World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:6, 2012

Comparison of Artificial Neural Network Architectures in the Task of Tourism Time Series Forecast

João Paulo Teixeira, Paula Odete Fernandes

***Abstract*—**The authors have been developing several models based on artificial neural networks, linear regression models, Box-Jenkins methodology and ARIMA models to predict the time series of tourism. The time series consist in the "Monthly Number of Guest Nights in the Hotels" of one region. Several comparisons between the different type models have been experimented as well as the features used at the entrance of the models. The Artificial Neural Network (ANN) models have always had their performance at the top of the best models. Usually the feed-forward architecture was used due to their huge application and results. In this paper the author made a comparison between different architectures of the ANNs using simply the same input. Therefore, the traditional feed-forward architecture, the cascade forwards, a recurrent Elman architecture and a radial based architecture were discussed and compared based on the task of predicting the mentioned time series.

***Keywords*—**Artificial Neural Network Architectures, time series forecast, tourism.

I. Introduction

COUNTLESS empirical studies have been undertaken and published in the field of tourism in recent years, and they are unanimous in considering that the forecasting of tourism demand has an important role to play in the planning, decision-making and control of the tourism sector [1]-[4].

Tourism has been seen as a strategic sector in the future, for the Portuguese economy, and should influence all decision makers in this subject area to take policies that ensure their profitability and sustainability [5]. In this sense, tourism is a truly strategic interest for the Portuguese economy because of their ability to create wealth and employment. This is a sector that showed clear competitive advantages as with few other [6]. According to the World Tourism Organization (WTO), Portugal will reach 18.3 million foreign visitors in 2020. Tourism is at present one of the most important activities. Apart from its impact on the balance of payments and GDP, and its role on employment generation, investment and revenue, it is also recognized as the "engine" for development and other economic activities [7]. Similarly to Portugal also the Northern region of Portugal is ruled to be a very different region that offers an interesting alternative to the so called 'mass tourism', focusing on the provision of a wide variety of tourism products that range from the beach, the mountains, the thermal/health spas not forgetting the rural tourism, which had a significant increase in recent years [3].

Teixeira, João Paulo is Adjunct Professor at the Polytechnic Institute of Bragança - Portugal in the Electrical Department. Campus de Sta, Apolónia 5300-301 Bragança (phone: +351 273 30 3129, email: joaopt@ipb.pt).

Fernandes, Paula Odete is with the Economics and Management Department, Polytechnic Institute of Bragança and NECE-Research Unit in Business Sciences (UBI) CO 5300-857 Bragança, Portugal (phone: +351273303103; fax: +351273313051; e-mail: pof@ipb.pt).

In this respect, and given the substantial growth of this sector in the North of Portugal, it will be at all useful the development of models that could be used to make reliable forecasts of tourism demand, as it assumes an important role in the process of planning and decision-making both within the public and the private sector.

Currently available in the field of forecasting are a wide range of methods that have emerged in response to the most varied situations, displaying different characteristics and methodologies and ranging from the simplest linear regression model to the most complex approaches. The Box-Jenkins forecasting models belong to the family of algebraic models known as ARIMA models, which make it possible to make forecasts based on a given stationary time series. The methodology considers that a real time series amounts to a probable realization of a certain stochastic process. The aim of the analysis is to identify the model that best depicts the underlying unknown stochastic process and which also provides a good representation of its realisation, i.e. of the real time series. The ANN methodology also has had countless applications in the most diverse areas of knowledge and has been used in the field of forecasting as an alternative to the classical models. These non-linear models first appeared as an attempt to reproduce the functioning of the human brain, with the complex system of biological neurones being their main source of inspiration.

The aims of this current research consist in the research and highlight the best architecture of the Artificial Neural Networks methodology as an alternative to the classical models such as the Box-Jenkins methodology or the linear regression models in analysing tourism demand. The ANN methodology has aroused great interest in the field of economic and business sciences, since, from the research work undertaken so far, it can be seen that this represents a valid alternative to classical forecasting methods, providing a response to situations that would be difficult to treat through classical methods [8] and [9]. Hansen *et al*. [10] state that ANN demonstrate a capacity for improving time series forecasting through the analysis of additional information, reducing its size and lessening its complexity. To this end, each of the above-mentioned methodologies is centred on the treatment, analysis and modelling of the tourism time series: "Monthly Number of Guest Nights in the Hotels". Due to its characteristics, the time series is considered a significant indicator of tourist activity, since it provides information about the number of visitors that have taken advantage of tourist facilities.

II. The Time Series - Monthly Number of Guest Nights in the Hotels of the North Region of Portugal

The time series most used by the author [3], [11] – [13] are the Monthly Number of Guest Nights in the Hotels in the region of North of Portugal.

Therefore this time series will be used as the object to compare the architectures of the ANN models.

Fig. 1 displays the time series between January 1987 and December 2010.

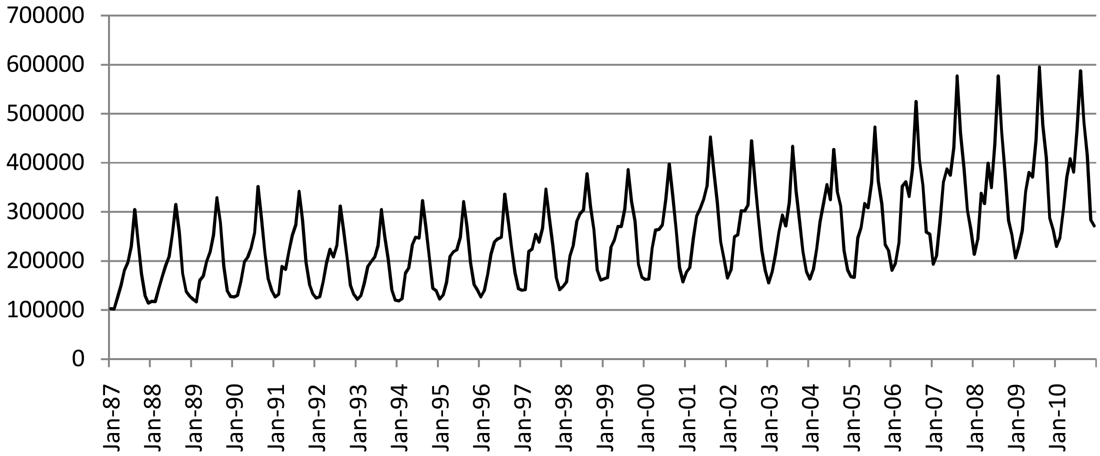

Fig. 1 Monthly Number of Guest Nights in the Hotels in the North region of Portugal

This time series is considered a significant indicator of tourist activity, since it provides information about the number of visitors that have taken advantage of tourist facilities, in this case in the North Region of Portugal.

The data observed cover the period between January 1987 and December 2010, corresponding to 288 monthly observations over the 24-years period. The values for the time series were provided by the Portuguese National Statistical Office (INE) [14].

Analysing the behaviour of the series it can be verified that there is seasonality (higher values during the summer months and lower values in winter). It is also clear that there is a progressive increase over the period in question. An increase from 1997 to 2001 is also apparent, and then there is a slight decrease until 2004, and then a significant growth from 2005 to 2009. The trend is a result of economic growth and investment in the tourism sector, which have occurred in northern Portugal in recent years. However, this trend is not apparently linear. This increase may be the result of investments made in marketing variables that promoted the region both nationally and internationally. Namely, in 1998 Lisbon hosted the World Exposition 1998. It was an official specialised World's fair from 22 of May to 30 September 1998. The theme of the fair was "the Oceans, a Heritage for the Future", was chosen to commemorate the 500 years of Portuguese discoveries. The Expo'98 received around 11 Million visitors in 132 days, and 155 countries and organizations were presented. Although Lisbon is outside the North region of Portugal, it is reasonable to consider an increase in the tourism demand also in the North region during the exposition. In 2001 the city of Oporto shared the designation of European Culture Capital. Oporto is the second largest city of Portugal and belongs to the North region.

In 2004 Portugal hosted the EURO 2004, the 2004 UEFA European Football Championship, a quadrennial tournament of European National Teams. Five of the ten stadiums were located in the North region. The tournament occurred between 12 June and 4 July. Although the number of overnights in 2004 suffered a decrease, it is believed that the outside image of Portugal and its organization capacity was improved, contributing to future incomes of tourists.

A. Training, Validation and Test Sets

The input of the AAN model is the previous 12 months of the time series. Therefore the 12 months of the year of 1987 was used only as the input for further months. The remaining data of the time series was divided in a set to perform the training of the ANN, corresponding to the period between January 1988 and December 2008, in a total of 252 input/ output pairs. The 12 months of the year 2009 was used to stop training early by a cross validation process. It is known as the validation set. The 12 months of the year of 2010 was used to measure the performance of the ANN with a set of data nor seen during the training process. It is known as the Test set.

III. The Artificial Neural Network Models

ANNs are models frequently associated within the broad field of knowledge related to artificial intelligence. They are based on mathematical models with an architecture that is broadly similar to that of the human brain. A neural network is composed of a set of interconnected artificial neurons, nodes, perceptrons or a group of processing units, which process and transmit information through activation functions. The connections between processing units are known as *synapses*.

The functions most frequently used are the linear and the sigmoidal functions - the logistic and hyperbolic tangent functions [3], [15] – [17]. It should also be mentioned that the neurons of a network are structured in distinct layers Fig. 2, (better known as the input layer, the intermediate or hidden layer and the output layer), with the ones most commonly used for the forecasting of time series being the multilayer perceptron or MLP [15], so that a neuron from one layer is connected to the neurons of the next layer to which it can send information. The connections between neuron i and j of previous and following layers are associated with a weight W_{ij}. Each neuron also was a bias b_i associated. Depending on the way in which they are linked between the different layers, networks can be classified as either feedforward, cascade forward, recurrent and radial networks [16].

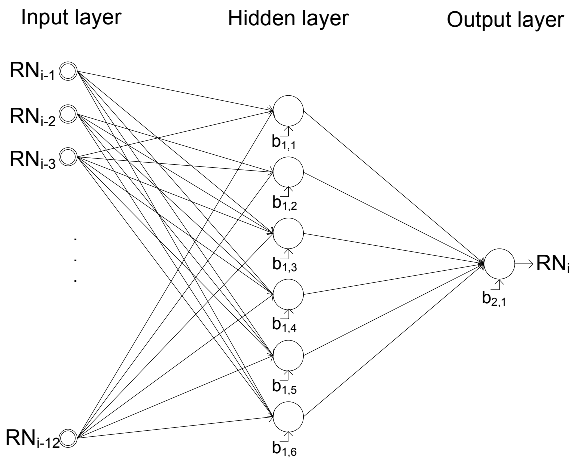

Fig. 2 Connection of neurons in a multilayer ANN

The specification of the neural network also includes an error function and an algorithm to determine the value of the parameters that minimise the error function. In this way, there are two central concepts: the physical part of the network, or, in other words, its architecture, and the algorithmic procedure that determines its functioning, or, in other words, the way in which the network changes according to the data provided by the environment [16].

It is also important to mention that for the ANN to learn with experience they have to be submitted to a process known as training, for which there are different training algorithms. One of the most frequently used algorithms in the forecasting of time series is the backpropagation algorithm [17] or its variants. During the training process, there is a "teacher" that provides a set of training cases, input/output pairs. Learning involves the minimisation of the performance function. This performing function usually is the mean squared error between output of the ANN and the target output vector. The minimization of the performance function is achieved by adjusting the matrix of weights W_{ij} of the connections and the neuron bias b_i according to a certain rule. The variants of the backpropagation algorithm most commonly used are based in the Levenberg-Marquardt Algoritm [18] and [19] or in the Resilient backpropagation [20].

In short, the output value produced by a feedforward network, with a hidden layer, can be expressed as follows [11]:

$$Y_t = b_{2,1} + \sum_{j=1}^{n} \alpha_j f\left(\sum_{i=1}^{m} \beta_{ij} y_{t-i} + b_{1,j} \right) \tag{1}$$

where,

m, number of nodes in the input layer;

n, number of nodes in the hidden layer;

f, sigmoidal activation function;

$\left\{\alpha_j, j = 0,1,\dots,n\right\}$, vector of weights that connects the nodes of the hidden layer to those of the output layer;

$\left\{\beta_{ij}, i = 0,1,\dots,m; j = 1,2,\dots,n\right\}$, weights that connect the nodes of the input layer to those of the hidden layer;

$b_{2,1}$ and $b_{1,j}$, indicate the weights of the independent terms (*bias*) associated with each node of the output layer and the hidden layer, respectively.

The equation also indicates the use of a linear activation function in the output layer.

For the other architectures used in this paper the detail will be discussed in respective section.

A. *Common Background*

In order to have the same conditions in the comparison of the architectures the input and the output was considered the same. Concretely, the input consists in the previous 12 months of the time series and the output consists in one node with the predicted value for next month of the time series.

In the architecture types where a fixed number of nodes in the hidden layer should be defined previously the number of 6 nodes was used. This is the case of the feedforward, cascade forward and recurrent architectures.

The input and number of nodes in the hidden layer was based in previous studies where these parameters were optimized using feedforward AAN [3] and [11] – [13].

The experiment was developed under Matlab environment [21].

Except in the Radial architecture all other architectures were trained 50 times and selected the best performance in the validation set.

B. *Feedforward Architecture*

This architecture is the most commonly used in artificial intelligence problems because it usually can solve properly the problems that other architectures not always can. Also in the previous studies of the authors it was always this architecture the selected one. Fig. 3 presents the general architecture of this ANN. This architecture has the common background, of the 12 nodes in the entrance, 6 nodes in the hidden layer and one node in the output. The activation function in the nodes of the hidden layer is the hyperbolic tangent, and in the output layer has the linear function as activation function. The ANN was trained with the Levenberg-Marquardt backpropagation algorithm [18] and [19].

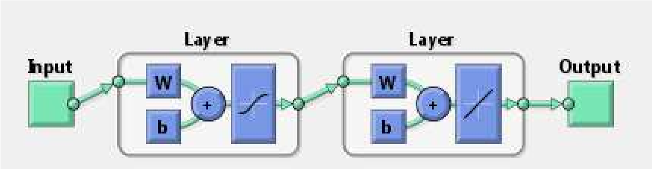

Fig. 3 Feedforward architecture of the ANN

C. Cascade Forward Architecture

This architecture, very rarely used, is similar to feedforward but has an additional connection between the input nodes and the output layer, as depicted in Fig. 4. This gives more direct connection between the input and the output in change by one additional matrix of weights. It is expected that this effect can be performed by the feedforward architecture with additional nodes in the hidden layer.

This ANN architecture also has 12 nodes in the input, 6 nodes in the hidden layer and one node in the output. The activation function in the nodes of the hidden layer is the hyperbolic tangent, and the linear function in the output layer.

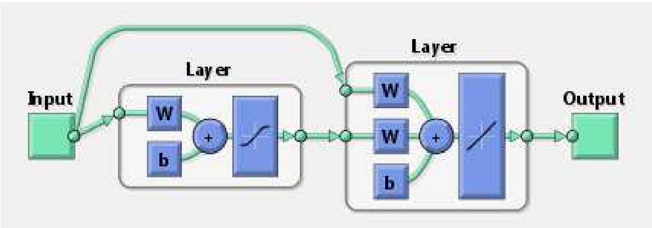

Fig. 4 Cascade forward architecture of the ANN

D. Elman Recurrent Architecture

This architecture, Fig. 5, is recurrent because there is a backward connection between the output of the hidden layer and its input. The delay in this connection stores values from the previous time step, which can be used in the current time step. Because the network can store information for future reference, it is able to learn temporal patterns as well as spatial patterns. The output layer is simple a linear layer that receives the output of the recurrent layer. The hidden layer has the hyperbolic tangent activation function. The ANN also has 12, 6 and 1 nodes in the input, hidden and output layers, respectively. The training Levenberg-Marquardt training algorithm is not adequate for this architecture, therefore the recommended algorithm as used (the gradient descendent with variable learning rate).

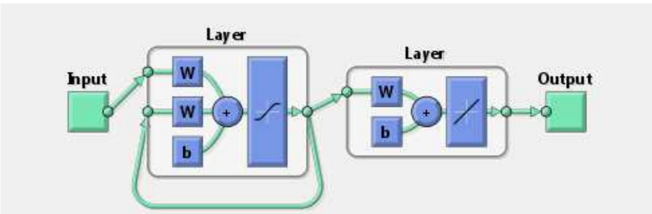

Fig. 5 Elman recurrent architecture of the ANN

E. Radial Basis Architecture

This ANN has its name based in the activation function of the hidden layer. The radial function is depicted in Fig. 6. The spread of the function a, is a parameter defined in the ANN.

The larger that spread is the smoother the function approximation will be. Too large a spread means a lot of neurons will be required to fit a fast changing function. Too small a spread means many neurons will be required to fit a smooth function, and the network may not generalize well.

The hidden layer is a radial basis layer, as shown in Fig. 7. The output of this layer is determined by the result of the radial function with the product between the bias b and the difference between the weights vector and the input vector. The complete architecture is presented in Fig. 8. The hidden layer is a radial layer and the output is a linear layer. The number of nodes is determined during the training process. Its number will increase until the output fits the target pattern unless a predetermined distance error.

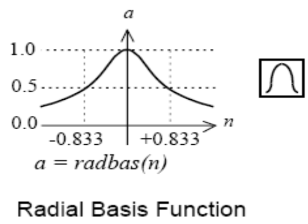

Fig. 6 Radial basis function

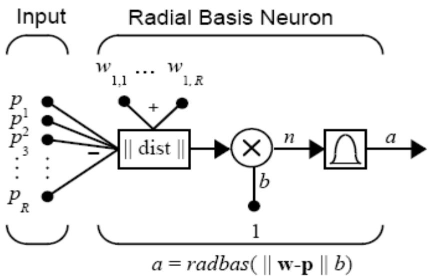

Fig. 7 Radial basis layer

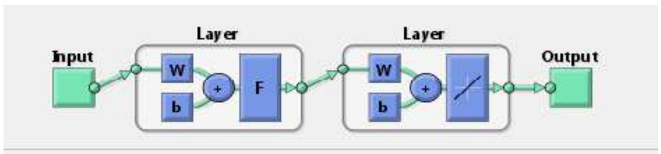

Fig. 8 Radial basis architecture of the ANN

IV. Discussion of the Results

In this section, the results for the test and validation sets will be analysed, comparing the real values observed with the forecast values for the time series using the different architectures. It should be mentioned that the validation set consist in the 12 months of the year 2009, and this set was used to stop the training with a cross validation method. This set was also used to select the best model along the 50 training sessions. The test set was never seen by the training process and was simple used to compare the predicted data with the ANNs with a new data set.

The mean absolute percentage error (MAPE) was used to measure the error distance between the predicted values and the target values of the time series.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|T_i - P_i|}{T_i}(\times 100) \quad (2)$$

Where N is the length of the set, T and P are the target and predicted values for month i.

The Radial basis architecture was not adequate for the present problem of prediction of future values based in the past. Fig. 9 presents the output of this architecture ANN. The ANN could fit exactly the data of the training set increasing the number of hidden nodes. The final number of 275 nodes was reached for the output of the network fit the real observed values of the time series with a null error. Anyhow the network was not able the predict future values given a zero values in the successive output. In this case only one training session was performed because all sessions gave the same result.

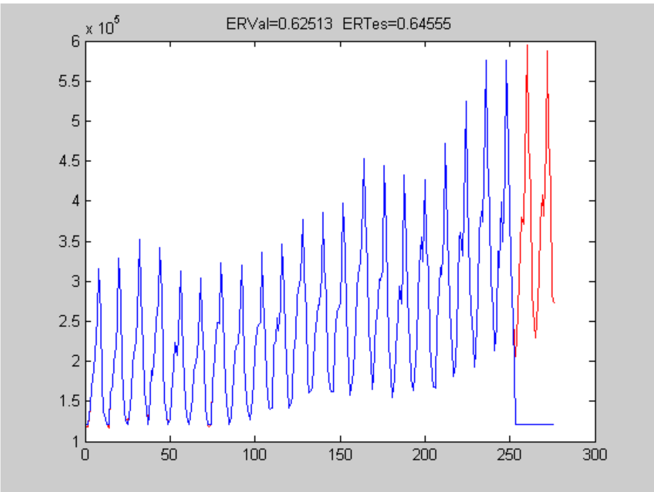

Fig. 9 Radial basis architecture output

TABLE I
MAPE FOR THE ARCHITECTURES ALONG THE VALIDATION AND TEST SETS

Architecture	Validation set	Test set
feedforward	4,15%	5,27%
cascade forward	4,36 %	5,36%
recurrent (Elman)	6,01%	4,57%

Therefore Table I presents the MAPE for the other architectures in the validation and test sets.

Fig. 10, presents the output of the feedforward architecture and the target (real) values along the validation (2009) and test (2010) sets. Fig. 11 presents the same values with the cascade forward ANN, and Fig. 12 the values with the Elman recurrent ANN. The MAPE presented for the three architectures are at the same level denoting that all architectures are adequate to predict this time series. The level of the error about 5% denotes that the three models can make prediction with relative high degree of accuracy.

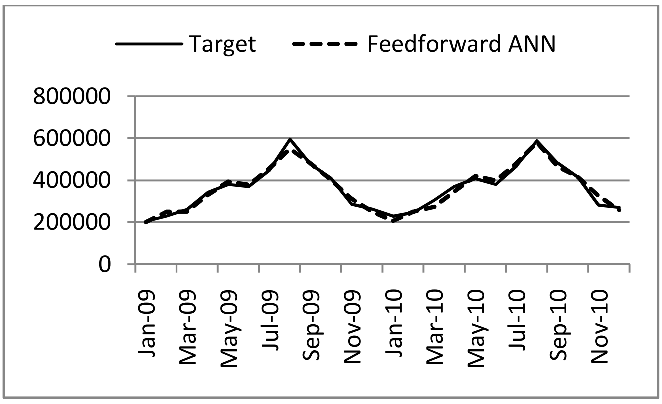

Fig. 10 Feedforward output for validation and test sets

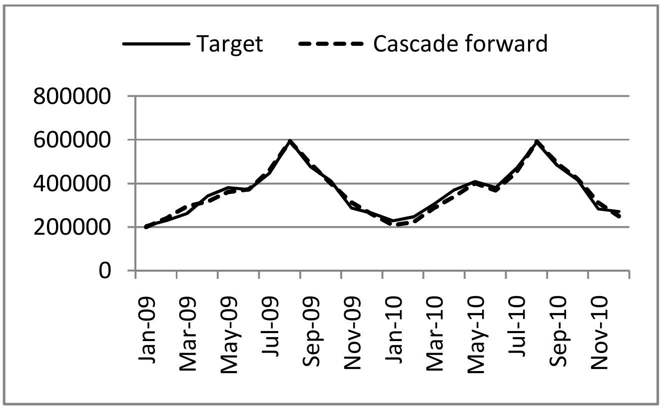

Fig. 11 Cascade forward output for validation and test sets

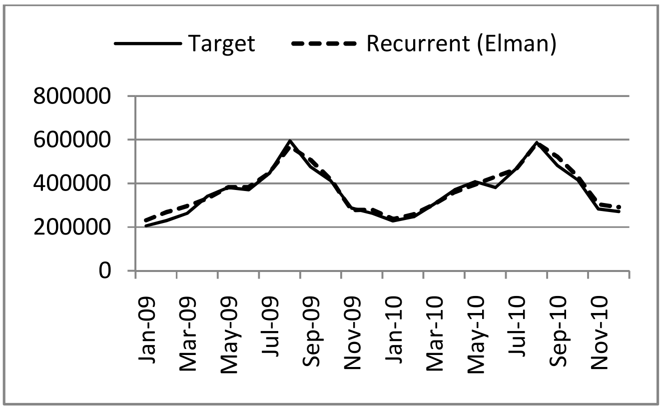

Fig. 12 Recurrent (Elman) output for validation and test sets

The plots of the output presented in the Fig. 10 to 12 confirm the highly competence of the ANNs models to fit a predicted data set of this tourism time series.

Comparing the three architectures by the MAPE, the cascade produces a relatively lower performance than the feedforward architecture in both sets. The recurrent architecture did not achieve the same performance in the validation set, but with the new data of the set the results became improved for 4,57%.

Fig. 13 presents the predicted values with the three architectures along the entire time series.

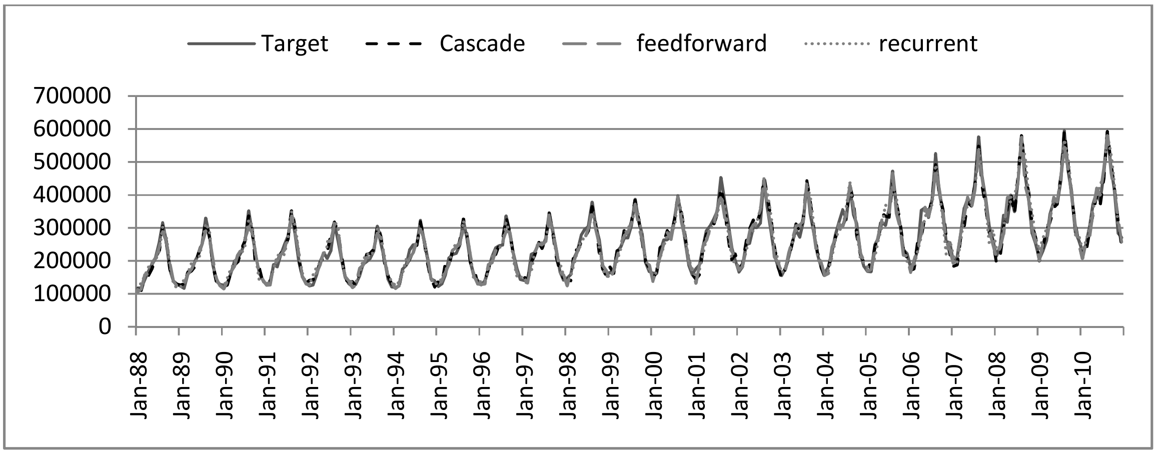

Fig. 13 Target and predicted values with the architectures along the all data set

V. Conclusions

The authors have been using different methodologies to predict the tourism time series of "Monthly Number of Guest Nights in the Hotels". The ANN models always had been compared with linear regression models, ARIMA and Box-Jenkins models. The feedforward architecture always were used with satisfactory results comparing with the other methodologies and never was experimented other architectures, although, their variations along different number of nodes in the input and in hidden layer. Also different features were experimented in the input, such as the time index [22] or the number of hour of sunshine [13] or even with economic features [23]. Now, different architectures were compared using the traditional input and number of nodes in the hidden layer. Specifically, the feedforward architecture, the cascade forward, the recurrent and the radial basis ANN were compared using the same time series.

The radial basis architecture was discard because it was shown inadequate to predict the future of a time series.

Considering the remaining three architectures, all demonstrated to have good ability to predict this time series. The results were at the same level of performance, although the feedforward achieved general better results considering both validation and test sets. Although the recurrent ANN produced better performance in the test set, it weren't consistent along the validation set.

References

[1] Witt, Stephen F. and Witt, Christine A.. *Forecasting tourism demand: a review of empirical research*. International Journal of Forecasting. N.º 11, pp.447/475, 1995.

[2] Wong, K. F.. *Introduction: Tourism Forecasting State of the Art.* Journal of Travel and Tourism Marketing; N.º 13 (1/2), pp.1/3. 2002.

[3] Fernandes, Paula Odete. *Modelling, Prediction and Behaviour Analysis of Tourism Demand in the North of Portugal*. Ph.D. Thesis in Applied Economy and Regional Analysis. Valladolid University - Spain.. 2005.

[4] Yu, Gongmei and Schwartz, Zvi. *Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models*. Journal of Travel Research. N.º 45, pp. 194/203, 2006.

[5] Dolgner, R. & Costa, A.. *Turismo, Sustentabilidade e Flexibilidade Laboral*. 16º Congresso da APDR Universidade da Madeira, Funchal, pp. 801-818. 2010.

[6] Ministério da Economia e da Inovação. *Plano Estratégico Nacional do Turismo – Para o desenvolvimento do Turismo em Portugal*. Lisboa. 2006.

[7] WTO; United Nations World Tourism Organization, Tourism Market Trends. [online]. UNWTO, 2006. Available in URL: http://www.unwto.org 02/2011.

[8] Thawornwong, S. and Enke, D.. *The adaptive selection of financial and economic variables for use with artificial neural networks.* Neurocomputing. N.º6, pp. 205/232. 2004.

[9] Hill, T.; O'connor, M. and Remus, W.. *Neural network models for time series forecasts.* Management Science. Vol. 42 (7), pp. 1082/1092. 1996.

[10] Hansen, J. V., Mcdonald, J. B. and Nelson, R. D.. *Time series prediction with genetic-algorithm designed neural networks: an empirical comparison with modern statistical models*. ComputlIntell. N.º15, pp. 171/184. 1999.

[11] Fernandes, P. and Teixeira, J.. *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.

[12] Fernandes, Paula O.; Teixeira, João Paulo - *Applying the artificial neural network methodology to tourism time series forecasting*. In 5th International Scientific Conference in 'Business and Management. Vilnius, Lithuania. ISBN 978-9955-28-267-9. 2008.

[13] Teixeira, J. P. & Fernandes, P. O. *A Insolação como Parâmetro de Entrada em Modelo Baseado em Redes Neuronais para Previsão da Série Temporal do Turismo*. CLME' 2011, Maputo.

[14] INE. *Anuário Estatístico da Região Norte 2010.* Instituto Nacional de Estatística, Lisboa. 2011.

[15] Bishop, C. M.. *Neural Networks for pattern recognition*. Oxford University Press. Oxford. London. 1995.

[16] Haykin, Simon. *Neural Networks. A comprehensive foundation*. New Jersey, Prentice Hall. 1999.

[17] Rumelhart, D. E. and McClelland, J. L.. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Volume 1: Foundations. The Massachusetts Institute of Technology Press, Cambridge. 1986.

[18] Hagan, M. T. and Menhaj, M.. *Training feedforward networks with the Marquardt algorithm, IEEE Transactions on Neural Networks*, vol. 5, nº 6, pp.989-993. 1994.

[19] Donald Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics* 11 (2): 431–441. 1963.

[20] Riedmiller, M. and Braun, H.. *A direct adaptive method for faster backpropagation learning: The RPROP algorithm. Proceedings of the IEEE International Conference on Neural Networks*. 1993.

[21] Demuth, H. and Beale, M.. *Neural Network Toolbox, for use with Matlab* – User's Guide, version 4, by the Math Works. 2000.

[22] Fernandes, Paula O.; Teixeira, João Paulo - *New approach of the ann methodology for forecasting time series: use of time index*. In International Conference on Tourism Development and Management. Kos, Greece. 2011.

[23] Fernandes, Paula O.; Monte, Ana Paula; Teixeira, João Paulo - Previsão da procura turística utilizando um modelo não linear. In XIII Congreso Internacional de Investigación en Ciencias Administrativas. Mexico. 2009.