NEURAL NETWORKS ESTIMATE ATMOSPHERIC VARIABLES IN THE CARIBBEAN BASIN

3.4

Nazario D. Ramírez-Beltrán*
University of Puerto Rico, Mayaguez, PR

Amos Winter
University of Puerto Rico, Mayaguez, PR

Felipe Chong Rengifo,
University of Puerto Rico, Mayaguez, PR

and

Anthony Veneros
University of Puerto Rico, Mayaguez, PR

Nazario Ramirez Escalante
University of Puerto Rico, Mayaguez, PR

ABSTRACT

A new statistical method for downscaling is introduced. The method is based on the application of neural network models and kriging algorithm to estimated atmospheric variables in three-dimensional space. Artificial neural networks (ANN) were used to interpolate atmospheric data in the vertical domain while the kriging algorithm was used to perform horizontal interpolation. This methodology was used to convert a very limited radiosonde data set for the Caribbean basin to a gridded set of data at 11 vertical levels. The proposed estimation scheme can be used to establish the initial and boundary conditions to run a regional atmospheric modeling system (RAMS). Estimation of atmospheric variables can also be used to identify short- and long-term correlations between upper-air variables with surface climate events, such as hurricane tracking and intensity. Atmospheric variables estimated solely by radiosonde observations were compared with National Center for Environmental Prediction (NCEP) data. Sixteen days were randomly selected from two dry and two wet months in an arbitrary year, 1995. Estimation based on radiosonde observations was in agreement with NCEP data for the eleven mandatory pressure levels. Furthermore, cross-validation techniques with five years of radiosonde observations were used to test the proposed methodology and satisfactory results were obtained. Cross-validation results indicate that the proposed methodology is a potential tool to obtain estimation of atmospheric variables with a high resolution. The suggested estimation scheme can be improved by including the NCEP data as an additional source of information. The proposed technique can be used to increase resolution by incorporating temporary soundings that may be obtained during the generation of a hurricane. The algorithm can also include airplane information of critical events.

1. INTRODUCTION

Here we discuss a new statistical method for downscaling, to increase resolution of atmospheric variables The method is based on the application of neural network models and kriging algorithm to estimate gridded atmospheric variables in three dimensional space which are derived from a few irregularly distributed radiosonde observations.

The relationships between local climatological events and resulting large-scale phenomena are highly nonlinear, often chaotic, and change with atmospheric circulation. This creates the need for novel mathematical techniques. Artificial neural networks (ANN) were used to interpolate atmospheric data in the vertical domain while the kriging algorithm was found to be an efficient estimator to interpolate horizontal data. The Kriging algorithm is one of the most prominent techniques to perform spatial interpolation, and was introduced by Georges Matheron (1979). This algorithm weights observations based on the spatial variability of the process realization. This technique has also been expanded to interpolate simultaneously several variables (Carr et al. 1985). We used the Kriging algorithm to perform horizontal interpolation of atmospheric variables because this method provides consistent and efficient estimates in the horizontal domain. This algorithm is also optimal in the sense that the interpolation error at the actual observed value is zero.

* Nazario D. Ramirez, Univ. of Puerto Rico, Dept. of Industrial Engineering, Mayaguez, PR 00680; e-mail: nazario@ece.uprm.edu

ANN is an emerging technique that can be used to classify a set of apparently non-related climatic variables or to model the multivariate inputs and outputs of a dynamic system. ANN has the advantage of learning from data that exhibits either highly nonlinear relationship since the inherent transfer functions are nonlinear in nature. To develop a nonlinear model via ANN reduces the problem of identifying the appropriate transfer function and the number of neuron in the hidden layer.

Several researchers have reported successful applications of the neural networks methodology to atmospheric sciences and climate dynamics. Snell (et al. 2000) pointed out that many climate studies require generating estimates of climate variables at a given location based on values from other locations. They suggest a methodology based on ANN to estimate temperatures for some locations given information from a lattice of surrounding locations. They found that ANN outperforms some conventional methods such as spatial average, nearest neighbor, and inverse distance method. Malmgren and Winter (1999) used rainfall and temperature data along with principal component analysis and neural network to identify climatic zones on Puerto Rico. Aviolat (et. al 1998) apply an ANN to describe the creation of clouds at different layers. Grecu and Krajewski (2000) proposed a methodology for detecting anomalous propagation echoes in radar data. They used a neural network model to classify the base scan radar echo into the anomalous propagation echoes or rain classes. They pointed out that the neural network approach presents a conceptual simple yet rigorous way to address the problem of detecting anomalous propagation echoes . Tangang (et al., 1998) applied neural network methodology to forecast the sea surface anomaly on three regions: el Niño 4, el Niño 3.5, and el Niño 3. Hesieh (2001) applied nonlinear principal component analysis (NLPCA) based on a neural network model to the Tropical Pacific SST. Hesieh and Wu (2001) applied nonlinear singular spectrum analysis (NLSSA) to the Southern Oscillation Index and found that the first mode exhibits a 52-month period while the second mode shows a 32-month period. They showed that NLSSA is superior to singular spectral analysis and to classical Fourier spectral analysis.

In this paper artificial neural networks and kriging algorithm were implemented to model a nonlinear atmospheric dynamic system in the Caribbean basin among eleven mandatory pressure levels to estimate atmospheric variables from the sea surface up to 17 km of elevation. The estimation scheme can be used for other purposes e.g. to derive correlations between upper-air variables with rainfall and temperature processes, and to develop initial and boundary conditions to run a regional climate numerical model.

2. METHODOLOGY

The estimation scheme includes three main tasks. (1) A quality control system that is designed to identify incorrect values and replace them with reasonable estimates. (2) The Kriging algorithm was used to perform horizontal interpolation for each atmospheric variable at each pressure level. (3) Neural network models were used to perform nonlinear vertical interpolations to properly estimate the atmospheric variables in the Caribbean.

2.1 Quality Control System.

Upper air observations began in the 1940's but a large numbers of observations became available only in 1957. Today, 90 countries operate about 1000 radiosonde stations that observe upper air parameters up to four times per day at internationally agreed-upon times. The study Caribbean region (10°N – 30°N and 60°W – 90°W) includes 67 radiosonde stations. The atmospheric variables measured by radiosonde are known as raob data and include the following variables: pressure (mb), geopotential height (m), air temperature (C^o), dew point (C^o), wind direction (degrees), and wind speed (m/s). Raob data occasionally exhibit assignable errors, whose origins are either instrumental or man made. A quality control system was designed to identify and correct the errors by studying a representative random sample. The selected random sample was drawn from five years of raob data to identify the appropriate percentiles. Four days of every month were randomly selected from the following years: 1994, 1995, 1998, 1999, and 2000. 1996 and 1997 were unavailable at the time the study was performed. Of the 67 radiosonde stations which were studied only seven are consistently working and about 10 stations provide raob data for any give day. Therefore, 240 days were selected that correspond to about 2400 soundings, which where taken at 12:00 UTC.

The Box-and-Whisker plot was used to identify the outliers in every variable for each pressure level. The values exhibited by the whisker plot suggested the 0.2 percentile could be used as a threshold to eliminate the instrument or the man made errors. Figure 2 exhibits the Box-and-Whisker plot for the temperature at 100 mb, which indicates that there exist three assignable errors in the upper level and none in the lower level. The assignable errors can be eliminated by selecting the 99.8 percentile in the upper level. Therefore, the corresponding thresholds for these variables are: the maximum = –579 (tenths of C^{o}) and the minimum = –890 (tenths of C^{o}).

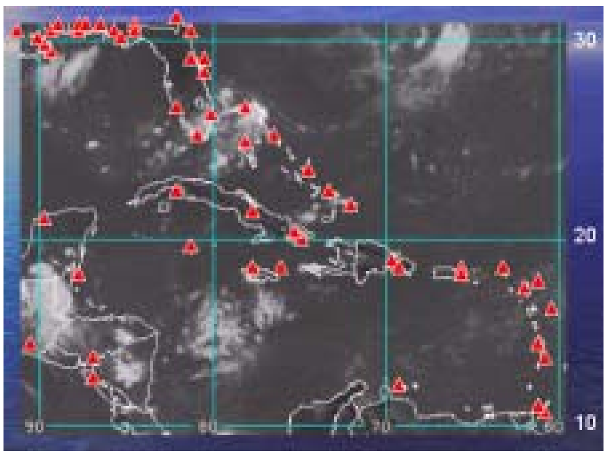

Figure 1. Location of Radiosonde Stations.

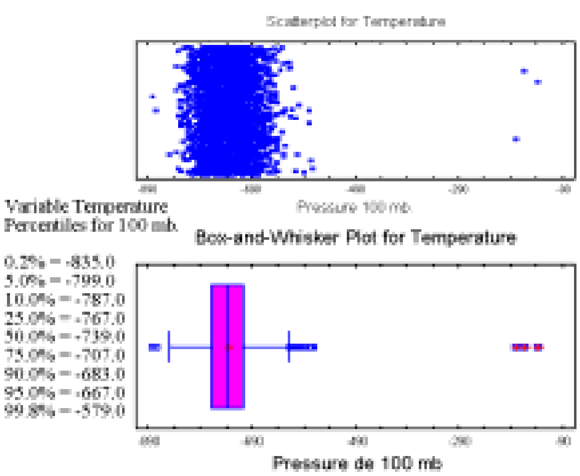

Figure 2. Threshold: Air Temperature at 100 mb

Upper and lower threshold for each variable and for each pressure level were identified by means of the Box-and-Whisker plot and a matrix that contains the upper and lower levels was developed. The quality control system consists of processing the raw data by the corresponding thresholds to identify and remove the assignable errors. In addition, the quality control system removes data files that contain less than 50% of good data

2 Horizontal Interpolation.

Radisonde observations for a given day were organized and the Kriging algorithm was used to perform horizontal interpolation up to the required grid size. This algorithm generates gridedd estimation for each variable at every pressure level (Bras, 1981; Federov, 1989). The implementation of the Kriging algorithm can be summarized as follows.

Let *"A"* be a set of known points $A = \{(x_i, y_i, z_i),\ \ i = 1, \ldots, n\}$ and each point is of the form (x_i, y_i, z_i), where x_i and y_i are the coordinates of the known point and z_i is the corresponding variable value. Let *B* be a set of unknown points $B = \{(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i),\ \ i = 1, \ldots, m\}$, where $(\tilde{x}_i, \tilde{y}_i)$ are the coordinates of unknown points and $\tilde{z}_i$ is the variable value to be estimated. The unknown values can be estimated as follow:

$$\tilde{z}_j = \sum_{i=i}^{n} w_{ij} z_i \quad j = 1, \ldots, m \tag{1}$$

where the weight values w_{ij} can be computed by solving the following system of equations:

$$\sum_{k=1}^{n} u_{ik} w_{kj} + \lambda_j = v_{ij} \quad i = 1, \ldots, n \quad j = 1, \ldots, m \tag{2}$$

$$\sum_{k=1}^{n} w_{kj} = 1 \quad j = 1 \ldots, m \tag{3}$$

where u_{ij} is the distance from the known point (x_i, y_i) to the known point (x_j, y_j) and v_{ij} is the distance from the known point (x_i, y_i) to the unknown point $(\tilde{x}_j, \tilde{y}_j)$, and λ_j is an additional parameter to be estimated from data.

The Kriging algorithm was implemented to estimate the atmospheric variables in a grid of one degree at each mandatory pressure level. The Kriging algorithm generates 600 grid points for each atmospheric variable at each pressure level, and each estimate is further used to derive the vertical interpolation. Estimation was performed at the studied area defined by 10°N to 30°N and 60°W to 90°W.

2.3 *Vertical interpolation*.

The gridded points generated by the Kriging algorithm were used to design a neural network model to perform a nonlinear vertical interpolation scheme. A feedforward neural network model was selected to express the relationships among climatic variables estimated at eleven pressure levels. Neural network models were chosen because successful applications have been reported in modeling highly nonlinear relationships among the variables of a dynamic system (Hesieh, 2001; Snell, 2000; Hagan et al., 1996; Ramirez-Beltran 1999a, 1999b, 2000a, 2000b, 2002).

Artificial neural network models were designed to express the relationships among the inputs and outputs of atmospheric dynamic systems. The first two modes of atmospheric variables from eleven mandatory pressure levels were used to train a neural network to obtain estimation at high resolution of the following atmospheric variables: geopotential height, air temperature, dew point, wind direction and wind speed. Air temperature and dew point were further combined to estimate relative humidity. Wind direction and wind speed also were used to estimate the U and V wind components. The designed scheme has the capability of estimating the mentioned atmospheric variables from the sea surface up to 17 km of elevation, (ie. from 1000 mb to 100 mb). Five years of radiosonde observations and cross-validation techniques were used to assess the capabilities of the proposed estimation scheme.

A feedforward neural network model is characterized by receiving input information to accomplish a modeling identification task without processing feedback information. The training patterns are presented to the network model several times until eventually the algorithm determines the optimal weights and biases that minimize the deviation between the network outputs and the established targets. The feedforward neural network model uses the backpropagation algorithm as the learning rule, which is based on the steepest descent algorithm. The backpropagation algorithm requires much more effort than the steepest descent on computing derivatives since the backpropagation was designed to work with nested functions and consequently it demands the use of the chain rule to compute the partial derivatives included in the sensitivity of the network. The errors are used to modify the searching direction and the gradient is computed at each layer starting from the last layer and finishing with the first layer, this is the reason for the backpropagation name.

A biological neuron is represented by a mathematical model, which is called an artificial neuron (Hagan et al, 1996). Usually, a set of interconnected artificial neurons is called a layer and a set of interconnected layers is a neural network model. The number of neurons in the hidden layer was identified by maximizing the prediction capability of the neural network model and the best results were found with five neurons in the hidden layer. The designed neural network has two layers and twenty-two neurons in the output layer since there are 22 outputs variables.

A neural network model was trained for every grid point to obtain simultaneously vertical interpolations for five variables. A carefully selection of the input patterns is required to train a neural network. The five-raob variables were organized as follows:

$$X_i = \begin{bmatrix} H_{ij} & T_{ij} & DP_{ij} & D_{ij} & V_{ij} \end{bmatrix} \quad i = 1,\ldots,\tau \quad j = 1,\ldots,L \tag{4}$$

where X_i represents the upper-air information for the i^{th} grid point. The variables H_{ij}, T_{ij}, DP_{ij}, D_{ij}, V_{ij}, represent geopotential height, air temperature, dew point, wind direction and wind speed, respectively, at the i^{th} grid point and at the j^{th} pressure level. L is the total number of pressure levels, and τ is the number of grid points in the studied area. In order to reduce the number of input patterns the principal modes of the i^{th} grid point were computed as follows:

$$Z_i = X_i U_i \qquad i = 1,\ldots,\tau \tag{5}$$

where U_i is the orthogonal matrix for the i^{th} grid point and its elements are the eigenvectors associated to the matrix X_i.

To accomplish the interpolation strategy the input patterns were selected and organized as follows:

$$\mathbf{P_i} = \begin{bmatrix} P_i \end{bmatrix} \quad \text{and} \quad \mathbf{T}_i = \begin{bmatrix} \tilde{Z}_i \end{bmatrix} \qquad i = 1,\ldots,\tau \tag{6}$$

Where $\mathbf{P}_i$ and $\mathbf{T}_i$ are the input and output patterns of the i^{th} grid point. P_i is the pressure level associated to the i^{th} grid point and $\tilde{Z}_i$ corresponds to the first two columns of Z_i. $\mathbf{P}_i$ is a vector of eleven elements and $\mathbf{T}_i$ is a vector with twenty-two elements. A neural network model was trained using the Backpropagation (BP) and the Levenberg-Marquardt (LM) algorithm. The LM algorithm improves significantly the performances of the BP algorithm (Hagan and Menhaj, 1994). The main drawback of the LM is the large memory that is required. This is one of the major reasons for training a ANN at each grid point. It should be noted that the neural network model identifies at each grid point a statistic model that correlate information from eleven pressure levels. The neural network model estimates climatic variables with one degree of resolution at any pressure level that falls in the range: 1015 mb to 100 mb. Figure 3 exhibits the neural network modeling process.

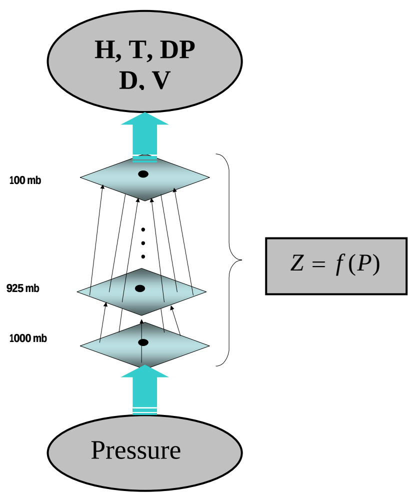

Figure 3. Neural Network Modeling Process

3. RESULTS.

Atmospheric estimates derived from this work were compared with data from the National Center for Environmental Prediction (NCEP) to test the veracity of our method. Sixteen days were randomly selected from two dry and two wet months in 1995. Estimations based on radiosonde observations were in agreement with NCEP data for the eleven mandatory pressure levels. Figure 4 shows the estimates of wind speed at 200 mb on February 15, 1995 at 12 UTC. Figure 5 exhibits estimates of wind speed obtained solely on radiason data and using NN estimation. Figure 5 exhibits comparisons of wind speed at 200 mb between NCEP and neural networks estimations. These figure shows similar results, however, the there are some differences since NCEP is given at 2.5x2.5 degrees on the horizontal, while neural networks estimates are given at 1x1 degree. This exercise indicates that . comprehensive estimates can be obtained by including NCEP data and radiosonde date to derive upper air estimates at any pressure level and at one degree of resolution.

One of the major advantages of the ANN scheme is that additional information from local sounding and from . airplanes can be incorporated into the ANN algorithm to improve resolution of upper air variables. A second advantage is the ANN can be used to perform short-term prediction of the upper air behavior.

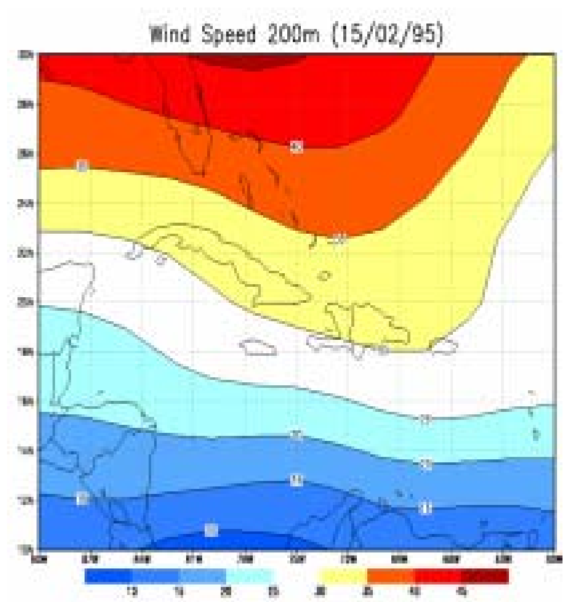

Figure 4. NCEP Estimates (wind speed 200 mb)

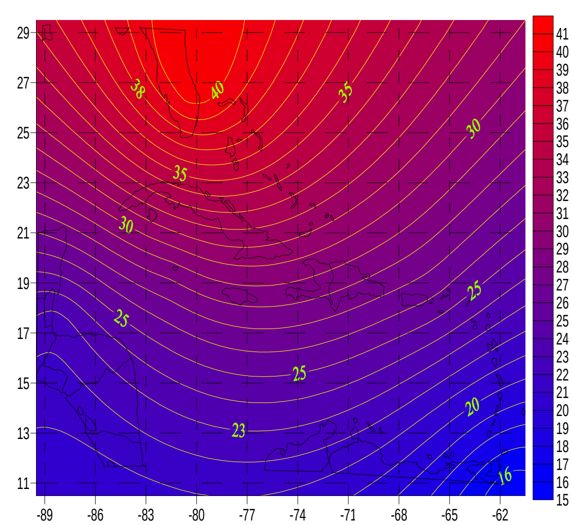

Figure 5. NN Estimates (wind speed 200 mb)

Hurricane Marilyn that occurs on September 1995 was used to illustrate one application of the ANN estimation scheme. Radiosonde and NCEP data were used to estimate upper air variables that were very close to the center of the storm. A . 10 degrees square with the center located over the storm was estimated every six hours. It has been shown that the dip layer is a better source than the individual layers to estimate the possible track of a hurricane. A dip layer was developed by using the weighs suggested by Neumann (1988) Thus, a dip layer of geopotential height for U and V components were computed across the hurricane track. Using the location of the storm as the origin, four quadrants were defined. Each quadrant has 25 grids and the first mode of the dip layer for each quadrant and for each variable was computed. The displacement of the storm was developed every six hours on longitude and latitude. It was studied whether or not there is any relationship between the first mode of the dip layers and the displacement of the storm.

It was found that there is a strong correlation between the dip layer of geopotential height, and wind components with the displacement of the hurricane. Figures 6 and 7 show the model fitting between the longitude and latitude displacements, respectively.

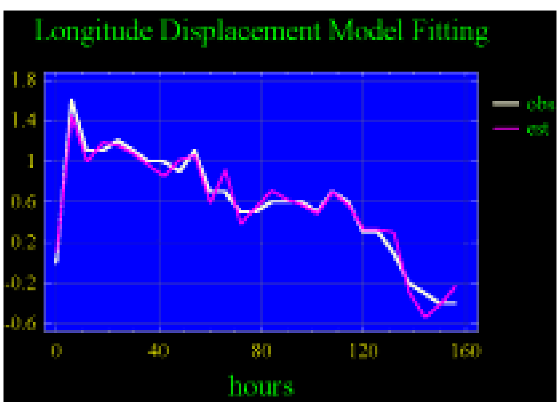

Figure 6. Longitude displacement correlation.

Figures 6 and 7 show in the horizontal axis time of the storm in hours and in the vertical axis the difference in Longitude (or Latitude) between two consecutives observations. The error in locating the center of the storm based on the moving gridded estimates can be determined by using the following equation

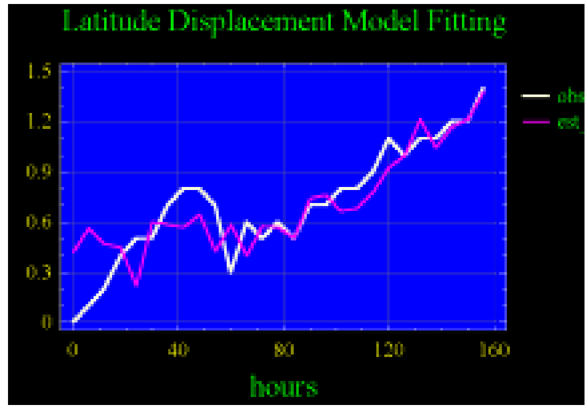

Figure 7. Latitude displacement correlation

$$e_t = 60\cos^{-1}\left[\sin(La_{1,t})\sin(La_{2,t}) + \cos(La_{1,t}) * \cos(La_{2,t})\cos(Lo_{1,t} - Lo_{2,t})\right] \qquad (7)$$

where e_t is the distance between two points at time t, the two points are defined by the observed and the estimated points. $La_{1,t}$, $La_{2,t}$, are the observed and estimated latitude at time t, respectively. $Lo_{1,t}$, $Lo_{2,t}$, are the observed and estimated longitude at time t, respectively

The distances provided by equation (7) are given in nautical miles. The estimated average error in locating the center of the storm was 18.55 Km. It should be noted that the location of a hurricane can be predicted if the ANN scheme is able to predict the upper air in a short-term period.

.

4. DISCUSION AND CONCLUSIONS.

The proposed methodology is a potential tool to mitigate the availability of very limited radiosonde information. At a given point in time the number of weather stations that provide sounding information in the Caribbean is approximately 10. We are planning to include NCEP data to improve our estimation scheme. In this preliminary work we did not exploit the use of satellite data. It should be noted that with an additional computational effort the proposed methodology could be extended to predict atmospheric profiles including satellite data.

One of the major contributions of this research effort is the design of a neural network algorithm to estimate climatic variables at any pressure level that falls in the range of 1050 mb to 100mb. Preliminary results show evidence that the neural network models based on radiosonde observations and other sources of information are capable of estimating upper-air variables. Comparing network predictions with real observations provides an assessment of the capability of the neural network. It was noted that air temperature and dew point were predicted with high accuracy. The average absolute errors for those variables were 1.19 ^{o}C and 1.75 ^{o}C, respectively. The average absolute error for geopotential height was 24.8 m, the wind speed average absolute error is in the range of 2.55 m/s whereas the wind direction average absolute error was 45 degrees. In summary, prediction errors are small enough to indicate that the proposed methodology provides satisfactory results. It should be noted that relative humidity could also be predicted based on air temperature and dew point predictions. Therefore, the proposed neural network approach is a potential tool to develop the atmospheric boundary conditions for a regional numerical model.

The future work of this research would be to develop a short-time prediction scheme The prediction scheme will provide the possibility of predicting the hurricane tracks and hurricane intensity.

ACKNOWLEDGMENTS

The authors appreciate the support from NASA EPSCOR grant NCC5-595 and from the University of Puerto Rico.

REFERENCES.

Aviolat, F., T. Cornu, and D. Cattani, 1998: Aoutomatic Clouds Observation Improved by an Articicial Neural Network., *Journal of Atmopheric and Oceanic Technology*, Vol. 15, 114-126.

Bras, R. and J. Kafristas, 1981: *The Practice of Kriging.* Cambridge, Massachusetts, MIT Press.

Federov, V.V., 1989: *Kriging and Other Sapatial Field Characteristics: with Special Reference to Environmental Studies*, Great Britain, Pergamon Press.

Grecu, M., and W. Krajewski, 2000: An Efficient Methodology for Detection of Anomalous Propagation Echoes in Radar Reflectivity Data Using Neural Networks. *Journal of Atmospheric and Ocean Technology*, Vol. 17, 121-129.

Hagan, M.T., H.B. Demuth, and M. Beale, 1996: *Neural Network Design*, PWS Publishing Co.

Hagan, M.T., and M.B., Menhaj 1994: Training Feedforward Networks with the Marquardt Algorithm, *IEEE Transactions on Neural Networks*, Vol. 5, No. 6, 989-993.

Malmgren, B.A., and A. Winter, (1999). Climate Zonation in Puerto Rico Based on Principal Components Analysis and an Artificial Neural Network, *Journal of Climate*, Vol. , 977-985.

Matheron, G., (1971). The Theory of Regionalized Variables and its Applications: *Centre de Geostatistique,* Fontainebleau, France.

Neumann, C.J., (1988). The National Hurricane Center NHC83 Model, *NOAA Technical Memorandum NWS NHC 41.* National Hurricane Center, Coral Gables, FL

Ramírez-Beltran, N.D. and H. Jackson, 1999a: Applications of Neural Networks to Chemical Process Control, *Computers & Industrial Engineering*, Vol. 37, No. 1-2, 387-390.

Ramírez-Beltran, N.D., 1999b: Neural Networks to Minimize Lot-End Costs. *Proceedings of the 3rd International Conference on Engineering Design and Automation*, Vancouver, Canada, 352-358.

Ramirez-Beltran, N.D. and H. Rodriguez, 2000a: A Neural Network Approach to Estimate the Chemical Activity Coefficients. *Proceedings of Group Technology and Cellular Manufacturing*, World Symposium 349-354.

Ramirez-Beltran, N.D., and L. Olivares, 2000b: Neural Networks and Accelerated Tests to Predict Shelf Life of Drug Products. *Proceedings of Group Technology and Cellular Manufacturing*, World Symposium 355-360.

Ramirez-Beltran, N.D., and Montes, J. 2002 Neural Networks to Model Dynamic Systems with Time Delays.*IIE Transactions,* Vol. 34, 313-327.

Snell, S.E., Gopal, S. and R.K Kaufmann, 2000: Spatial Interpolation of Surface Air Temperature Using Atificial Neural Networks: Evaluating their use for Downscaling GCMs. *Journal of Climate,* Vol. 13, 886-895.

Tangang, F.T., B. Tang, A. Monahan, and W.W. Hsieh, 1998: Forecasting ENSO Events: A Neural Network-Extended EOF Approach., *Journal of Climate*. Vol. 11, 29-41.