14A.7 UPPER AIR INFORMATION AND NEURAL NETWORKS TO ESTIMATE HURRICANE INTENSITY

Nazario D. Ramirez-Beltran* and Anthony Veneros
University of Puerto Rico

1. INTRODUCTION

Upper air information and artificial neural networks (ANN) are used to predict hurricane intensity in the North Atlantic basin. The proposed prediction scheme uses historical data to identify the analog hurricanes. Self organized ANN is used to identify the storm analogs to the current hurricane. The analogs are based on: Julian date, storm location, intensity, and direction. Twenty eight (1975-2002) years of climatology and persistent data provide preliminary information to identify the hurricanes that best resemble the behavior of the current storm. Once the analog hurricanes are identified the historical NCEP reanalysis data are used to extract synoptic variables along the storm track. Historical data from analogs and current radiosonde observations are used to estimate the synoptic variables of the current hurricane. Persistence, climatological and synoptic observations of the analog hurricanes and the information of the current storm are combined to create the initial training set for the ANN. The original meteorological variables are used to generate nonlinear transformations and an optimization algorithm is used to identify the variables that are best correlated with storm intensity. The best variables obtained from the optimization algorithm are divided into two parts. The first part is used to identify the optimal transfer function and the number of neurons in the hidden layer. The second part is used to identify an optimal initial point and to predict hurricane intensity.

2. SOURCE OF DATA AND PREDICTORS

Results of this study are based on the following sources of information: Hurricane best track, NCEP Reanalysis, and Radiosonde. The best track and NCEP reanalysis data provide historical information while radiosonde includes historical and current information. We are working with aircraft reconnaissance, and satellite data, which provide on-line upper air information to improve the working data set. Persistence predictors were derived from the best track and synoptic predictors from NCEP Reanalysis and Radiosonde data. Estimation of the synoptic observations for the current storm was estimated based on the synoptic observations of analog hurricanes. Radisonde observations were only used to improve the synoptic estimates when the current storm was located on the Caribbean, since Radiosonde observations were not provided in the open Ocean.

The selected predictors are the following:

- Persistence predictors: Julian date, storm intensity, direction of storm, location of center of the storm latitude, and longitude.
- Synoptic predictors: tropospheric vertical wind shear, maximum possible intensity, average momentum at 850 mb, average momentum at 200 mb, wind components at 850 mb, and 200 mb, k-index, and total totals.

Corresponding authors address: Nazario D. Ramirez, University of Puerto Rico, Dpto. Industrial Engineering, Mayaguez, PR, 00680, e-mail: nazario@ece.uprm.edu

3. METHODOLOGY

The suggested methodology to predict the hurricane intensity includes three major steps: 1) identifying analog hurricanes, 2) Identifying the best predictors, and 3) Designing a neural network to predict the hurricane intensity.

3.1 *Analog hurricanes*

The proposed prediction scheme uses historical data to identify the analog hurricanes. An analog is defined as a storm that best resemble the meteorological behavior of the current storm. The analogs are based on climatology and persistence variables. At the early stage of the storm, the analog set is derived based on the first 6 observations, and as soon as the storm life increases the number of observations for selecting analogs will also increase by one unit at a time up to 10 observations. When more than 10 observations are available the size of the moving window is maintained fixed to 10 observations. The period from 1975 to 2002 was selected to identify the analogs because most of the synoptic variables were complete

A self organized ANN with the Kohonen learning rule was designed to identify the storm analogs to the current hurricane. Ten neurons were used to characterize each persistence variable of each hurricane. The characteristics of the persistence variables of the current storm are the key to identify the potential analog set. The majority voting rule was used to identify a single code for each hurricane. All hurricanes that have the same code to the current storm were selected to be processed to a second ANN. The second ANN has the Kohonen learning rule and two neurons and again the assigned code to the current storm was used to select the set of analog storms. This process is repeated three times in order to increase the sample size and to derive a robust estimation, i.e., three sets of analog hurricanes were identified at every point in time.

3.2 *The best predictors*

Every set of analogs are organized to create an information matrix, which has 21 columns and the number of rows depends on the selected number of analogs. The first six observations are associated to the first analog. The next 6 rows represent the second analog and the last 6 rows represent the current hurricane. The number of rows of the information matrix will increase depending of the available number of observations. The first 5 columns of the information matrix correspond to the persistence variables and the last 14 columns correspond to the synoptic variables.

The information matrix is organized into the appropriate lagged order depending on the length of the desired lead time. Twelve different information matrixes were generated after applying a nonlinear transformation to each column of the original information matrix. The information vector is a column vector that contains the corresponding hurricane intensity of the analog storms included in the information matrix.

A regression approach was used to identify the best predictors, where the columns of the information matrix are the predictors and the information vector is the predictand. It should be noted that the number of predictors significantly exceeds the number of available observations. Therefore, the following strategy was used to identify the best predictors. The number of columns of the information matrix were divided into smaller groups such that the number of columns of each group is equal or less that n/5; where n is the number of available observations at specific time. The best predictors were selected for each group and the winners generated an improved information matrix. The improved information matrix was divided into smaller number of columns and again the best predictors were selected. This process was repeated over and over until the number of columns of the information matrix had n/5 columns or less. These columns were called the best subset for each information matrix.

It should be noted there were three sets of analogs and each set generates 12 different best subsets, then the total number of best subsets are 36, at each time interval.

3.3 *Neural network to predict hurricane intensity*

A feedforward neural network with Levenberg-Marquardt (LM) training algorithm was implemented to predict hurricane intensity. LM algorithm was used to increase the speed of convergence (Hagan, Demuth, and Beale, 1996). A cross-validation technique was used to identify the best transfer functions and the number of neurons in the hidden layer. Once the structure of the network was identified the best initial point was obtained after performing a random search exploration.

Each best subset was used to train three times an ANN and a hurricane intensity prediction was obtained. Since the ANN is a nonlinear optimization algorithm and is highly depending on the initial point, the ANN will provide different results after every training process. Thus an individual best subset was used to perform 3 predictions and its median is selected as the prediction from the best subset. The best prediction from the 36 subsets is selected based on avoiding the multicollinearity problem and minimizing the mean square of the prediction errors. The multicollinearity problem is detected by using the eigensystem criterion (Montgomery, et al, 2001).

4. RESULTS

Five hurricanes were arbitrarily selected to implement the proposed prediction method. The prediction errors are shown in Table 1. Performances of the prediction model are compared with the official prediction. The official prediction errors are smaller than the proposed method. We are exploring to include data from the Advance Microwave Sounding Units sensor to reduce prediction errors. Table 2 shows the comparison of the proposed model with existing hurricane intensity prediction algorithms. These results indicate that the proposed prediction scheme is a potential tool to improve the existing prediction algorithms.

Table 1. Prediction Errors (kt)

	Proposed Method		Official Results	
Hurricane	12 hrs	24 hrs	12 hrs	24 hrs
Isidore	12.08	20.97	4	16
Erin	7.49	14.1	5	10
Lili	11.58	12.5	9	13
Kyle	5.85	11.81	4	7
Felix	7.2	8.7	6	11
Average error(kt)	**8.84**	**13.616**	**5.6**	**11.4**

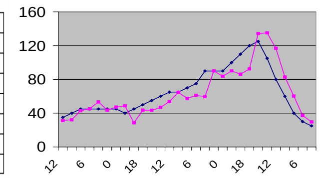

Figure 1. Hurricane Lili (2002)

Table 2. Predicted and observed intensity of Hurricane Lili (Sept 2002),

Hurricane Lili	**Forecast interval(hr)**	
Intensity errors (kt)	**12**	**24**
SHF5	13 (33)	17 (29)
SHIP	13 (33)	16 (29)
GFDI	14 (33)	18 (29)
AVNI	13 (33)	19 (29)
UKMI	15 (27)	20 (24)
Proposed Method	11.58(30)	12.5(25)
OFCL	9 (33)	13 (29)
Official mean (1992-2001)	7 (2198)	11 (1963)

5. CONCLUSIONS

Preliminary results show that the proposed prediction scheme is a potential tool to increase the accuracy in predicting hurricane intensity and reducing the prediction errors. The proposed prediction scheme is useful for the following lead time: 6, 12, 18, and 24 hours.

It has been shown that the design of an ANN requires a special procedure to obtain appropriate application, otherwise, the ANN will provide misleading results. The combination of an ANN with Kohonen rule is an appropriate technique to identify analog hurricanes. It has been confirmed that the LM algorithm significantly increases the speed of convergence of the regular backpropagation algorithm.

6. AKNOWDLEDGEMENT

This research has been supported by NASA-EPSCoR grant NCC5-595, by the NOAA-CREST grant NA17AE1625 and also by the University of Puerto Rico at Mayaguez. Authors appreciate and recognize the funding support from these institutions.

7. REFERENCES

Montgomery, D.C., E.A. Peck, G.G. Vining, (2001). *Introduction to Linear Regression Analysis*, 3rd Ed., John Wiley & Sons, Inc., New York.

Hagan, M.T., H.B. Demuth, M. Beale. (1996). *Neural Network Design*. PWS Publishing Co., Boston.