Journal of Computer Science & Computational Mathematics, Volume 6, Issue 4, December 2016
DOI: 10.20967/jcscm.2016.04.005

Predicting Climatic Meteorological Parameters by Using the Artificial Dynamics Neural Networks: Case Study, Bushehr City

Sarkhosh Seddighi Chaharborj[1*], Masoumeh Mohajeri Borazjani[2], Yousof Gheisari[3]

[1]School of Mathematics and Statistics, Carleton University, 1125 Colonel By Dr, Ottawa, ON K1S 5B6, Canada
[1,2,3]Department of Computer, Engineering, Science and Research Branch, Islamic Azad University, Bushehr, Iran

[*]Corresponding author email: sseddighi2014@yahoo.com.my

Abstract: Climate change is found to be one of the main catastrophes to which human encountered. It serves as a threat to the Planet Earth and to predict its components has great deal of importance to planning on irrigation, controlling pests and diseases, drought as well as crisis management among many others. Since both temperature and humidity are the most important meteorological parameters so that other atmospheric changes are function of these two parameters, the present research tries to put forward appropriate prediction on them by using models of non-linear Autoregressive Neural Network and Autoregressive Network with Exogenous inputs (NARX). For this purpose, metrological data for Bushehr province in south of Iran for the years 2012-2013 and model performance criteria including R^2, RMSE and NRMSE were used. Different architectures for dynamic artificial neural network models were investigated through comparing the root mean square error. Models of performance validation suggested that Nonlinear Autoregressive Exogenous (NARX) forecasts temperature and humidity more accurate than Nonlinear Autoregressive Neural Network (NAR).

Keywords: forecasting, meteorology, dynamics of neural networks, NARX, NAR, Bushehr.

1. Introduction

To the best of our knowledge many researches in literature assessed neural network efficiency for prediction of meteorological parameters. Maqsood *et al.*, in 2004 [1] forecasted wind speed, relative humidity and temperature for the next 24 hours in Canada, Southern Saskatchewan using neural network forecasting models including MLPN, Elman ERNN, Radian Basic RBFN, Hopfield HFM and regression techniques and concluded that the neural network model compared to the regression models predicted climatic parameters more accurately and it can easily be extended and obviated needs to additional complicated calculations. Rather than ERNN and MLRN, RBFN predicts wind speed, relative humidity and temperature more accurately, while HFM model does not.

Smith *et al.*, in 2005 [2] developed a mode to forecast temperature using neural network. In this model, the past 24 hours data were used as input and short term temperature prediction. Alijani and Yosof in 2005 [3] compared and predicted the expected annual temperature in Tabriz with global temperature anomalies during the statistical period of 54 years (1951-2003) using linear regression and artificial neural networks and showed that non-linear models (ANNs) are much stronger than the linear and quasi-linear ones.

In a study entitled as "Application of neural networks in forecasting the weather", Hayati in 2007, [4] applied a two-layer perceptron neural network and radial basis function neural network as a tool to forecast the weather. In this research project, to enhance the forecasting accuracy some meteorological data from Kermanshah for ten years (1996-2005) were used. As results demonstrated, both networks are capable to forecast the weather accurately, at the same time multi-layer perceptron neural and radial basis function are proven to be suitable to forecast the temperature and humidity respectively. Salahi *et al.*, in 2010 [5] used multilayer perceptron (MLP) network to predict maximum temperatures in Ardabil city (northwest of Iran). Out of 21-years period (1985-2005), 85 percent of data was used to network training and the remaining was dedicated to network testing phase. Their results implied that the maximum temperature is forecasted in acceptable accuracy and the best ANN model was a 3-layer perceptron model with five neurons in the input layer, 3 neurons in the hidden layer, one neuron in the output layer and Marquardt-Levenberg training algorithm.

Esfandiari *et al.*, in 2011 [6] predicted average monthly temperature for years (2002-2005) to calculate model error by using monthly average temperature data from synoptic station Sanandaj (western Iran) in the 38 years period (1964-2001) as multi-layer perceptron networks inputs,.Then they validated the model of performance through statistical criteria such as regression and correlation coefficients between the observed and predicted temperature values as well as average relative error percentage. The results indicated appropriate performance and acceptable precision of artificial neural networks in temperature prediction. Correlation coefficient and model error average percent were estimated about 0.99 and 1.97% respectively. Olaiya and Adeyemo, in 2012, [7] in a study entitled "Application of data mining techniques to predict weather and climate changes" evaluated data mining techniques potential to predict the meteorological parameters (temperature, precipitation, wind speed and evaporation). They used neural network and decision tree algorithms.This study uses data from synoptic stations in Nigeria for years 2000 to 2009. C5 decision tree classification algorithm was considered to generate decision trees and rules for classification climatic parameters such as maximum temperature, minimum temperature, precipitation, evaporation, and wind speed in months and years of temporal scale. The results showed that in case of enough data, data mining techniques can be used to predict weather and climate change. Using NARX neural network in the hourly manner, Lubna *et al*., in 2013, [8] forecasted solar radiation of Amman, Jordan and then its performance was compared with various training algorithms. For this purpose, the meteorological data from 2004 to 2007 was used to train NARX and meteorological data from 2008 was applied to test it. The results obtained in this study showed that Levenberg-Marquardt training algorithm characterized with the lowest RMSE and the highest R value in both training and validation period in NARX model show the best performance.

Through metrological data from synoptic weather stations Mehrabad Tehran as input, Khalaji *et al*., in 2013, [9] forecasted the wind speed using Neural Network. As a result, the wind speed was predicted at different time intervals to use its power in order to establish new wind farm. In this paper, we deal with data by correlation coefficients and determining the wind speed dependence on input data. Statistical parameters such as mean absolute percent error were used to calculate error/s of training neural network. The obtained results indicate that given the forecasting process, error ranged 5-9% in acceptable interval. Many studies use static neural networks for prediction purposes. Di Piazza *et al.*, in 2013 and 2014, [10, 11] forecasted wind speed using the dynamic Recurrent Dynamic ANN neural network FTDNN and NARX. In this study, to predict the architecture of NARX architecture, a series-parallel architecture was used for learning and parallel architecture to forecast daily wind speed according to the stepwise forward prediction. Data on wind speed and minimum and maximum daily temperature data (2012-2010) recorded by weather stations in North East Sicily Palermo was used for training neural network. The results showed that both neural network models have good performance and accurate and reliable results were obtained to predict the wind speed.

Therefore, in the present study, we used dynamics neural network (NARX, NAR) to predict meteorological parameters (temperature and humidity) in Bushehr city.

2. Materials and Methods

As climate influences substantially human's social and personal lifestyle, enormous research centers are studying and dealing with climate issues throughout the world. The main topic on which such research institutions focus, is on weather forecasting i.e. the daily variations in troposphere. Such weather forecasting is based on the present and projected future values of important metrological parameters. Temperature and humidity are the most important and most basic climate parameters. In recent decades, most researchers have used statistical methods such as multiple, polynomial regression, AR, MA, ARIMA to model and predict meteorological processes. Indeed, the models incorporate parameters linearly in decision-making processes so that most often, it can properly analyze complicated climate issues, hence it seems necessary to introduce more efficient models to predict complicated nonlinear phenomena [12-17]. Formulas of AR, MA, ARIMA models, respectively can be presented as follows:

$$\left(1-\sum_{i=1}^{p}\alpha_i L^i\right)X_t = \varepsilon_t$$

$$X_t = \left(1+\sum_{i=1}^{q}\beta_i L^i\right)\varepsilon_t$$

$$\left(1-\sum_{i=1}^{p}\alpha_i L^i\right)(1-L)^d X_t = \left(1+\sum_{i=1}^{q}\beta_i L^i\right)\varepsilon_t$$

Weather forecasting is facilitated through nonlinear complex systems without any mathematical operations due to system variability over time forecasting methods get impossible via routine methods. On the other hand, weather forecasting receives substantial importance in various areas including economics, military, agricultural, etc.

Artificial Neural Network (ANN) is one of the efficient techniques in climate science leading to a tremendous surge in research activities. A neural network model is a natural parallel distributed process and its most important feature is the ability to model the nonlinear and complicated relationships requiring no prior assumptions on data relationship [18].

Neural networks fall in two categories: dynamic and static. The formers such as Artificial Neural Networks (ANN) are characterized with no feedback and their output is directly calculated by inputs having feed forward connections. However as for latters, such as Autoregressive Neural Networks (NAR) and neural network regression with exogenous variables (NARX), the output depends on present and past values of inputs, outputs, and the network architecture [19].

2.1 Autoregressive Neural Networks (NAR)

NAR (nonlinear autoregressive) neural networks can be trained to predict a time series from that series past values. In this model, time series future output y(t) are predicted as a nonlinear function based on past time series. Figure 1 illustrates NAR architecture. NAR equation is as the following:

$$y(t) = f\left(y(t-1), y(t-2), \ldots, y(t-d)\right). \tag{1}$$

Where d and t represent the number of output delay and time (days) respectively. *y(t)* is the predicted output values at time *t* dependent on the previous output values. *f* is a nonlinear function approximated using Artificial Neural Network Multi-layer Perceptron (MLP).

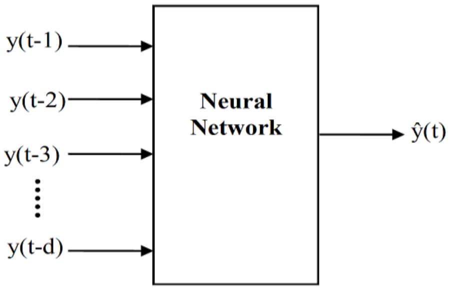

Figure 1. Structures of the NAR network.

2.2 Nonlinear Autoregressive Network with Exogenous Input (NARX)

Dynamic artificial neural networks or a nonlinear autoregressive exogenous model (NARX) is a nonlinear autoregressive model which has exogenous inputs. This means that the model relates the current value of a time series which one would like to explain or predict both: past values of the same series and current and past values of the driving (exogenous) series of the externally determined series that influence the series of interest. In addition, the model contains an "error" term which is related to the fact that knowledge of the other terms will not enable the current value of the time series to be predicted exactly. Hence, the artificial neural network NARX is a recurrent dynamic network, with feedback connections by which output layer to the input layer is connected with time delay. Neural network model NARX is based on the linear model ARX, which is commonly used in time series modeling. The equation for NARX model is the relation 2 [20]:

$$y(t) = f\left(y(t-1), y(t-2), \ldots, y(t-d_y), u(t), \mathrm{u}(\mathrm{t}-1), \mathrm{u}(\mathrm{t}-2), \mathrm{u}(\mathrm{t}-\mathrm{d}_u)\right) \quad (2)$$

Where: $d_y \geq 1$ and $d_u \geq 1$ denote number of delays in output and input respectively. *u(t)* is the input to the network at time *t*. *y(t)* is the predicted output value at time *t* dependent on the past output or input values. *f* is a non-linear function oapproximated using artificial Multilayer Perceptron Network (MLP). Figure 2 illustrates dynamic artificial neural network NARX where z^{-1} time delay unit in inputs and outputs [21]. Here time delay is expressed in day.

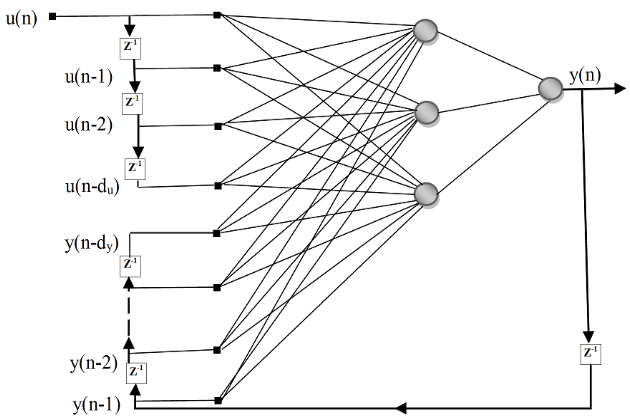

Figure 2. NARX model with tapped delay line at input.

2.3 Training Algorithm

Training manner in such networks is based on error correction training law. This method uses a random set of initial weights to carry out training. Having been determined the model output for each of the patterns in the training set, the error from difference between model output and the calculated expected values is corrected through recursion in the network in reverse direction (input to output). Levenberg-Marquardt method was used to train network weights. It applies following approximation for weight correction.

$$X_{k+1} = X_k - \left[H + \mu I\right]^{-1} g_k \cdot$$

(3)

Here, H is Hessian matrix, μ is learning rate, X_k previous weight changes, g_k is present rate gradients and *k* denotes learning repetition counter. LM algorithm is found to be the fastest and most accurate method for a medium networking (with a few hundred effective variables). In this study, values of parameters for the learning algorithm are presented in Table 1.

Table 1. The learning algorithm parameters

Maximum number of epochs to train	1000
Performance goal	0
Maximum Validation Checks	6
Minimum Gradient	1e-05
Initial mu	0.001
Mu Decrease Ratio	0.1
Mu Increase Ratio	10
Maximum mu	1.0000e+10

In addition, in the neural network used in this study, hyperbolic tangent activation function is applied as the transfer function in the hidden layer equation 4 and a linear function in the output layer.

$$f(x) = \frac{2}{1+e^{-2x}} - 1 \quad (4)$$

2.4 Data Sets and Evaluation Metrics

The data included temperature and humidity data for the period 2012 to 2013 collected from the Meteorological Agency of Bushehr province. The data consists of 731 weather data record information such as temperature and humidity daily (6, 9 and 15 GMT) recorded by synoptic station at Bushehr. The statistics are reliable and used without any modification since they have been collected from meteorology station. Also, in order to compare the prediction model performance, the following evaluation criteria models are considered.

$$R^2 = 1 - \frac{\sum(\hat{y}_t - y_t)^2}{\sum \hat{y}_t^2} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum(\hat{y}_t - y_t)^2}{n}} \quad (6)$$

$$NRMSE = \frac{RMSE}{\bar{y}_t} \quad (7)$$

As a whole, the parameters $R^{2,}$ RMSE and NRMSE are of the most widely used measures of the model accuracy in prediction among the others. In many studies, the RMSE criterion has been used as the best measure of fitting models since this measure is the average of the MSE and all the MSE features, including incorporating outliers to compare different models and as it is square root of MSE, and lessens error difference [22]. The normalized root mean square (NRMSE) is the relative difference between observed and predicted data to average observations. If the NRMSE is less than 10%, the simulation is very good, if more than 10% and less than 20%, it is good as well, if more than 20% and less than 30%, it is relatively good and if over 30% indicates weak simulation [13].

In evaluation criteria models y_t, $\hat{y}_t$ and n represent the target value (actual observation), the model output and observations number, respectively. Clearly the best value for R^2 equals to 1 and zero is for others.

3. Results

3.1 Results of Neural Network Model NAR Prediction

To predict air average temperatures for the next 10 days using the network NAR, the 731 data on average temperatures, about 70% of the data to train network and 20% to test the network and 10% to validate network were dedicated. Having completed the training and repeating errors, 20 neurons in the hidden layer and 8 neurons in the output delays were determined (Figure 3). The average temperature prediction by NAR in every day is a function of the average temperature for past 8 days. Hence the temperature values in the past days and next 10 days are considered as the network input and output respectively.

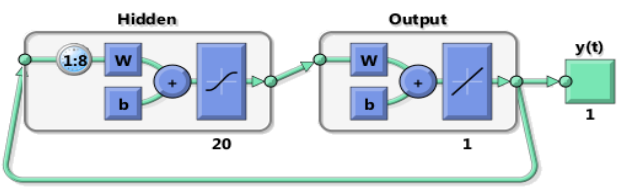

Figure 3. NAR network architecture in the prediction of the average air temperature.

To predict daily average humidity using the network NAR, about 70% of the data to train network and 20% to test the network and 10% to validate network were dedicated. The architecture designed in this network was obtained through trial and error during which using various values, different hidden layers and related neurons as well as number of delays in network output were created to predict average humidity. Architecture chosen after repeated trial and error is as 24 neurons in the hidden layer and 4 time delays caused by the output feedback (Figure 4).

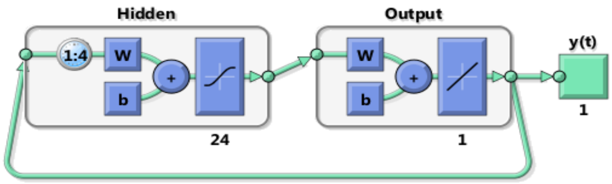

Figure 4. NAR network architecture in the prediction of average air humidity.

Table 2 presents the results of NAR networks to predict the average temperature and humidity.

Table 2. Assessment of the NAR neural network to predict the average temperature and humidity, T: Temperature and H: Humidity.

MODEL	RMSE		NRMSE		R^2	
	T	H	T	H	T	H
NAR	0.59	5.94	2.2%	9%	0.97	0.63

As it can be seen in Table 2, the regression coefficient between the observed and the predicted temperature data is 0.97, i.e. neural network NAR accounted for 97% of the variation. This indicates the good performance of the model in predicting the temperature. The NRMSE value for temperature is 2.2%, and given that NRMSE index value is less than 10%, this network accurately predicted the temperature.

As for humidity, R^2=0.63, i.e. NAR network explained about 63% of total variations, suggests that the neural network NAR has high forecast accuracy and cannot predict only 37% of humidity variables. Figure 5 and Figure 6 illustrate actual and predicted average values of temperature and humidity respectively.

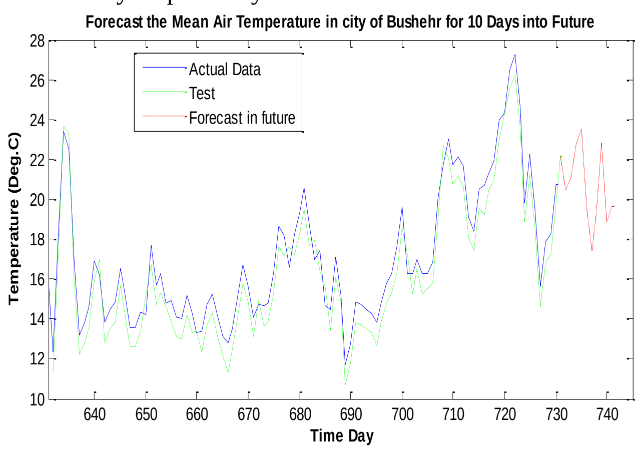

Figure 5. Comparison of average temperatures predicted by neural network NAR to actual data

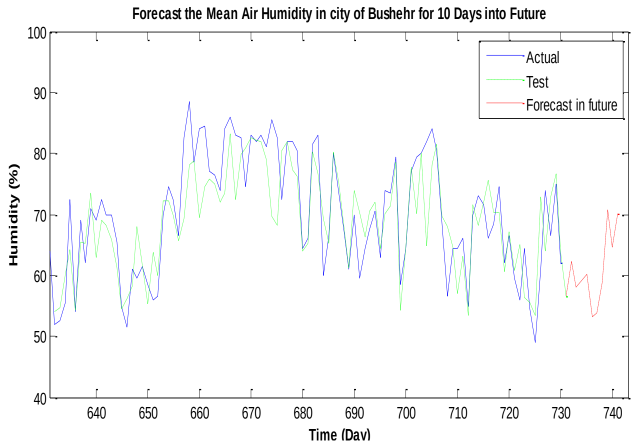

Figure 6. Comparison of average humidity predicted by neural network NAR to actual data.

3.2 Results of Neural Network Model NARX Prediction

To predict air average temperatures, 731 data on average temperatures (2012-2013), about 70% of the data to train network and 15% in individually for test and validate network were dedicated. In foregoing network exogenous and endogenous inputs included, average humidity and average temperature, and the network output was average temperatures in the 10 forthcoming days. It should be noted that the main reason behind the selection of humidity as model input was its contribution in precision of the temperature prediction. In this model, two structures of the series-parallel and parallel were used.

Series-parallel structure is undertaken to network training, model the temperature behavior and parallel structure to predict temperature parameter. All data, except those belonging to the last 100 days, are used for learning the temperature behavior. Different tests were used to calculate the best number of neurons in the hidden layer as well as the number of delays in model input and output. The best network architecture 2-26-1 with 4 delays input and 6 delays in output were obtained experimentally. Figure 7 shows the NARX network architecture for predicting average temperatures.

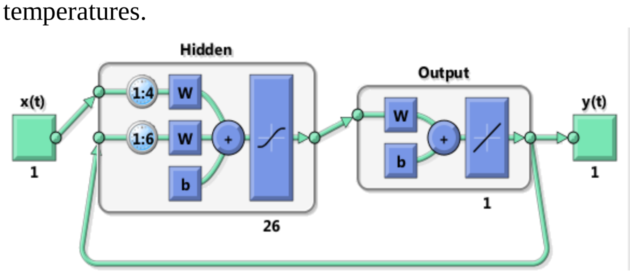

Figure 7. NARX network architecture in the prediction phase of the average air temperature.

Additionally, to model and to predict the average relative humidity, out of 731 humidity data, 70%, 15% and 15% were allocated to train, test and validation respectively. In this network, the temperature and humidity values were considered as an input (exogenous needed) and as output parameters (endogenous output), respectively.

Then the number of neurons in the hidden layer and the number of delays in the input and output layers are determined. Using trial and error, the best network architecture as 2-24-1 with 8 delays in input and output layers are obtained. Figure 8 shows the structure of the neural network NARX architecture for humidity prediction.

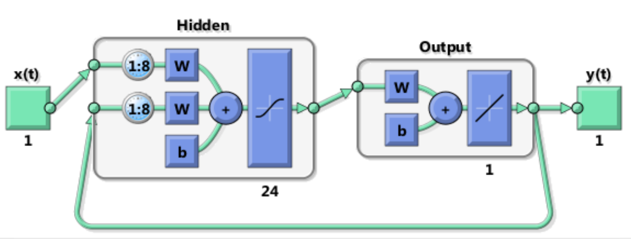

Figure 8. NARX network architecture to predict average air humidity.

Table 3 presents the results obtained from NARX neural networks performance to predict average temperature and humidity.

Table 3. Assessment of the NARX neural network to predict the average temperature and humidity, T: Temperature and H: Humidity.

MODEL	RMSE		NRMSE		R^2	
	T	H	T	H	T	H
NARX	0.35	5.81	1.3%	8.8%	0.99	0.65

As shown in Figure 9, due to network precision in temperature prediction, curves for actual and predicted data overlap completely, implying suitable the network performance in temperature prediction. At the same time, correlation coefficient between the actual and predicted data was estimated to be 0.99 and given that the error rate in RMSE is 0.35 and predicted values are close to the actual ones (Figure 9). It can be stated that the designed network could accurately simulate the temperature variations. Also, evaluating the performance assessment of NARX network to predict average air humidity, Table 3 showed that root mean square error and coefficient of determination were found to be 5.81 and 0.65 respectively, and given that the NRMSE is less than 10%, the network simulated the average humidity in accurate manner. Figure 10 shows the actual and the predicted values on average humidity.

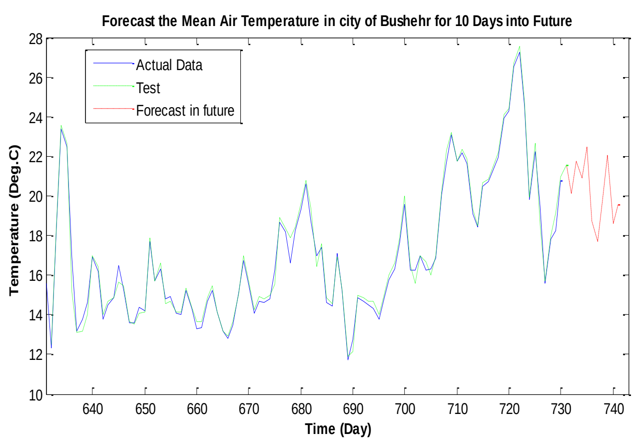

Figure 9. Comparison predicted and actual average temperatures data by neural network NARX.

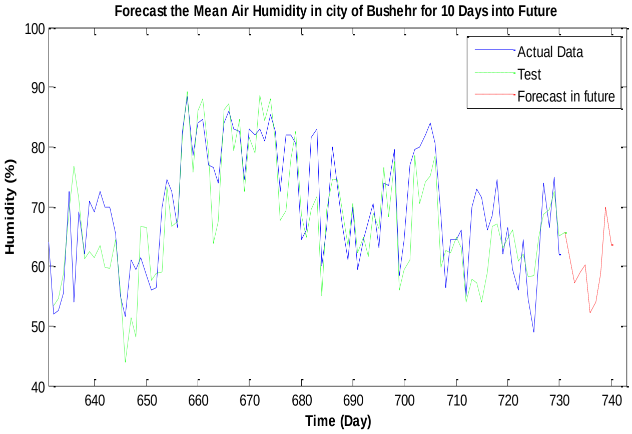

Figure 10. Comparison predicted and actual average humidity data by neural network NARX.

3.3 Comparison of Results

Results of prediction and comparative validation of NARX and NAR models are presented in Table 4.

Table 4. Comparative characteristics of artificial neural network NARX and NAR, T: Temperature and H: Humidity

MODEL	RMSE NRMSE				R^2	
	T	H	T	H	T	H
NAR	0.59	5.94	2.2%	9%	0.95	0.63
NARX	0.35	5.81	1.3%	8.8%	0.99	0.65

As Table 4 indicates, the NARX neural network due to having the highest correlation coefficient and the lowest root mean square error outperforms NAR. The NARX based predictions are much more accurate compared to those based the NAR.

The last but not the least, in light of above discussion, it can be stated that the results of NARX model tend much close to actual values and hence outperform ARMA model and it can be concluded that the application of artificial neural networks NARX as a non-linear method simulates temperature trend in a better time series.

4. Concluding Remarks and Future Perspectives

The main objective for this study is to predict the substantial climate parameters i.e. average temperature and average relative humidity. To this end, the two dynamic neural networks of NAR and NARX and Levenberg–Marquardt algorithm (LMA) were used. For the first time in this study, NAR and NARX networks were adapted to predict the climatic parameters in the study area. The comparison between the two models (networks) showed that NARX has more accurately for predicting the parameters. However, both networks predicted meteorological parameters very strongly (temperature and humidity) and simulated variations trend in appropriate manner.

In future researches it will be more promising to use other meteorological parameters to calibrate for NARX model in weather forecasting. Also, wavelet theory and or integrating of neural network models with time series as a new initiative in the meteorology field seems to be essential. Additionally, an algorithm to select the optimal number of neurons in the hidden layer and the number of input and output delays in neural networks is presented. Therefore, time required by this method for determining the optimal model structure is significantly reduced.

References

[1] I. Maqsood, R.K Muhammad, A. Ajith, “An ensemble of neural networks for weather forecasting”, *Neural Computing and Applications*, vol.13 no.2, pp.112-122, 2004.

[2] A. B. Smith, R. W. McClendon, G. Hoogenboom, “An Enhanced Artificial Neural Network for Air Temperature Prediction”, *In IEC (Prague)*, pp. 7-12, 2005.

[3] B. Alijani, G. R. Yosof, “Compare and forecast annual temperature changes Tabriz earth temperature anomalies using linear regression and artificial neural networks”, *Journal of Geography and Development*, vol. 3, no. 6, pp. 21-38, 2005.

[4] M. Hayati, “Application of neural networks in weather forecasting”, *Kermanshah Branch, Islamic Azad University Press*, 2007.

[5] B. Salahi, H. Seided Assad, S. Hossein, S. Behrooz., “Maximum prediction using artificial neural network. The city of Ardabil”, *Journal of Geographical Research*, vol. 98, pp.78-56, 2010.

[6] D. F. Esfandiari, S. A. Hosseini, M. M. Azadi, Z. Hejazizadeh, “Prediction of Monthly Average Temperature Through Artificial Neural Network Multi-Layer Perceptron (MLP)”, Geography, vol. 8, no. 27, pp. 45-65, 2011.

[7] F. Olaiya, A. Adeyemo, “Application of Data Mining Techniques in Weather Prediction and Climate Change Studies”, *Information Engineering and Electronic Business*, vol. 4, no. 1, pp. 51-59, 2012.

[8] M. Lubna, H. Mohhamad, E. Abdelhafez, S. Walid. “Hourly Solar radiation prediction based on nonlinear autoregressive exogenous (NARX) neural network”, *JJMIE*, vol. 7, no. 1, pp. 11-18, 2013.

[9] A. M. Khalaji, N. Younos, S. Sherwin, “Wind speed forecasting using neural networks in order to install a wind farm”, *The first National Conference on Renewable and Clean Energy*. Hamedan: iscconferences.ir., 2013.

[10] A. Di. Piazza, M. C. Di. Piazza, G. Vitale, “Solar radiation estimate and forecasting by neural networks-based approach”, In *XIII Spanish-Portuguese Conference on Electrical Engineering (XIII CHLIE), Valencia*, July, 2013.

[11] A. Di. Piazza, D. P. Maria Carmela, V. Gianpaolo, "Estimation and forecast of wind power generation by FTDNN and NARX-NET based models for energy management purpose in smart grids algorithms", *International Conference on Renewable Energies and Power Quality. Cordoba (Spain),* April 8-10, 2014.

[12] S. Belaid, A. Mellit, "Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate", *Energy Conversion and Management*, vol. 118, pp. 105-118, 2016.

[13] J. C. Savage, "The Parkfield prediction fallacy", *Bulletin of the Seismological Society of America*, vol. 83, no. 1, pp. 1-6, 1993.

[14] M. A. Khorshid Dust, G. R.Yousef, "Prediction of precipitation and determination of the wet and dry winter season in western Azarbayjan province", *Journal of Geographical Research*,Mashhad, vol. 72, pp. 25-36, 2008.

[15] T. Kim, T., V. Juan, "Nonlinear model for drought forecasting based on conjunction of wavelet transforms and neural networks", *Journal of Hydro. Eng. ASCE*, vol. 8, no. 6, pp. 319-328, 2003.

[16] H. S. Nogay, C.A. Tahir, E. Marija, "Application of artificial neural networks for short term wind speed forecasting in Mardin", *Turkey: Journal of Energy in Southern Africas*, vol. 23, no. 4, pp. 2-7, 2012.

[17] M. Zhu, Y. Pan, Z. Huang, P. Xu, "An alternative method to predict future weather data for building energy demand simulation under global climate change", *Energy and Buildings*, vol. 113, pp. 74-86, 2016.

[18] A. A. Dehghani, A. Reza, "Estimation of watershed discharge using artificial neural networks", *First International Conference on Water Crisis. Zabul*, 2008.

[19] L. Medsker, J. Lakhmi, "Recurrent neural networks: design and applications", *Boca Raton,.FL:CRC press*, 2000.

[20] S. Mandal, N. Prabaharan, "Ocean Wave forecasting using recurrent neural networks", *Ocean engineering*, vol. 33, no. 10, pp. 1401-1410, 2006.

[21] S. Haykin, "Neural Network: A Comprehensive Foundation", *2nd ed. Prentice Hall*, vol. 4, pp. 156-256, 1999.

[22] D. Swanson, T. Jeff, B. Thomas, "MAPE-R: a Rescaled Measure of Accuracy for Cross-Sectional Subnational Population Forecasts", *Journal of Population Researc*, vol. 28, no. 2-3, pp. 225–243.