

## **APPLYING THE ARTIFICIAL NEURAL NETWORK METHODOLOGY FOR FORECASTING THE TOURISM TIME SERIES**

**Paula Fernandes<sup>1</sup>, João Teixeira<sup>2</sup>**

<sup>1</sup>*Department of Economics and Management, School of Technology and Management  
Polytechnic Institute of Bragança, Campus de Sta. Apolónia,  
Apartado 134, 5301-857 Bragança, Portugal, e-mail: pof@ipb.pt*

<sup>2</sup>*Department of Electrical Engineering, School of Technology and Management  
Polytechnic Institute of Bragança, Campus de Sta. Apolónia,  
Apartado 134, 5301-857 Bragança, Portugal, e-mail: joaopt@ipb.pt*

**Abstract.** This paper aims to develop models and apply them to sensitivity studies in order to predict the demand. It provides a deeper understanding of the tourism sector in Northern Portugal and contributes to already existing econometric studies by using the Artificial Neural Networks methodology. This work's focus is on the treatment, analysis, and modelling of time series representing "Monthly Guest Nights in Hotels" in Northern Portugal recorded between January 1987 and December 2005. The model used 4 neurons in the hidden layer with the logistic activation function and was trained using the *Resilient Backpropagation* algorithm. Each time series forecast depended on 12 preceding values. The analysis of the output forecast data of the selected ANN model showed a reasonably close result compared to the target data.

**Keywords:** artificial neural networks, time series forecasts, tourism, backpropagation, feedforward, training.

### **1. Introduction**

Several empirical studies in the tourism scientific area have been performed and published in the last decades. These studies agree that the forecast process in the tourism sector must be done with particular care.

In this context, and related to tourism demand in Northern Portugal, a study has been carried out with the reference temporal sequence –“Monthly Guest Nights in Hotels”.

The object sequence – “Monthly Guest Nights in Hotels” consist in the registered number of the overnight in the all category Hotels of the North region of Portugal (basically in the region of North of Douro River) in each month.

The study consisted in develop a model for predict the next year sequence knowing the values of the previous year. For this purpose the model used is supported in Artificial Neural Networks (ANN) methodology and was developed in the Matlab environment.

The ANN methodology was inspired in the biologic theories of human brain function [1]. The human brain is composed of several non-linear processors densely interconnected operating in parallel.

This paper is organized in the following structure: first, there is an overview section that examines the theoretical foundation of neural networks. Then, the

following section, in particular, assesses the use of ANN models as a forecasting tool for business applications. Based on the theoretical analysis, a neural network is developed for forecasting tourism demand in Northern Portugal. Real data from official publications in Portugal is used for the neural network development. The model development process, the empirical and analysis results of forecasting are described in section 4. The quality of forecasting results is measured in mean absolute percentage error. Some concluding remarks are given in the final section.

### **2. Analogy between artificial neuron and biological neuron**

The biological neuron is the basic building block of the nervous system. The human nervous system consists of billions of neurons of various types and lengths relevant to their location in the body [1]. The Fig 1 shows a schematic of an oversimplified biological neuron with three major functional units: dendrites, cell body and axon. The cell body has a nucleus that contain information about heredity traits; the dendrites receive signals from other neurons and pass them over to the cell body; and the axon, which branches into collaterals, receives signals from the cell body and

carries them away through the synapse to the dendrites of neighbouring neurons.

Because a neuron has a large number of dendrites/synapses, it can receive and transmit many signals at the same time; these signals may either excite or inhibit the firing of the neuron [2]. This basic system of signal transfer was the fundamental step of early neuro-computing development and the operation of the building unit of Artificial Neural Networks.

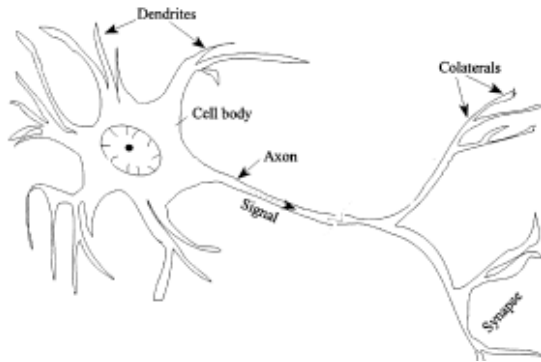


Fig 1. Schematic of biological neuron [2]

The basic analogy between artificial neuron and biological neuron is that the connections between nodes correspond to the axons and dendrites, the connection weights represent the synapses and the threshold approximates the activity in the soma.

3. Neural Networks Approach

The theory of neural network computation provides interesting techniques that mimic the human brain and nervous system, like we showed before. Neural networks are an information technology capable of representing knowledge based on massive parallel processing and pattern recognition based on past experience or examples. The artificial neural networks have been extensively studied and used in time series forecasting. The pattern recognition ability of a neural network makes it a good alternative classification and forecasting tool in business applications [3, 4, 5]. In addition, a neural network is expected to be superior to traditional statistical methods in forecasting because a neural network is better able to recognize the high-level features, such as serial correlation, if any, of a training set.

An additional advantage of applying a neural network to forecasting is that a neural network can capture the non-linearity of samples in the training set [2, 6, 7]. Using a neural network to forecast non-linear tourist behaviour could achieve a lower mean absolute percentage error, lower cumulative relative absolute error, and lower root mean square error than classical models [6, 8].

Artificial Neuronal Networks has been developed as generalizations of mathematical models of human cognition or neural biology, based on the assumptions that [9]:

- a. Information processing occurs at several simple elements that are called neurons;
- b. Signals are passed between neurons over connection links;
- c. Each connection link has an associated weight, which, in a typical neural net, multiplies the signal transmitted;
- d. Each neuron applies an activation function (usually nonlinear) to its net input (sum of weighted input signals) to determine its output signal.

Through replicate learning process and associative memory, the ANN model can accurately classify information as pre-specified pattern. A typical ANN consists of a number of simple processing elements called neurons, nodes or units. Each neuron is connected to other neurons by means of directed communication links. Each connection has an associated weight. The weights are the parameters of the model being used by the net to solve a problem. ANNs are usually modelled into one input layer, one or several hidden layers, and one output layer [10]. Fig 2 demonstrates a simplified neural network with three layers.

In Fig 2, each node in the hidden layer computes  $y_j (j = 1, 2, 3)$  according to expression (1) [9]:

$$f_j = \sum_{i=1}^2 x_i w_{ji} \tag{1}$$

In addition, a *sigmoid function* ( $y_j$ ), in the following form, is used to transform the output that is limited into an acceptable range. The purpose of a *sigmoid function* is to prevent the output being too large, as the value of  $y_j$  (for  $j=1, 2, 3$ ) must fall between 0 and 1:

$$y_j = \frac{1}{1 + e^{-f_j}} \tag{2}$$

Finally,  $Y$  in the node of the output layer in Fig 2 is obtained by the following function:

$$Y = \sum_{j=1}^3 y_j w_j \tag{3}$$

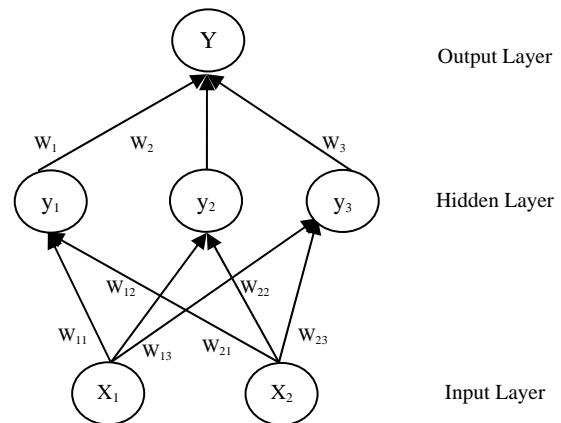


Fig 2. A neural network model [6]

Nodes in the input layer represent independent parameters of the system. The hidden layer is used to add an internal representation handling non-linear data. The output of the neural network is the solution for the problem. A feedforward neural network learns from a supervised training data to discover patterns connecting input and output variables. Feedforward recall is a one-directional information processing neural network in which the signal flows from the input units to the output units in a forward direction [11, 12, 13].

Much of the current interest in the neural network technique forecasting tool can be traced to the development of the backpropagation learning algorithm. This algorithm gives the network the ability to form and modify its own interconnections in a way that often rapidly approaches a goodness-of-fit optimum. The technique was first developed by Werbos and further advance by Rumelhart, Hinton and Williams [9].

Many variants of Backpropagation training algorithm were developed. In our case we adopted the *Resilient Backpropagation*, because it can combine fast convergence, stability and generally good results [14].

Usually, the learning process involves the following stages [4, 6]:

1. Assign random numbers to the weights;
2. For every element in the training set, calculate output using the summation functions embedded in the nodes;
3. Compare computed output with observed values;
4. Adjust the weights and repeat steps (2) and (3) if the result from step (3) isn't less than a threshold value; alternatively, this cycle can be stopped early by reaching a predefined number of iterations, or the performance in a validation set does not improve.
5. Repeat the above steps for other elements in the training set.

#### 4. Tourism demand forecasting in Northern Portugal

##### 4.1. Methods and procedures

For the selection of data we used the secondary source published in the Portuguese National Statistical Institute. According the study published by Fernandes [6] and Fernandes and Teixeira [5], the original time series suggests a power transformation (Fig 3a). Therefore we take logarithms of the data to stabilize the seasonality and variance, and we have another time series that was used during the work, namely Transformed Original Data (TOD) (Fig 3b).

The ANN model used in this study is the standard three-layer feedforward network. Since the one-step-ahead forecasting is considered, only one output node is employed. The activation function for hidden nodes

is the logistic function [*Logsig*]:  $f(x) = \frac{1}{1 + e^{(-x)}}$ ; and

for the output node the identity function (pure linear function) [*Lin*]:  $f(x) = x$ .

Bias terms are used in both hidden and output layer's nodes. The fast *Resilient* Backpropagation algorithm provide by the MATLAB neural network toolbox is employed in training process. The ANN is randomly initialised with weights and bias values. For selecting the architecture several experiments with different architectures were carried out (train and test) and the better architectures were selected according to the results in a validation set using hundreds of training sessions. So, the architecture consists of 12 input nodes in the entrance layer, 4 hidden nodes in the second layer and one node in the output layer – (1–12;4;1). The input of the model consists of the 12 previous numbers – corresponding to the last 12 months over-nights. The output is the predicted overnights for the next month. The selection of this model is supported in the author's work Fernandes [6] and Fernandes and Teixeira [5].

To make monthly predictions we have combined the following assumptions: consider as delayed inputs the most previous observations of the month we are predicting; due to the seasonal behaviour of the series we use a period of twelve months.

The data set was divided in a sub-set for training, a sub-set for validation and a sub-set for test. The data set between January 1988 until December 2001 (in a total of 168 observations) was used for training. In the training process of an ANN different end points are achieved, although with similar performance, for different initial values. Therefore, several training sessions for each identified situation have been performed with different initial weights. From this number of training sessions we retain the ANN (concerning its weights) that obtain better forecast results in each situation under the validation set (see [6]).

It must be noted that the data between January and December 1987 was used as the input data for predicting January 1988 till December. The data between January and December 2002 was used for the validation set. This set is used for early stop training if the root mean squared error (RMSE) does not decrease in a number of 5 training iterations. This early stop training condition avoids the ANN to over fit the training data without improvements in a data not used in the training phase. Finally, the data between January 2003 till December 2005 (36 observations) was used as the target. This data was never seen under the training and selection process and was only used to present the results of the model with never seen data.

For an ANN model the prediction equation for computing a forecast of  $Y_t$  using selected past observations can be written as [6]:

$$Y_t = b_{2,1} + \sum_{j=1}^n w_j f \left( \sum_{i=1}^m W_{ij} y_{t-i} + b_{1,j} \right) \quad (5)$$

where,  $m$  is the number of input nodes;  $n$  is the number of hidden nodes;  $f$  is a sigmoid transfer function

such as the logistic, used in the hidden layer nodes;  $\{w_j, j = 0, 1, \dots, n\}$  is a vector of weights from the hidden to output nodes;  $\{W_{ij}, i = 0, 1, \dots, m; j = 1, 2, \dots, n\}$  are weights from the input to hidden nodes;  $b_{2,1}$  and  $b_{1,j}$  are the bias associated with the nodes in output and hidden layers, respectively.

The equation shows a linear transfer function used in the output node and the resilient backpropagation algorithm was used for training the ANN.

**4.2. Empirical results**

In this section we will observe the results of each ANN under the test set. For this purpose we will compare the predicted data of each ANN with the target values for the last three years 2003 to 2005 (the tests set). We should emphasize that the target data is the original data of the time series and was never seen by the model in the training phase. The selection process of the better ANN is governed by the minimum RMSE in the training set (see [6]).

In order to compare the performance, the RMSE between the observed and predicted values are used as

the agreement index, according to equation (5). The other agreement index used in this paper is the coefficient of correlation between the observed and predicted values, equation (6).

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - P_t)^2}{n}}; \tag{5}$$

where  $A$  is the target value,

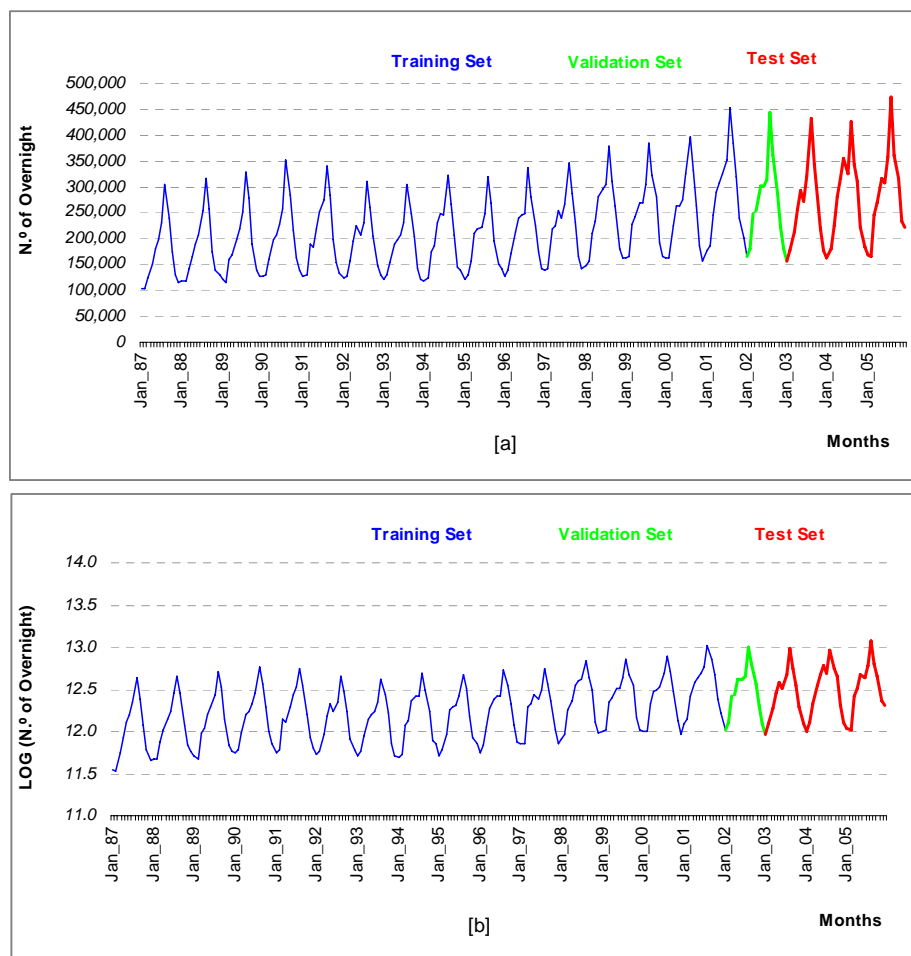
$P$  denotes the value of prediction and  $n$  the total number of observations.

$$r_{A,P} = \frac{\sum_{t=1}^n (A_t - \bar{A})(P_t - \bar{P})}{\sqrt{\sum_{t=1}^n (A_t - \bar{A})^2 \sum_{t=1}^n (P_t - \bar{P})^2}}; \tag{6}$$

where  $A$  is the target value,

$P$  denotes the value of prediction and  $n$  the total number of observations.

Comparing the results, presented in Table 1, of the prediction and considering that this model resulted from a selection of several different architectures we can say that the final results are stable and has an interesting performance, but we can never say that this is the better model. Therefore, this model is selected based only in the training set.



**Fig 3.** Monthly Guest Nights in Hotels in the North of Portugal from 1987:01 to 2003:12: (a) Original Data; (b) Natural Logarithms

**Table 1.** Results of ANN model.

Time Serie	Performance Measured			
	Training set (n=168)		Test set (n=36)	
	<i>r</i>	<i>RMSE</i>	<i>r</i>	<i>RMSE</i>
TOD	0.989	12.268	0.9615	21.8715

We should look now at the performance in the test set. Regarding the performance in the test set presented in Table 1 the RMSE becomes deteriorated now, but the correlation coefficient stills at a relatively high level.

The summary statistics for the out-of-sample (data used as the test set) forecasting performance of model are given in Table 2. MAPE is the Mean Absolute Percentage Error given by the expression (7).

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100. \quad (6)$$

**Table 2.** Out-of-sample (test set) forecasting error measures.

Year	<i>r</i>	<i>RMSE</i>	<i>MAPE (%)</i>
2003	0.983	18.969	6.39 %
2004	0.958	23.065	5.77 %
2005	0.967	23.315	6.89 %

Judged by all three overall accuracy measures and across three test sets we observe that the forecasting performance of this model in 2005 is not as good as in the other years but still at the very same level of accuracy. Therefore, according to the Criteria of MAPE for Model Evaluation in Lewis [15], the predicted data with the selected model has a highly accurate forecast, because the results are lower than 10 %.

Fig 4 displays the original and predicted time series for the entire sequence. As expected the model predicts better the values under the period of training set than for the period of the test set, that were not seen previously. We can observe an additional difficulty for the model imposed by the fact that years 2001 till 2005

have had an increasing number of overnights, and this increasing phenomenon was present in the training set only in 2001 and was difficulty to reflect for the following years. This phenomenon was due to the fact that the city of Guimarães and the Douro Region were considered World Cultural Heritage, and the city of Porto was the European Capital of Culture in 2001.

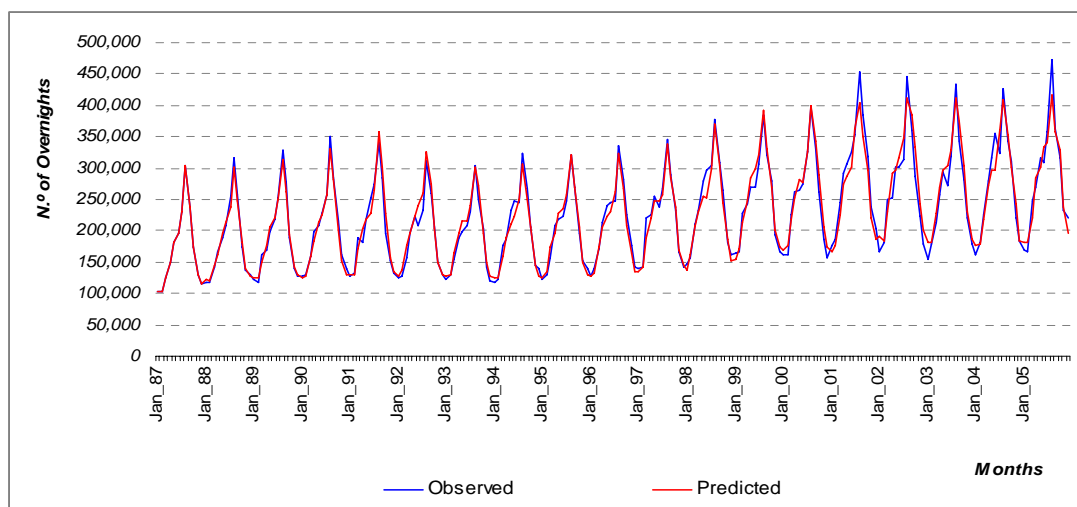
**5. Conclusions**

This paper presents the process of modelling tourism demand for the north of Portugal, using an artificial neural network model. The time series was considered in the logarithmic transformed data. This series was separated into a training data set to train the neural network, in a validation set, to stop the training process earlier and a test data set to examine the level of forecasting accuracy.

The model has 4 neurons in the hidden layer with the logistic activation function and was trained using the *Resilient Backpropagation* algorithm (a variation of backpropagation algorithm). The ANN model has the 12 preceding values as the input. The analysis of the output forecast data of the selected ANN model showed a reasonably close result compared to the target data. In other words, the model produced, according to the criteria of MAPE a highly accurate forecast. Therefore it can be considered adequate for the purpose of prediction in the reference time series. But the best model in most real world forecasting situations should be the one that is robust and accurate for a long time horizon and thus users can have confidence to use the model frequently.

To test the robustness of a model, it is critical to employ multiple years in the test set.

Finally and considering the results, the artificial neural network based models represent an effective alternative to classical models in tourism forecasting. This methodology becomes interesting to forecast because it allows the use of a non linear model for seasonal time series.



**Fig 4.** Comparison between Original Data and Predicted Values, from 2001:01 to 2005:12.

**References**

1. RUMELHARD, D. E.; MCCLELLAND, J. L. Parallel Distributed Processing – Explorations in the Micro-structure of Cognition. *Foundations*, Vol 1, 1986. 294 p.
2. BASHEER, I.A.; HAJMEER, M. Artificial Neural Networks: Fundamentals, Computing, Design and Application. *Journal of Microbiological Methods*, No 153, 2000, p. 3–31.
3. CHIANG, W.C.; URBAN, T.L.; BALDRIDGE, G.W. A Neural Network Approach to Mutual Fund Net Asset Value Forecasting. *Omega, The International Journal of Management Science*, Vol 215, No 2, 1996, p. 205–215.
4. ZHANG, G. P.; PATUWO, E.B.; HU, M.Y. Forecasting with Artificial Neural Network: The State of The Art. *International Journal Forecasting*, No 115, 1998, p. 35–62.
5. FERNANDES, P. O.; TEIXEIRA, J. P. A New Approach to Modelling and Forecasting Monthly Overnights in the Northern Region of Portugal. Proceedings of the 15th International Finance Conference (CD-ROM). Université de Cergy. Hammamet, Medina, Tunísia. 2007.
6. FERNANDES, P. O. Modelling, Prediction and Behaviour Analysis of Tourism Demand in the North of Portugal. Ph.D. Thesis in Applied Economy and Regional Analysis. 2005. 159 p.
7. HILL, T.; O’CONNOR, M.; REMUS, W. Neural Network Models for Time Series Forecast. *Management Science*, Vol 152, No 7, 1996, p. 1082–1092.
8. PATTIE, D. C.; SNYDER, J. Using a Neural Network to Forecast Visitor Behaviour. *Annals of Tourism Research*, Vol 23, No 1, 1996, p. 151–1615.
9. HAYKIN, S. Neural Networks. A comprehensive foundation. 1999. 231 p.
10. TSAUR, S; CHIU, Y.; HUANG, C. Determinants of Guest Loyalty to International Tourist Hotels-a Neural Network Approach. *Tourism Management*, No 23, 2002, p. 397-1505.
11. KUAN, C.; WHITE, H. Artificial Neural Network: An Econometric Perspective. *Econometric Reviews*, No 13, 1994, p. 1-91.
12. NAM, K.; SCHAEFER, T. Forecasting International Airline Passenger Traffic Using Neural Networks. *Logistics and Transportation Review*, Vol 31, No 3, 1995, p. 239–251.
13. YAO, J.; LI, Y.; TAN, C. Option Price Forecasting Using Neural Networks. *Omega, The International Journal of Management Science*, No 28, 2000, p. 1555–1566.
14. REIDMILLER, M; BRAUN, H. A Direct Adaptive Method for Faster Backpropagation Learning: The RPRO Algorithm. Proceedings of the IEEE International Conference on Neural Networks. 1993, p. 135–147.
15. LEWIS, C.D. Industrial and Business Forecasting Method. London: Butterworth Scientific., 1982. 305 p.