other countries						
1 month	3.39	0.90	5.96	1.13	0.77	2.30
3 months	1.53	0.95	3.06	3.05	0.96	2.79
6 months	5.68	1.06	2.88	2.72	*0.73*	2.27
Total						
1 month	5.15	3.22	37.67	1.22	0.50	1.74
3 months	7.17	3.78	12.32	2.40	0.49	2.88
6 months	11.63	3.72	38.35	1.69	*0.47*	2.55

*Note: Italics: best model for each country. * Best model.*

Table 5. MAPE (2010:04–2012:02) SA Levels vs. SA Rates

	Levels (seasonally adjusted)			Rates (seasonally adjusted)		
	MLP	RBF	Elman	MLP	RBF	MLP
France						
1 month	0.10	*0.04*	0.12	2.13	2.86	12.70
3 months	0.09	0.05	0.12	4.83	1.64	15.86
6 months	0.15	0.05	0.10	3.91	1.33	16.81
United Kingdom						
1 month	0.36	0.36	0.44	1.50	0.77	4.80
3 months	0.43	0.51	*0.33*	6.19	1.36	6.10
6 months	0.47	0.48	0.41	6.73	1.91	10.94
Belgium and the NL						
1 month	0.12	*0.05*	0.11	0.76	0.65	2.52
3 months	0.11	0.06	0.12	1.86	0.96	3.06
6 months	0.18	0.07	0.12	2.19	1.15	2.88
Germany						
1 month	*0.08*	0.12	0.10	3.20	2.10	8.56
3 months	0.12	0.15	0.16	4.24	1.66	13.53
6 months	0.17	0.15	0.13	2.91	2.30	14.87
Italy						
1 month	0.32	0.07	0.21	29.84	8.97	48.33
3 months	0.33	0.06	0.18	23.98	11.66	51.24
6 months	0.41	*0.05*	0.26	41.43	11.29	45.96
US and Japan						
1 month	*0.15*	0.25	0.22	1.46	0.69	2.13
3 months	0.25	0.30	0.16	1.37	1.14	3.46
6 months	0.27	0.30	0.17	2.40	1.08	4.15
Northern countries						
1 month	0.43	*0.16*	0.30	2.52	1.49	4.32
3 months	0.27	0.20	0.30	2.87	2.09	4.47
6 months	0.31	0.23	0.24	3.72	2.55	5.23
Switzerland						
1 month	0.11	*0.05*	0.13	0.59	0.63	2.18
3 months	0.14	0.07	0.09	1.04	1.06	1.36
6 months	0.17	0.07	0.08	1.38	1.06	1.96
Russia						
1 month	*0.22*	0.46	0.27	0.53	0.31	0.95
3 months	0.37	0.55	0.38	1.01	0.62	1.13
6 months	0.52	0.57	0.38	1.44	0.74	1.34
Other countries						
1 month	0.24	*0.16*	0.24	1.13	0.77	2.30
3 months	0.25	0.18	0.24	3.05	0.96	2.79
6 months	0.26	0.21	0.18	2.72	0.73	2.27
Total						
1 month	0.11	0.04	0.15	1.22	0.50	1.74
3 months	0.13	0.04	0.13	2.40	0.49	2.88
6 months	0.16	*0.03**	0.11	1.69	0.47	2.55

*Note: Italics: best model for each country. * Best model.*

Table 6. Diebold-Mariano loss-differential test statistic for predictive accuracy

	Levels vs. Seasonally adjusted levels			Growth rates vs. Seasonally adjusted rates			Seasonally adjusted levels vs. SA growth rates		
	MLP	RBF	Elman	MLP	RBF	Elman	MLP	RBF	Elman
	France								
1 month	3.69	4.57	5.30	6.26	2.75	5.35	−4.37	−2.97	−2.70
3 months	4.80	4.48	3.98	5.34	6.98	−0.47	−2.73	−1.92	−3.94
6 months	2.84	5.14	3.73	4.70	6.54	2.61	−2.73	−1.74	−2.13
	UK								
1 month	1.65	1.49	2.46	2.45	4.58	2.41	−2.28	−1.21	−2.28
3 months	1.99	1.28	2.24	0.63	2.20	−1.22	−1.74	−3.43	−2.71
6 months	1.19	−0.51	1.98	1.82	0.21	1.06	−1.41	−2.57	−1.35
	Belgium and the NL								
1 month	4.23	3.91	6.58	3.12	5.45	2.34	−3.31	−4.16	−4.79
3 months	3.90	4.34	3.44	2.56	2.33	3.59	−1.93	−18.65	−3.08
6 months	2.08	3.14	4.15	5.50	1.63	2.22	−2.69	−25.33	−2.38
	Germany								
1 month	3.39	2.59	9.19	3.39	3.46	4.92	−1.27	−1.93	−1.59
3 months	4.11	2.76	3.35	4.11	3.38	3.18	−1.50	−2.57	−1.48
6 months	2.46	0.84	3.14	5.15	3.24	5.24	−2.41	−1.99	−1.33
	Italy								
1 month	1.25	3.14	5.59	1.56	3.32	2.84	−2.18	−1.82	−1.68
3 months	2.74	3.61	3.51	0.81	1.84	2.35	−2.41	−2.01	−1.76
6 months	1.94	5.25	2.86	1.25	2.11	0.49	−2.58	−2.14	−1.48
	US and Japan								
1 month	4.25	1.19	1.25	2.62	2.44	1.76	−2.24	−3.17	−4.55
3 months	3.21	1.25	7.18	1.06	1.52	1.27	−4.46	−2.75	−2.60
6 months	2.18	0.26	3.97	2.41	1.85	1.40	−2.46	−4.61	−2.01
	Northern countries								
1 month	1.12	5.20	3.98	2.23	3.79	1.06	−2.87	−2.11	−2.85
3 months	2.37	3.38	3.79	−0.46	1.95	0.61	−3.36	−3.09	−2.91
6 months	1.54	1.55	2.36	−0.70	1.25	0.25	−3.52	−2.77	−3.12
	Switzerland								
1 month	3.21	3.09	5.22	5.73	6.06	2.37	−5.20	−5.94	−3.21
3 months	4.97	3.66	7.54	3.25	3.47	2.94	−4.55	−15.64	−5.62
6 months	2.57	2.94	4.96	3.00	3.06	3.65	−5.96	−12.33	−5.82
	Russia								
1 month	2.60	2.35	5.64	2.20	4.91	3.07	−2.63	1.47	−2.83
3 months	2.96	0.62	2.75	0.11	4.17	0.88	−2.31	−0.81	−3.27
6 months	2.31	−1.35	3.92	0.58	−0.90	1.13	−2.04	−1.62	−3.80
	Other countries								
1 month	1.39	2.70	0.67	3.76	3.01	3.89	−3.61	−1.61	−2.23
3 months	1.80	2.62	2.15	0.25	1.89	0.86	−1.82	−1.36	−2.46
6 months	0.89	−1.26	2.01	2.29	1.14	−0.51	−2.09	−1.92	−3.96
	Total								
1 month	3.35	5.33	4.78	2.10	1.58	3.49	−7.92	−4.70	−5.43
3 months	4.25	6.30	3.42	0.36	1.92	1.82	−4.73	−6.82	−3.49
6 months	2.31	4.34	3.41	3.53	1.80	2.10	−5.92	−6.09	−3.04

Notes: 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. * Significant at the 5% level.

In contrast to previous studies (Hamzaçebi 2008; Alon *et al.* 2001; Franses, Draisma 1997), we find that ANNs produce significantly more accurate forecasts when built with deseasonalized data. Our finding on the suitability of working with seasonally adjusted levels for neural network forecasting confirms previous research by Zhang and Kline (2007), Zhang and Qi (2005), Virili and Freisleben (2000) and Nelson *et al.* (1999). The fact that we do not find evidence in favour of detrending could partly be explained by the fact that the data used for the analysis do not present a strong trend component.

Conclusions

Tourism demand forecasting has become essential in one of today's fastest growing industries. Accurate forecasts of tourist arrivals are crucial to develop a sustainable tourist model at the destination level. In this context, Artificial Neural Networks are a very useful technique for forecasting purposes. This study analyzes the effects of data pre-processing in the forecast performance of ANNs when using seasonal time series. This is an important issue in order to improve the accuracy in neural network-based time series forecasting. We implement a multiple-input multiple-output approach to predict international tourism demand in order to compare the forecasting performance of three different neural network architectures (multi-layer perceptron, radial basis function and the Elman neural network). We repeat the experiment using alternative approaches for input pre-processing (levels, growth rates, seasonally adjusted levels and seasonally adjusted growth rates) to analyze the effects of data pre-processing on the forecast accuracy of the different ANN models. To assess the differences between each two competing series we compute the Diebold-Mariano loss-differential test statistic for predictive accuracy.

When comparing the forecasting accuracy of the different input pre-processing techniques for each visitor market and for different forecasting horizons, we obtain significantly better forecasts with levels than with growth rates. We also find that seasonally adjusted series significantly improve the forecasting performance of the networks, indicating the importance of deseasonalizing when using neural networks for forecasting purposes. These results reveal that, when using seasonal data, neural networks performance can be significantly improved by working directly with seasonally adjusted levels. When comparing the forecasting accuracy of the different techniques, we find that multi-layer perceptron and radial basis function ANNs outperform the Elman networks. These results suggest that issues related with the divergence of the Elman neural network may arise when using dynamic networks with forecasting purposes.

The forecasting out-of-sample comparison reveals the suitability of applying multi-layer perceptron and radial basis function neural networks models for tourism demand forecasting. A 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, may improve the forecasting performance of practical neural network-based tourism demand forecasting.

Acknowledgements

References

Akkemik, K. A. 2012. Assessing the importance of international tourism for the Turkish economy: a social accounting matrix analysis, *Tourism Management* 33(4): 790–801. http://dx.doi.org/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(3): 147–156. http://dx.doi.org/10.1016/S0969-6989(00)00011-4

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(1): 68–81. http://dx.doi.org/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(2): 213–229. http://dx.doi.org/10.1177/0047287510362787

Athanasopoulos, G.; de Silva, A. 2012. Multivariate exponential smoothing for forecasting tourist arrivals, *Journal of Travel Research* 51(5): 640–652. http://dx.doi.org/10.1177/0047287511434115

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(5): 1761–1769. http://dx.doi.org/10.1016/j.matcom.2008.09.004

Bishop, C. M. 1995. *Neural networks for pattern recognition.* Oxford University Press. 482 p.

Cang, S. 2014. A comparative analysis of three types of tourism demand forecasting models: individual, linear combination and non-linear combination. *International Journal of Tourism Research* 16(6): 596-607. http://dx.doi.org/10.1002/jtr.1953

Cang, S.; Yu, H. 2014. A combination selection algorithm on forecasting, *European Journal of Operational Research* 234(1): 127–139. http://dx.doi.org/10.1016/j.ejor.2013.08.045

Celotto, E.; Ellero, A.; Ferretti, P. 2012. Short-medium term tourist services demand forecasting with rough set theory, *Procedia Economics and Finance* 3: 62–67. http://dx.doi.org/10.1016/S2212-5671(12)00121-9

Chen, K. 2011. Combining linear and nonlinear model in forecasting tourism demand, *Expert Systems with Applications* 38(8): 10368–10376. http://dx.doi.org/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(1): 215–226. http://dx.doi.org/10.1016/j.tourman.2005.12.018

Cho, V. 2003. A comparison of three different approaches to tourist arrival forecasting, *Tourism Management* 24(3): 323–330. http://dx.doi.org/10.1016/S0261-5177(02)00068-7

Chu, F. 2008. A fractionally integrated autoregressive moving average approach to forecasting tourism demand, *Tourism Management* 29(1): 79–88. http://dx.doi.org/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(6): 1414–1420. http://dx.doi.org/10.1016/j.tourman.2011.01.018

Claveria, O.; Torra, S. 2014. Forecasting tourism demand to Catalonia: neural networks vs. time series models, *Economic Modelling* 36(1): 220–228. http://dx.doi.org/10.1016/j.econmod.2013.09.024

Claveria, O.; Monte, E.; Torra, S. 2015. Tourism demand forecasting with neural network models: different ways of treating information, *International Journal of Tourism Research* 17(5): 492–500. http://dx.doi.org/10.1002/jtr.2016

Du Preez, J.; Witt, S. F. 2003. Univariate versus multivariate time series forecasting: an application to international tourism demand, *International Journal of Forecasting* 19(3): 435–451. http://dx.doi.org/10.1016/S0169-2070(02)00057-2

Franses, P. H.; Draisma, G. 1997. Recognizing changing seasonal patterns using artificial neural networks, *Journal of Econometrics* 81(1): 273–280. http://dx.doi.org/10.1016/S0304-4076(97)00047-X

Goh, C.; Law, R.; Mok, H. M. 2008. Analyzing and forecasting tourism demand: a rough sets approach, *Journal of Travel Research* 46(3): 327–338. http://dx.doi.org/10.1177/0047287506304047

Hadavandi, E.; Ghanbari, A; Shahanaghi, K; Abbasian, S. 2011. Tourist arrival forecasting by evolutionary fuzzy systems, *Tourism Management* 32(5): 1196–1203. http://dx.doi.org/10.1016/j.tourman.2010.09.015

Hamzaçebi, C. 2008. Improving artificial neural networks' performance in seasonal time series forecasting, *Information Sciences* 178(23): 4550–4559. http://dx.doi.org/10.1016/j.ins.2008.07.024

Haykin, S. 1999. *Neural networks. A comprehensive foundation*. Prentice Hall. 823 p.

Hong, W.; Dong, Y.; Chen, L.; Wei, S. 2011. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting, *Applied Soft Computing* 11(2): 1881–1890. http://dx.doi.org/10.1016/j.asoc.2010.06.003

Kon, S.; Turner, L. 2005. Neural network forecasting of tourism demand, *Tourism Economics* 11(3): 301–328. http://dx.doi.org/10.5367/000000005774353006

Law, R. 2000. Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting, *Tourism Management* 21(4): 331–340. http://dx.doi.org/10.1016/S0261-5177(99)00067-9

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(2): 14–24. http://dx.doi.org/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(5): 563–577. http://dx.doi.org/10.1002/(SICI)1099-1255(199909/10)14:5<563::AID-JAE530>3.0.CO;2-R

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(2): 372–383. http://dx.doi.org/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(5): 531–541.

Nelson, M.; Hill, T.; Remus, W.; O'Connor, M. 1999. Time series forecasting using neural networks: should the data be deseasonalized first?, *Journal of Forecasting* 18(5): 359–367.

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(2): 142–153. http://dx.doi.org/10.1177/0047287511400754

Pai, P.; Hung, K.; Lin, K. 2014. Tourism demand forecasting using novel hybrid system, *Expert Systems with Applications* 41(8): 3691–3702. http://dx.doi.org/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(5): 781–790. http://dx.doi.org/10.1016/j.tourman.2005.05.006

Pérez-Rodríguez, J. V.; Torra-Porras, S.; Andrada-Félix, J. 2005. STAR and ANN models: forecasting performance on the Spanish IBEX35 stock index, *Journal of Empirical Finance* 12(3): 490–509. http://dx.doi.org/10.1016/j.jempfin.2004.03.001

Ripley, B. D. 1996. *Pattern recognition and neural networks*. Cambridge University Press. 403 p. http://dx.doi.org/10.1017/cbo9780511812651

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. http://dx.doi.org/10.1016/j.knosys.2013.01.014

Sigala, M.; Chalkiti, K. 2014. Investigating the exploitation of web 2.0 for knowledge management in the Greek tourism industry: an utilisation–importance analysis, *Computers in Human Behavior* 30: 800–812. http://dx.doi.org/10.1016/j.chb.2013.05.032

Song, H.; Li, G. 2008. Tourism demand modelling and forecasting – a review of recent research, *Tourism Management* 29(2): 203–220. http://dx.doi.org/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. http://dx.doi.org/10.1016/j.protcy.2012.09.049

Tsaur, S.; Chiu, Y.; Huang, C. 2002. Determinants of guest loyalty to international tourist hotels: a neural network approach, *Tourism Management* 23(4): 397–405. http://dx.doi.org/10.1016/S0261-5177(01)00097-8

Tsaur, R.; Kuo, T. 2011. The adaptive fuzzy time series model with an application to Taiwan's tourism demand, *Expert Systems with Applications* 38(8): 9164–9171. http://dx.doi.org/10.1016/j.eswa.2011.01.059

Tsui, W. H. K.; Ozer-Balli, H.; Gilbey, A.; Gow, H. 2014. Forecasting of Hong Kong Airport's passenger throughput, *Tourism Management* 42: 62–76. http://dx.doi.org/10.1016/j.tourman.2013.10.008

Virili, F.; Freisleben, B. 2000. Nonstationarity and data preprocessing for neural network predictions of an economic time series, in S. I. Amari, C. Lee Giles, M. Gori, V. Piuri (Eds.). *Proceedings of IJCNN 2000*, 24–27 July 2000, Como, 5: 129–136. http://dx.doi.org/10.1109/IJCNN.2000.861446

Wu, Q.; Law, R.; Xu, X. 2012. A sparse Gaussian process regression model for tourism demand forecasting in Hong Kong, *Expert Systems with Applications* 39: 4769–4774. http://dx.doi.org/10.1016/j.eswa.2011.09.159

Yu, G.; Schwartz, Z. 2006. Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models, *Journal of Travel Research* 45(2): 194–203. http://dx.doi.org/10.1177/0047287506291594

Zhang, G. P.; Kline, D. M. 2007. Quarterly time series forecasting with neural networks, *IEEE Transactions on Neural Networks* 18(6): 501–514. http://dx.doi.org/10.1109/TNN.2007.896859

Zhang, G. P.; Qi, M. 2005. Neural network forecasting for seasonal and trend time series, *European Journal of Operational Research* 160(2): 501–514. http://dx.doi.org/10.1016/j.ejor.2003.08.037

Oscar CLAVERIA. Associate professor of Statistics and Econometrics, University of Barcelona. Research fellow at the Research Institute of Applied Economics (AQR-IREA). His research focuses on time series analysis, business surveys and tourism economics.

Enric MONTE. Associate Professor of Signal Theory and Communications, Polytechnic University of Catalunya (UPC). He received PhD in Digital Signal processing from UPC and a degree in Philosophy and Mathematics from UNED. His research areas include digital signal processing, automatic speech recognition and neural networks.

Salvador TORRA. Associate professor of Statistics and Econometrics, University of Barcelona. Research fellow at the group Multivariate Statistical Techniques and Optimization in Economics. His research focuses on finance, artificial intelligence and neuronal econometric methodology.