Acta Universitaria
DIRECCIÓN DE APOYO A LA INVESTIGACIÓN Y AL POSGRADO

Forecast of Wind Speed with a Backpropagation Artificial Neural Network in the Isthmus of Tehuantepec Region in the State of Oaxaca, Mexico.

Ernesto Cortés Pérez*, Airel Nuñez Rodríguez*, Rosa Edith Moreno De La Torre*, Orlando Lastres Danguillecourt* and José Rafael Dorrego Portela*

ABSTRACT

This paper presents the preliminary results of setting up an artificial neural network (ANN) of the feed forward type with the backpropagation training method for forecast wind speed in the region in the Isthmus of Tehuantepec, Oaxaca, Mexico. The database used covers the years from June 2008 - November 2011, and was supplied by a meteorological station located at the Isthmus University campus Tehuantepec. The experiments were done using the following variables: wind speed, pressure, temperature and date. At the same time were done seven tests combining these variables, comparing their mean square error (MSE) and coefficient correlation *r*, with data the predicting and experimental. In this research, is proposed a ANN of two hidden layers, for a forecast of 48 hours.

RESUMEN

Este trabajo presenta los resultados preliminares de la configuración de una red neuronal artificial (ANN), de tipo alimentación hacia adelante con el método de entrenamiento de retro-propagación para pronosticar la velocidad de viento en la región del Istmo de Tehuantepec, Oaxaca, México. La base de datos utilizada abarca los años comprendidos entre Junio 2008 - Noviembre 2011, y fue suministrada por una estación meteorológica ubicada en la Universidad del Istmo campus Tehuantepec. Los experimentos se realizaron utilizando las siguientes variables: velocidad del viento, presión, temperatura y fecha. Al mismo tiempo se hicieron siete pruebas combinando estas variables, comparando su error cuadrático medio (MSE) y el coeficiente de correlación r, con los datos de predicción y experimentales. En esta investigación, se propone una ANN de dos capas ocultas, para un pronóstico de 48 horas.

Recibido: 8 de Diciembre del 2011
Aceptado: 14 de Febrero del 2012

INTRODUCTION

Decisions concerning the programming and organization of conventional electric generating plants as well as marketing locations are typically made in the morning. These decisions are related to predictions of the amount of energy generated at a given wind speed within forecast time frames of at least 48 hours [1]. An accurate forecast of wind speed provides decisive information on the availability of wind energy during the two days following a forecast, and this facilitates the efficient supply of wind-generated energy via existing supply systems. The benefit is obvious. Since the control of the amount of energy supplied can be reliably balanced, the economic value of the wind-generated energy increases, and it contributes to making renewable sources of energy more competitive in comparison to conventional energy sources.

At present, wind energy is used principally to produce electric energy with wind generators. Wind generators generate approximately 1% of the energy consumed worldwide [2].

In order to take advantage of wind energy, it is important to be aware of daytime, evening and seasonal wind variations, such as the variation of wind speed relative to its height above the ground, the occurrence of wind gusts during brief periods, and the maximum values that occur in data history series. However, the principal inconvenience is the wind speed intermittence [3, 4, 5]. In this study is proposed a model for predicting wind-speed intermittence by

Palabras clave:
Predicción, viento, velocidad, red, neuronal, artificial.
Keywords:
Forecast, wind, speed, artificial, neural, network.

*Departamento de Ingeniería en Computación, e Instituto de Estudios de la Energía, UNISTMO, Universidad del Istmo, Campus Tehuantepec, Oaxaca, Ciudad Universitaria S/N, Barrio Santa Cruz, 4a. Sección Sto. Domingo Tehuantepec, Oax., México C.P. 70760 Email: {cope_144, aireln, remt_84, olastres, r.dorrego.63}@hotmail.com

using an Artificial Neural Network (ANN) type feed forward in the Isthmus of Tehuantepec region in the state of Oaxaca, Mexico.

To develop the experiments, was used a data base provided by a meteorological station located on the Tehuantepec campus of the University of the Isthmus (see figure 1). This data bank covers the period of June 2008 to November 2011.

The meteorological station takes wind-speed readings every minute. In this research project the data base was constructed by calculating ten-minute averages.

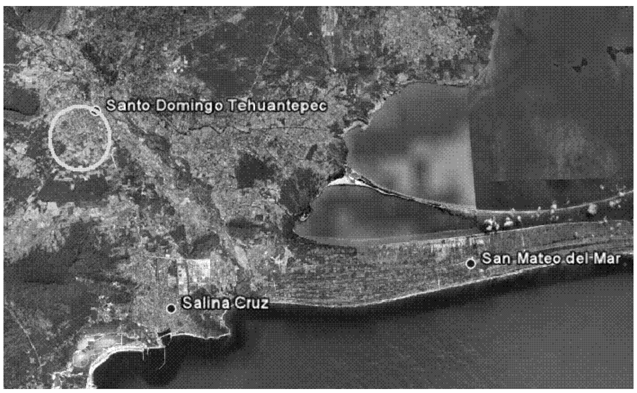

Figure 1 . Isthmus of Tehuantepec region.

RELATED WORKS

A number of studies have reported successful results from the application of neural networks to wind velocity forecasts. For example, in 2009 [2] Monfared, Rastegar and Hossein applied a strategy based on fuzzy logic and neural networks to forecast wind velocity. They employed statistical properties such as standard deviation, mean , and variable calculation relation gradient as neuro-fuzzy predictor model input. They did so by using real-time data taken obtained in Rostamabad, northern Iran from 2002 to 2005. Readings were taken in average intervals of 30 minutes. In 2009 Cadenas and Rivera used intervals averaging one hour to forecast wind velocity by using data collected by the Federal Commission of Electricity (Comisión Federal de Electricidad, CFE) during a period of seven years in the La Venta region of Oaxaca, Mexico [4]. This was done using a backpropagation y Madaline neural network with a mean square error measured at 0.0039 during the training process. In another study, Adbel Aal, Elhady y Shaahid [10] performed wind velocity forecasts by using tardy neural networks (TTND), principally for the maintenance of wind farms where researchers used a wind velocity data base with intervals averaging one hour that was compiled in the Dhahran region of Saudi Arabia from 1994 to 2005. The proposed model was evaluated with data from May of 2006, and the forecasting time was from 6 to 24 hours. For this, the network precision measurement parameter was a mean absolute error (MAE) of 0.85 m/s, and the correlation coefficient was equal to 0.83 between the actual value and that of the forecast. In 2009 [9] Sancho Salcedo, Ángel Pérez and Emilio Ortiz employed a multilayered network with the Levenberg-Marquadt training method in which the input variables were wind velocity chosen at two points of interest with wind direction at one point, temperature at one point, and solar radiation at two points. Hence, there are six values in the input layer, a hidden layer with six neurons with a Sigmoidal logarithmic activation function, and an output layer with one neuron. The forecasting time was 48 hours. The data had been collected since 2006 in Albacete province in southern Spain.

ARTIFICIAL NEURAL NETWORK (ANN)

Neuronal networks have come about as a means of developing systems that emulate the functioning of the human brain [6, 7, 8]. Nowadays, artificial neural networks can be utilized to carry out forecasts at varying degrees of success. ANN depend solely on initial data histories for training in the expectation that hidden dependencies are discovered that can be used to make future predictions. In other words, ANN is not represented as an explicit model. Rather, it is a black box that is capable of learning. The advantage of using ANN for prediction is that it learns from examples, and after it learns it is capable of generalizing as it encounters hidden non-linear patterns, even when there is noise in the collection of training examples.

Another advantage of neural networks is their capacity to deal with new data in novel situations, or with incomplete or erroneous data sets. One of the principal disadvantages is that they can learn valid dependencies only within a limited period, and in general prediction error cannot be estimated.

The most important characteristics of ANN are:

- **Learning ability.**
- **Generalization.**
- **Abstraction.**

The neuron is the basic element of the brain. The computational model for an artificial neuron is as follows:

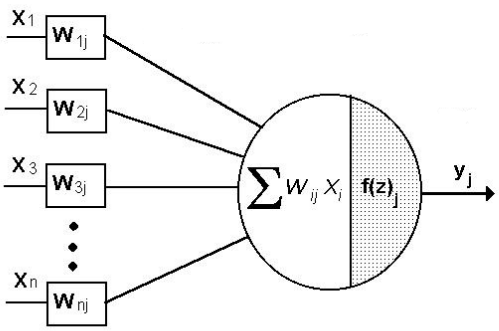

Figure 2 . Artificial Neuron Model.

Where:
X_i: are the input patterns.
$W_{i\ j}$: are weights assigned to each input pattern.
$\sum W_{ij}X_i$: is the total weight of input patterns.
$f(z)_j$: is the activation function.
y_j: is the output corresponding to each pattern.

INPUT PARAMETER ANALYSIS FOR THE FORECAST OF WIND SPEED

The wind speed vector (speed and direction) is surely the most important vector in terms of wind strength prediction.

For time series prediction one usually encounters single variable values in equidistant intervals with which prediction is attempted based on existing data histories. For this purpose the data bank for these variables must contain the information necessary to represent the behavior of test variable being studied (extreme values, possible cycles or periods, etc.)

Additional information can be found in others time series (temperature, pressure), this information can be used for a more exact prediction, and this information can be added using what is called variable intervention or indicator intervention, which provides additional information about the period in which the prediction is going to be made.

Nevertheless, more information does not always mean better prediction. At times this can worsen the teaching of the network, learning, generalization and prediction process. It is always necessary to generate information relevant to the ANN, whenever this is possible.

Some of the additional variables to consider were:

- **Temperature** (°C).
- **Pressure** (kPa).
- **Date** (mm/dd).

Correlation analysis of readings from August 2011 for the following variables:

- Wind speed / temperature.
- Wind speed / pressure.
- Temperature / pressure.

They indicate that exist some dependency (correlation). Figures 3-8 shows the correlations among these variables.

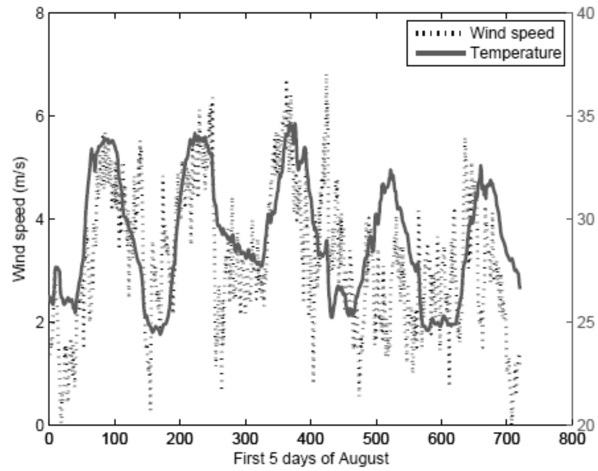

Figure 3 . Relationship between wind speed/temperature.

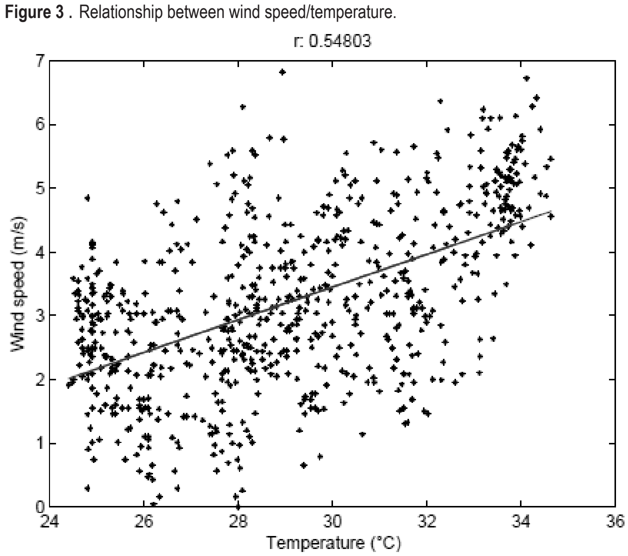

Figure 4 . Graph of relationship between wind speed and temperature, correlation r : 0.548.

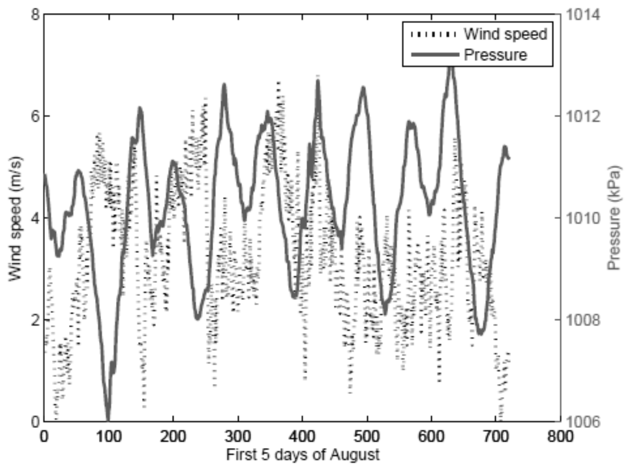

Figure 5 . Relationship between wind speed/pressure.

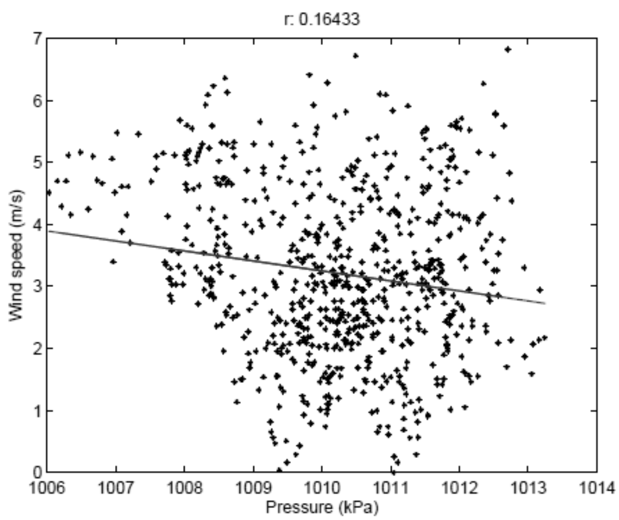

Figure 6 . Graph of relationship between wind speed and pressure, correlation r : 0.164.

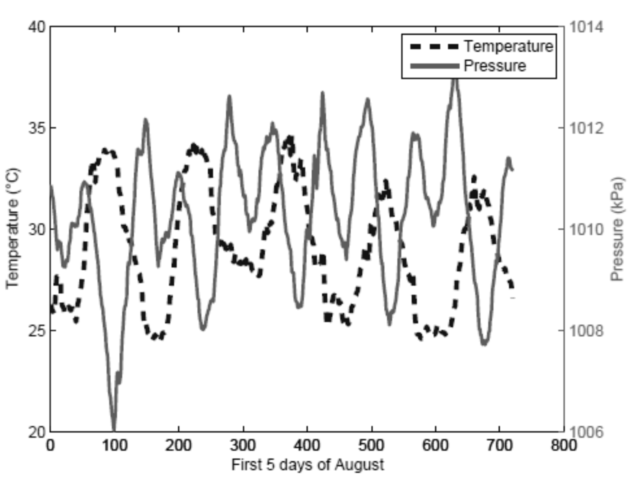

Figure 7 . Relationship between temperature/pressure.

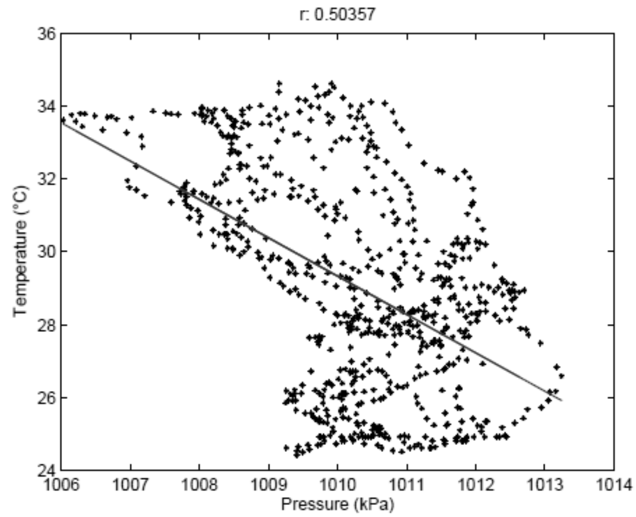

Figure 8 . Graph of relationship between temperature and pressure, correlation r : 0.503.

Table 1 .

Correlation among variables.

Variables	Correlation r
Wind speed / temperature	0.548
Wind speed / pressure	0.164
Temperature / pressure	0.503

These studies produced correlations among variables that demonstrate that the relationship between wind speed and temperature is closer than that of wind speed relative to pressure and of temperature relative to pressure. The table 1 describes the results of this analysis, which indicate that these variables are adequate for carrying out wind speed prediction.

The coefficient of determination *r* gives information about the correlation between variables, this coefficient is between [0, 1]. If the coefficient of determination is close to 1, it shows better correlation. On the other hand, this does not discount the use of variables such as humidity, solar radiation or wind direction.

PROPOSED MODEL: TIME SERIES MODELING FOR ARTIFICIAL NEURONAL NETWORK

Time series prediction using ANN consists of teaching the network the history of a variable in the chosen time interval while indicating the information to be learned in the future. Past data is entered as input, and future data to be used for represents ANN output.

An ANN is capable of predicting various kinds of data; nevertheless, this research project focuses on predicting wind speed time series in the Isthmus of Tehuantepec region, Oaxaca, Mexico. The wind-speed time series show the development of a value over time.

This value may be influenced by other factors such as pressure, humidity, temperature, radiation, among others.

In the time series it is necessary to be familiar with the values at point *t* in order to predict the future point of *t+P*. For this purpose, it is necessary to create a map with *D* examples of data points every Δ units in time. To forecast a future value *y(t + P)*, see Eq. 1.

$$y(t+p) = [x(t_{D-1} - \Delta), \ldots, x(t_1 - \Delta), x(t)] \quad (1)$$

Example for forecast 2 periods ahead with 4 sample data: $D = 4, \Delta = 1, P = 2$, see Eq. 2.

$$y(t+2) = [x(t_3 - 3) \quad x(t_2 - 2) \quad x(t_1 - 1) \quad x(t)] \quad (2)$$

Example for forecast 100 periods ahead with 4 sample data: $D = 4, \Delta = 1, P = 100$, see Eq. 3.

$$y(t+100) = [x(t_3 - 3) \quad x(t_2 - 2) \quad x(t_1 - 1) \quad x(t)] \quad (3)$$

The method for to build the matrix of examples W, is iterative, where examples number depend directly the series time size. The variable data taken by the meteorological station, covering the years June 2008 to November 2011 and are performed every minute, in this research, the matrix is constructed by calculating averages 10 minutes, the matrix has 118,694 examples with $D = 288, \Delta = 1 \quad y \quad P = 288$. The matrix can see in figure 9.

Wind Speed	1	2	3	. . .	118,692	118,693	118,694
Temperature	2						
Pressure	3						
Date	4						

June 2008 ⟵⟶ November 2011

Figure 9 . General layout of the array of examples W.

The architecture of the ANN, one can see in figure 10.

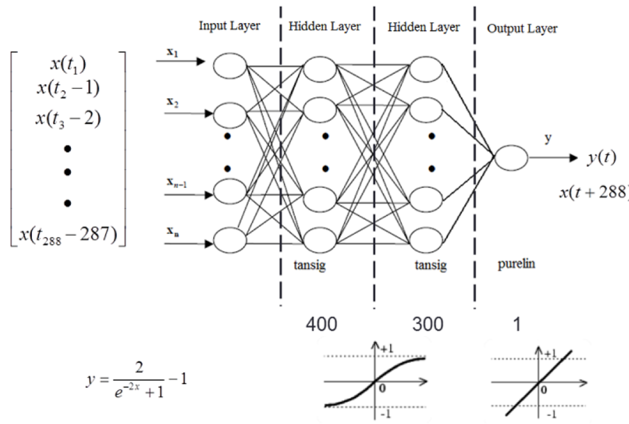

Figure 10 . Feed-forward ANN architecture.

Mapping of the input and output of ANN, one can see in figure 11.

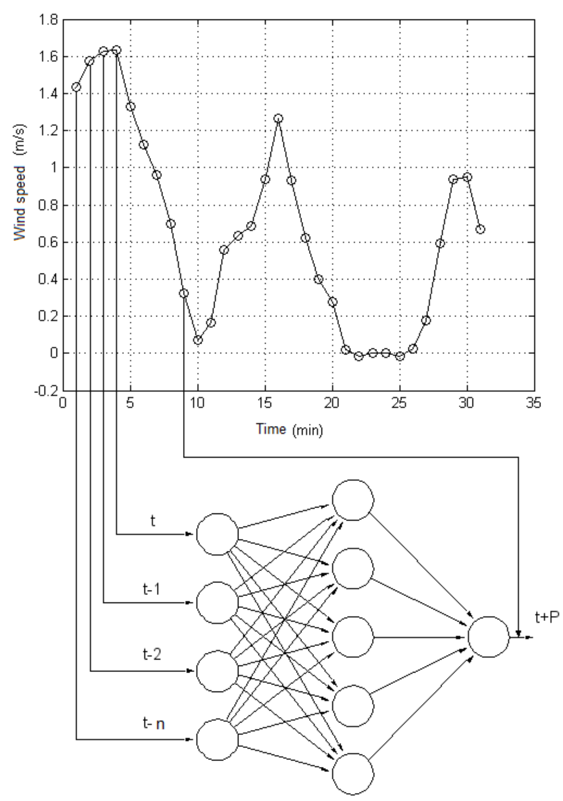

Figure 11 . Artificial Neuronal Network Input and output.

TRAINING AND PERFORMANCE OF ARTIFICIAL NEURONAL NETWORK.

A feed-forward network with a *backpropagation* algorithm is implemented using the MATLAB 7.10.0.499 toolbox for neural networks in order to develop the prediction model.

In general, there is no rule to determine the values of the network parameters. However, in accordance with various tests performed and the recommendations of different authors [2, 3, 5] [9, 10, 11, 12, 13], it was put into effect with two hidden layers. Both hidden layers had a sigmoidal tangent. A linear output layer had a single linear function neuron.

The structure proposed is given as 400-300-1. That is, there are 400 neurons in the first hidden layer, 300 in the second hidden layer and one neuron in the output layer. The sigmoidal tangent *(tansig)* function is employed in the hidden layers, and in the linear function output layer.

400 epochs were tested to obtain a quadratic average error decrease *(MSE)*, which became stable during the training with a tolerance of 0.00001. Figure 12 demonstrates the performance curve of the ANN.

If y_t is the current observation for period t and a_t is the forecast for the same period, then the error is defined as:

$$e_t = y_t - a_t \tag{4}$$

And the errors mean *(MSE)* is:

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \tag{5}$$

Where n is the number of periods in the time.

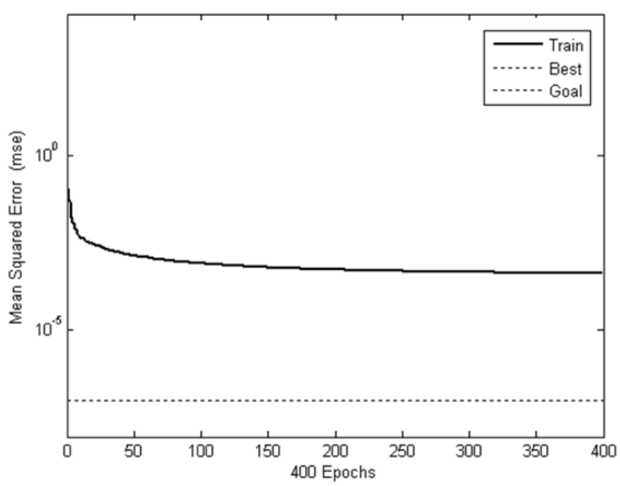

Figure 12 . Performance of Artificial Neural Network (error mse).

The accuracy of the network was tested with the four following variables:

- Date, *(D)*.
- Wind speed, *(Ws)*.
- Temperature, *(T)*.
- Pressure, *(P)*.

In different combinations, seven tests were carried out to forecast wind speed on September 2011. See table 2 for the results that were obtained combining the variables.

The parameters used during network training were:

- Training time.
- Error (*MSE*).
- Learning rate.
- Epochs.

In table 3 shows the parameters of training in each test done for the ANN. In figures 13 and 14 shows the results of the forecasting of the test number 7, which yielded the highest value for *r*, this with the *Wind speed (Ws), Date (D) and Temperature(T)*, variables.

Table 2 .

Comparison of forecast percentage r in seven tests.

Test num.	Variables (n)	Value for(D)	Value for(P)	r
1	(Ws)	288	288	0.479
2	(Ws-D)	576	288	0.459
3	(Ws-D-P)	864	288	0.529
4	(Ws-D-P-T)	1152	288	0.760
5	(Ws-T)	576	288	0.775
6	(Ws-P)	576	288	0.593
7	(Ws-D-T)	864	288	0.8

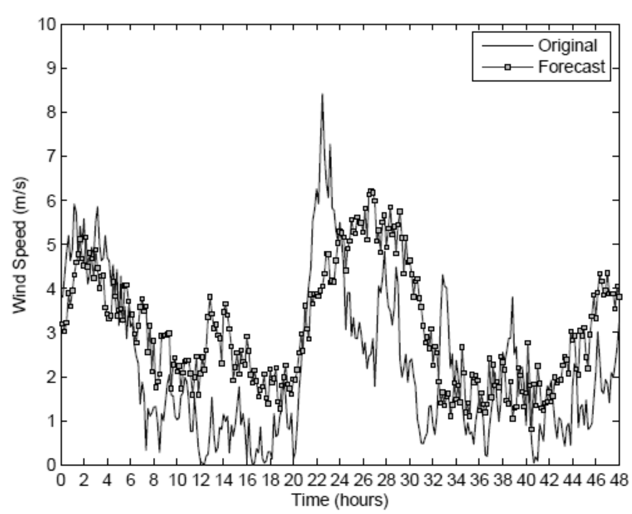

Figure 13 . Result forecast test number 7 (variables: Wind speed, date and temperature).

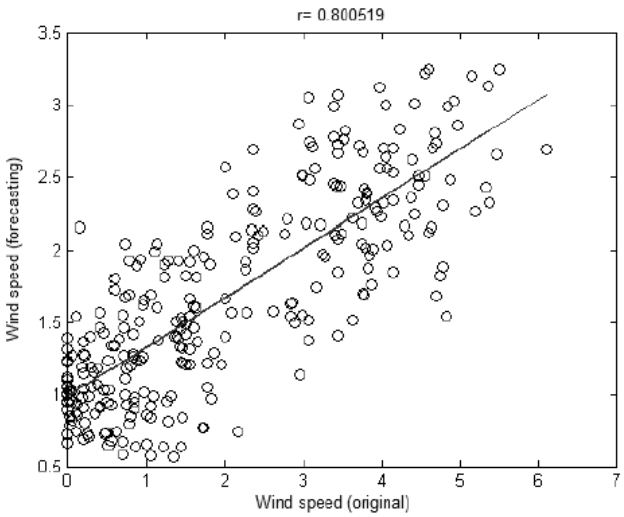

Figure 14 . Result forecast test number 7, correlation r=0.800519.

Table 3 .

ANN parameters and training time in each test.

Test number	Training time	Error (MSE)	Epochs
1	7.6 hours	0.0032	400
2	8.2 hours	0.0034	400
3	8.7 hours	0.0036	400
4	9.4 hours	0.0038	400
5	8.2 hours	0.0036	400
6	8.5 hours	0.0036	400
7	8.7 hours	0.0035	400

CONCLUSION

In the particular case of the Isthmus of Tehuantepec, Oaxaca, is presented a model that incorporates date and temperature variables. It has demonstrated improved sensitivity to adjustments in wind-speed forecasts, particularly in this research project.

It's worth mentioning that one of the characteristics that determine forecast accuracy is the modeling and quantity of data.

It has been demonstrated that there is not a good relationship between wind speed and variables such as pressure, where the similarity factor measurement produced a correlation r of only 0.2438. However, on the other hand, it was shown that wind speed/temperature and temperature/pressure variables had a good factor of similarity with a correlation r of 0.5034 and 0.4699 respectively. At the same time, it was necessary to add that the use of other variables has not been discounted.

It can also be concluded that an ANN with three layers in the order of 400-300-1 (Tangent Sigmoidal, Tangent Sigmoidal, Linear), respectively, seems to be a useful tool for the forecast of future data.

This research project obtained a forecast factor of correlation of r **= 0.80051**, for information used to predict wind speed in the two days following a forecast. This helped those in charge on a given shift to make decisions concerning the programming and organization of wind-powered electric generating plants. These decisions are related directly to the forecasts, which provide precise information regarding the availability of wind energy during the two days following prediction.

ACKNOWLEDGMENT

This work is supported by the project "Predicción del Recurso Eólico en el Istmo de Tehuantepec", FOMIX-CONACYT, FOMIX-CONACYT-GOB. DEL EDO. DE OAXACA, with the number 2009-C02-123395.

REFERENCES

[1] Lange, M., Focken, U. (2006). *Physical Approach to Short-Term Wind Power Prediction,* Springer.

[2] Monfared, M., Rastegar, H., Madadi, H. (2009). A new strategy for wind speed forecasting using artificial intelligent methods, *Renewable Energy,* 34:845-848.

[3] Bilgili, M., Sahin, B., Yasar, A. (2007). Application of artificial neural networks for the wind speed prediction of target station using reference stations data, *Renewable Energy,* 32:2350-2360.

[4] Cadenas, E., Rivera, W. (2007). Wind speed forecasting in the South Coast of Oaxaca, Mexico, *Renewable Energy,* 32:2116-2128.

[5] Carolin-Mabel, M., Fernandez, E. (2008). Analysis of wind power generation and prediction using ANN: A case study, *Renewable Energy,* 33:986-992.

[6] Jones, M. T. (2005). *AI Application Programming,* Second Edition, ed. Charles River Media.

[7] Jang, J. S. R. , Sun, C.T., Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence,* ed. Prentice Hall.

[8] Martin Del Rio, B. (2001). *Redes neuronales y sistemas borrosos,* ed. RA-MA.

[9] Salcedo Sanz, S., Perez Bellido, A.M., Ortiz Garcia, E.G., Portilla Figueroa, A., Prieto, L., Correoso, F. (2009). Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks. *Neurocomputing,* 72:1336-1341.

[10] Abdel-Aa, R.E., Elhadidy, M.A., Shaahid, S.M. (2009). Modeling and forecasting the mean hourly wind speed time series using GMDH-based abductive networks, *Renewable Energy,* 34:1686-1699.

[11] Ramirez-Rosado, I.J, Fernandez-Jimenez, L.A., Monteiro, C., Sousa, J., Bessa, R. (2009).Comparison of two new short-term wind-power forecasting systems, *Renewable Energy,* 34:1848-1854.

[12] Kavasseri, R.G., Seetharaman, K. (2009). Day-ahead wind speed forecasting using f-ARIMA models, *Renewable Energy,* 34:1388-1393.

[13] Cadenas, E., Rivera, W. (2009). Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks, *Renewable Energy* 34:274-278.