Aplinkos tyrimai, inžinerija ir vadyba, 2010. Nr. 1(51), P. 5 – 10 *ISSN 1392-1649*
Environmental Research, Engineering and Management, 2010. No. 1(51), P. 5 – 10

Weather Forecasting Using ANFIS and ARIMA MODELS. A Case Study for Istanbul

Mehmet Tektaş
Marmara University, Vocational School of Technical Sciences, Turkey

(received in February, 2010, accepted in March, 2010)

This paper presents a comparative study of statistical and neuro-fuzzy network models for forecasting the weather of Göztepe, İstanbul, Turkey. For developing the models, nine year data (2000-2008) comprising daily average temperature (dry-wet), air pressure, and wind-speed have been used. Adaptive Network Based Fuzzy Inference System (ANFIS) and Auto Regressive Moving Average (ARIMA) models have been applied. To ensure the effectiveness of ARIMA and ANFIS techniques, different models employing a different training and test data set have been tested. The criteria of performance evaluation are calculated for estimating and comparing the performances of ARIMA and ANFIS models. Hence, the paper explains briefly how neuro-fuzzy models can be formulated using different learning methods and then analyzes whether they can provide the required level of performance for a reliable model for practical weather forecasting. On the obtained the results the most suitable model and network structure are determined according to prediction performance, reliability and efficiency. The performance comparisons of ANFIS and ARIMA models due to MAE (Moving Average Error), RMSE- R^2 (Root-Mean-Square error) criteria, indicate that ANFIS yields better results.

Key words: *Weather forecasting, ARIMA, Adaptive Network Based Fuzzy Inference System (ANFIS).*

1. Introduction

Weather forecast systems are among the most complex equation systems that computer has to solve. A great quantity of data, coming from satellites, ground stations and sensors located around our planet send daily information that must be used to foresee the weather situation in next hours and days all around the world. Weather reports give forecast for next 24, 48 and *72* hours for wide areas (Pasero 2004).

Weather forecasts provide critical information about future weather. There are various techniques involved in weather forecasting, from relatively simple observation of the sky to highly complex computerized mathematical models.

Data recorded by a weather station at the prominent meteorological center of Göztepe, İstanbul have been used for the analysis and forecast applying ANFIS and ARIMA, they are evaluated and compared..

2. Literature Review

Weather forecasting has been one of the most challenging problems around the world because of both its practical value in meteorology and popular sphere for scientific research.

Every sign points to the facts that there is a recognized need for accurate estimates of the temperature on a variety of temporal and spatial scales. Among scientific research works related to the weather forecasting some should be mentioned. Hu (1964) initiated the implementation of ANN, an important Soft Computing methodology in weather forecasting. Cook and Wolfe (1991) developed a neural network to predict the average air temperatures. Fuzzy logic can also be of great use in the atmospheric data analysis and prediction. Being capable of dealing with linguistic variables, this methodology can be utilized in analyzing atmospheric variables. Özelkan and Duckstein (1996) compared the performance of regression analysis and fuzzy logic in studying the relationship between monthly atmospheric circulation patterns and precipitation. Liu and Chandrasekar (2000) developed a fuzzy logic and neuro-fuzzy system for classification of a hydrometeor type based on polarimetric radar measurements where fuzzy logic was used to infer a hydrometeor type, and the neural network-learning algorithm was used for automatic adjustment of the parameters of the fuzzy sets in the fuzzy logic system according to the prior knowledge.

3. Data Analysıs And Case Study Characterıstıcs

Data recorded by a weather station at a prominent meteorological center of Göztepe, İstanbul are used for the analysis in this study. Their initial analysis has shown that the most important weather parameters are the average temperature (dry and wet), wind-speed and pressure variables. These variables also represent a strong correlation with the other weather parameters.

3.1. Temperature

Temperature, in general, can be measured to a higher degree of accuracy relative to any of the other weather variables. Historical temperature data recorded in the above-mentioned metereological center are used. The average temperatures in the period of nine years (2000-2008) are plotted in Figure 1.

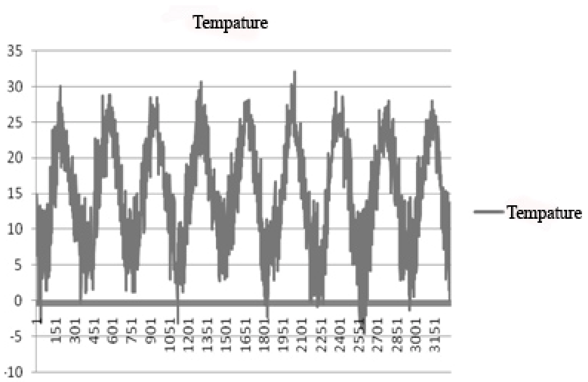

Fig. 1. Average temperatures in period 2000-2008

3.2. Wind-Speed

Wind is often considered as one of the most difficult meteorological parameters to forecast. It is significant in winter whether the temperature is low, which is roughly below 4.4 ^{o}C. When the temperature is below a freezing point, wind is a major factor determining the cooling rate. Figure 2 presents the wind-speed recorded in Göztepe, İstanbul between years 2000 and 2008.

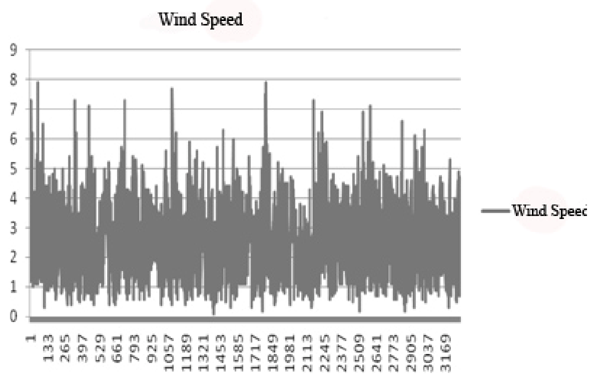

Fig. 2. Wind-speed in period 2000-2008

3.3. Pressure

Barometric pressure changes with local weather conditions, making the barometric pressure an important and useful weather forecasting tool. High pressure zones are generally associated with fair weather, while low pressure zones are generally associated with poor weather. For forecasting purposes, the absolute barometric pressure value is generally less important than the change in the barometric pressure. In general, the rising pressure indicates improving weather conditions, while the falling pressure indicates deteriorating weather conditions.The pressures in the period of nine years (2000-2008) are plotted in Figure 3.

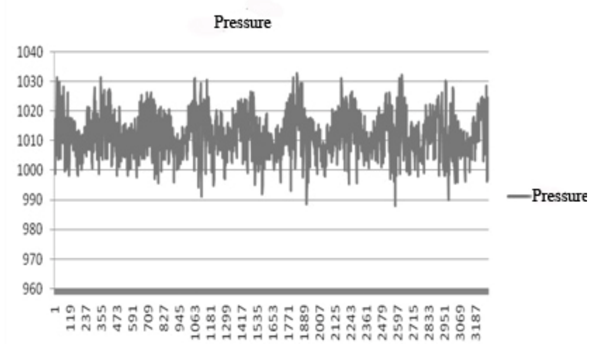

Fig. 3. Pressure in period 2000-2008

4. The ANFIS Model

ANFIS is an adaptive network that is functionally equivalent to a fuzzy inference system (Jang 1993), and referred to in literature as “adaptive network based fuzzy inference system” or “adaptive neuro fuzzy inference sytem”.

ANFIS is perhaps the first integrated hybrid neuro-fuzzy model (Jang 1992) and its architecture is very similar that presented in Fig. 4.

In the ANFIS model, crisp input series are converted to fuzzy inputs by developing membership functions for each input series. The membership function pattern used for the input series is of the Gaussian shape. The fuzzy inputs with their associated membership functions form the inputs to the neural network. These fuzzy inputs are processed through a network of transfer functions at the nodes of different layers of the network to obtain fuzzy outputs with linear membership functions that are combined to obtain a single crisp output, as the ANFIS method permits only one output in the model. Data from January 2000 to June 2008 are used to train the network to apply a hybrid method whose membership functions and parameters keep changing until the weather forecast error is minimized. Then the resulting model is applied to the test data of the period from July 2008 to December 2008.

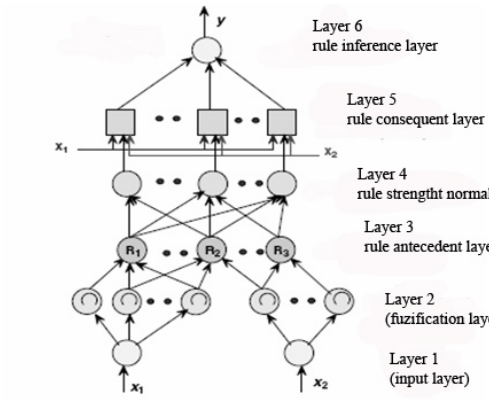

Fig. 4. Takagi Sugeno neuro-fuzzy system

Realized ANFIS model; Number of nodes: 55, Number of linear parameters: 80, Number of nonlinear parameters: 32, Number of training data pairs: 3123, Number of checking data pairs: 164 and Number of fuzzy rules: 16 are obtained.

Performance analysis of the ANFIS model is realized using MATLAB's Fuzzy Logic Toolbox and GUI editor.Training performance of a realized model obtaining from 20,000 epoch both for chkerr and step size parameters are shown in Fig 5a and b.

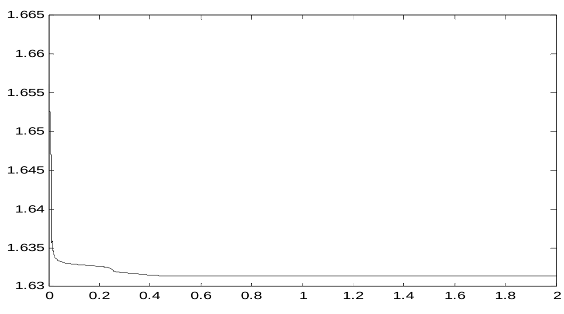

Fig. 5.a. Chkerr parameters of Training performance

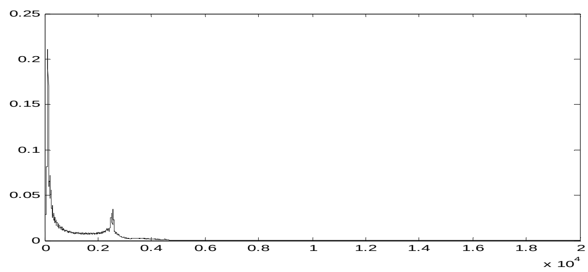

Fig. 5.b. Step size parameters of Training performance

Real and ANFIS system outputs related to the temperature forecasting are shown both for training and testing procedures in Fig.6.

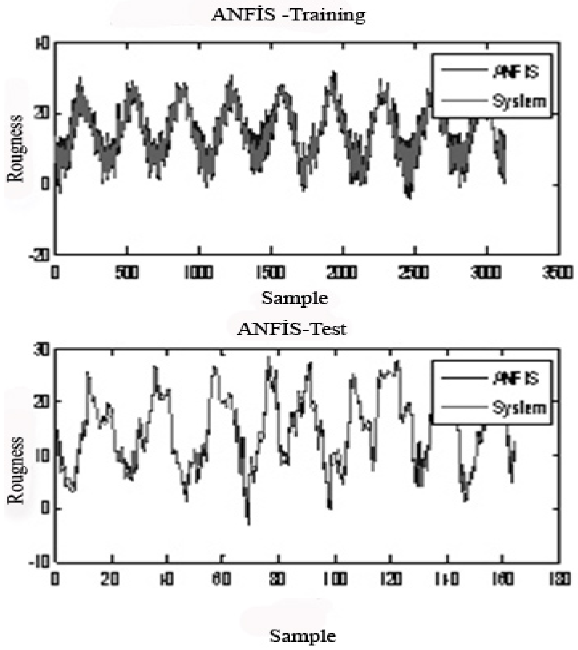

Fig. 6. Real and ANFIS system outputs related to tempature forecasting

Performance analysis of the ANFIS model is realized using MATLAB's ANFIS Toolbox and GUİ editor. The regression results of training and testing procedures are shown in Fig.7.

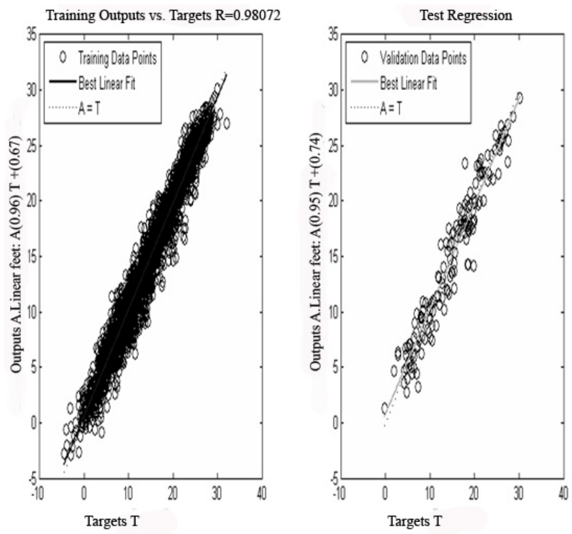

Fig. 7. Regression results

5. The ARIMA MODEL

In real world applications, many processes can be represented using the time series as follows:
X(t-p),…x(t-2),x(t-1),x(t)

For making a prediction using time series, a great variety of approaches are available. Prediction of scalar time series $\{\hat{x}(n)\}$ refers to the task of finding estimate $\hat{x}(n+1)$ of next future sample x(n+1) based on the knowledge of the history of time series, i.e. samples x(n), x(n-1) (Rank 2003)

Linear prediction, where the estimate is based on a linear combination of N past samples, can be represented as below:

$$\hat{x}(n+1) = \sum_{i=0}^{N-1} a_i . x(n-i)$$

with prediction coefficients α_i , *i=0,1, ...N-1.*

Introducing general nonlinear function f(.) $f(.) : R^n \to R$ applied to vector x(n)=[x(n), x (n-M), …,x(n-(N-1))M]^T of the past samples, nonlinear prediction approach $\hat{x}(n+1)$ =f(x(n)) (Rank 2003) is reached.

Traditionally, a time series forecasting problem is tackled using linear techniques such as Auto Regressive Moving Average (ARMA) and Auto Regressive Integrated Moving Average (ARIMA) models popularized by Box and Jenkins (1976).

General form of ARMA(p,q) model can be written as below:

$$X_t - \sum_{r=1}^{p} \phi . X_{t-r} = \sum_{s=0}^{q} \theta_s . \varepsilon_{t-s}$$

Where $\{\varepsilon_t\}$ is white noise . This process is stationary for appropriate ϕ, θ (Box and Jenkins 1976).

General form of the ARIMA model is given by:

$$Y_t = a_o + \sum a_i . y_{t-i} + \sum b_j . e_{t-j}$$

İ=1,2, ...p and j=0,1,...q

Where Y_t is a stationary stochastic process with non-zero average, a_0 - constant coefficient, e_i - white noise disturbance term, a_i represents autoregressive coefficients and b_j denotes moving average coefficients.

Performance analysis of the ARIMA model is done using the SPSS's package program. The Model obtained is ARIMA(2,1,1). Real and ARIMA system outputs related to tempature forecasting are in Fig.8.

The estimate equation obtained by the Arima(2,1,1) model and its result parameters are as follows:

$Z=0.4805160257+Z_{t-1}(0.5054977403)+$

$Z_{t-2}(0.5054977403)+ 0.5770235918\,\varepsilon_{t-1}$

$R^2=0,91$ D-W=2.036871 AIC=4.349352

Accuracy of the ARIMA(2,1,1) model, ACF and PACF corelograms are examined and shown in Fig.9.

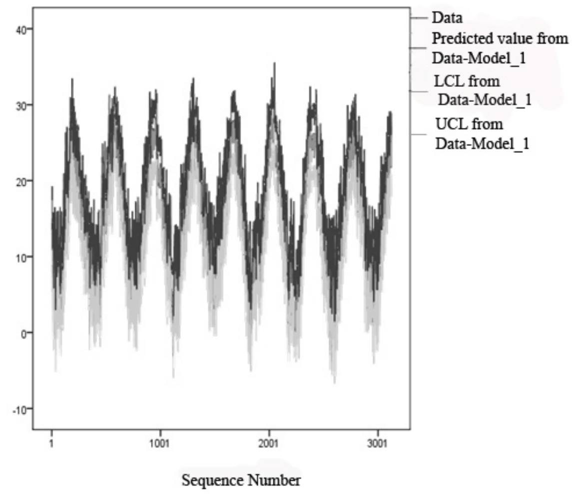

Fig. 8 Real and ARIMA system outputs related to tempature forecasting

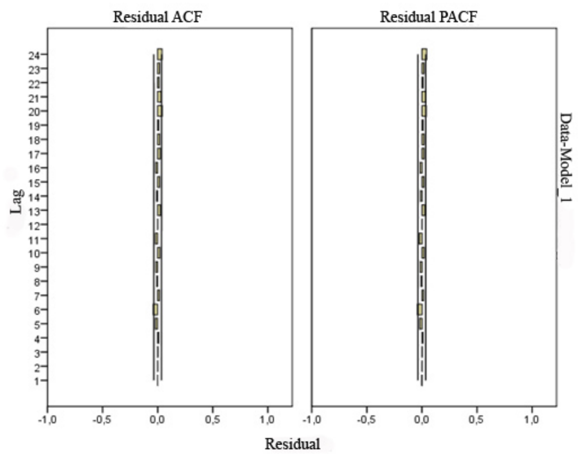

Fig. 9. Residual ACF and PACF corelograms of the ARIMA(2,1,1) model

ARIMA(2,1,1) model and graph of residual terms are shown in Fig.10.

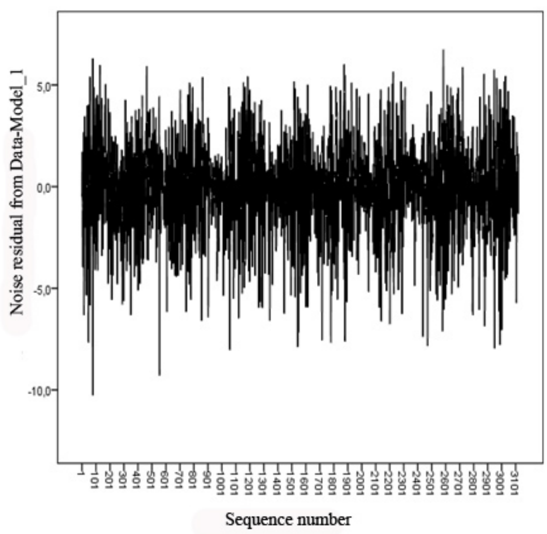

Fig. 10. Graph of residual terms

6. Discussions and Test Results

In this study, four inputs and one output were used to forecast the tempature in the ANFIS model. 5% of the data were used in testing and 95% of data were used in a training process. The same data were used by ARIMA(2,1,1) method to evaluate performance criteria. The performance comparisons of the ANFI and ARIMA models due to MAE,RMSE,R^2 CRİTERİA are shown in Tables1.a and .b.

Table.1. a. Performance comparison of ANFIS and ARIMA models

	MAE		RMSE	
	ANFIS	ARIMA	ANFIS	ARIMA
Training Data	1.2552	1.523	1.6312	2.075
Test Data	1.3212	1.178	1.7176	1.709

Table.1. b. Performance comparison of ANFIS and ARIMA models

	R^2
ANFIS	0.9802
ARIMA	0.91

As can be understood from Tables 1.a and b., compared to the ARIMA, the ANFIS due to MAE,RMSE,R^2 criteria gives better results.

7. Conclusions

ANFIS has gained a great popularity in time-series prediction because of its simplicity and reliability. The performance of ANFIS and ARIMA is compared. Compared to the ARIMA, the ANFIS can more efficiently capture dynamic behavior of the weather temperature, resulting in a more compact and natural internal representation of the temporal information contained in the weather profile.

References

Box,G.E.P. and G.M.Jenkins, 1976.Time Series Analysis : Forecasting and Control. Holden Day Inc. San Francisco,CA.

Cook, D.F. and Wolfe, M.L., 1991, “A back-propagation neural network to predict average air temperatures”, AI Applications, 5, p.p.40-46

H. Liu and **V.** Chandrasekar, Classification of Hydrometeor Type Based on Multiparameter Radar Measurements: Development of a Fuzzy Logic and Neuro Fuzzy Systems and In-situ Verification, *J. Atmos. Ocean Tech.,* Vol. 17, p.p.140-164, 2000.

Hu, M.J.C., 1964, Application of ADALINE system to weather forecasting, Technical Report, Stanford Electron.

Jang, J.S.R. (1993). ANFIS: Adaptive network-based fuzzy inference system. IEEE Trans. Syst., Man, Cybern., 23, p.p.665–685.

Jang.R, “Neuro-Fuzzy Modeling: Architectures, Analyses and Applications”, Ph.D. Thesis, University of California, Berkeley, 1992.

Ozelkan, E. C., F. Ni, and L. Duckstein, 1996: Relationship between monthly atmospheric circulation patterns and precipitation: Fuzzy logic and regression approaches. Water Resour. Res., 32, p.p.2097–2103.

Pasero,E., Moniaci,W, “ClMSA 2004 - IEEE lntemanonal Conference on Compumional Intelligence for Measurement System and Applications”,Boston. MA, USA, 14-16 July 2004 Artificial Neural Networks for Meteorological Nowcast

Rank,E.,2003 .Application of Bayesian trained RBF Networks to nonlinear time-series modeling. Signal Processing, 83: p.p.1393-1410.

Assoc. prof. Mehmet Tektaş at the Marmara University, Vocational School of Technical Sciences.
Research Areas: Fuzzy Logic,Artificial Neural Networks ,Forecasting, Statistic, Optimization
Adress: Marmara University,
Vocational School of Technical Sciences,
34722 Istanbul, Turkey.
Tel: +902164182504
Fax: +902164182505
E-mail: **tektas@marmara.edu.tr**

Orų prognozavimas ANFIS ir ARIMA modeliais. Stambulo pavyzdys

Mehmet Tektaş
Marmara universitetas, Profesinė techninių mokslų mokykla, Stambulas, Turkija

(gauta 2010 m. vasario mėn.; atiduota spaudai 2010 m. kovo mėn.)

Straipsnyje pateikiama lyginamoji orų prognozavimo statistinių ir neuroninių – neraiškiųjų (*angl. neurofuzzy*) tinklų analizė. Sudarant modelius, naudoti devynių metų (2000–2008 m.) Stambulo duomenys: vidutinė dienos temperatūra, oro slėgis ir vėjo greitis. Taikyti adaptyviosios neuroninės-neraiškiosios išvedimo sistemos (ANFIS) ir autoregresinio integruoto slankiųjų vidurkių metodo (ARIMA) modeliai. Atlikta ANFIS ir ARIMA modelių lyginimo analizė pagal MAE ir RMSE- R^2 kriterijus rodo, kad ANFIS patikimesnis.