P1.91 DEEP LAYER OF UPPER AIR AND MULTIVARIATE TIME SERIES MODELS TO PREDICT HURRICANE TRACKS

Nazario D. Ramirez Beltran*, and Joan Manuel Castro
University of Puerto Rico

1. INTRODUCTION

A multivariate time series model was used to predict hurricane tracks in the North Atlantic basin. Two types of data sets were developed to build the prediction scheme. The first data set was used to identify analog hurricanes to the current storm and the second data set uses steering and synoptic observations along the storm track to predict the hurricane displacements. An analog is a hurricane that has a similar meteorological behavior to the current tropical storm. The data sets included climatology, persistence, and synoptic information during 19 years (1984 – 2002) and the data sources were the Hurricane Best Track, NCEP Reanalysis, and Radiosonde data. The Best Track and NCEP reanalysis data provide historical information while Radiosonde data include archived and current information.

2. METHODOLOGY

The suggested methodology includes three major steps: 1) Identify the upper air information to generate the initial conditions, 2) Develop the upper air time series along the storm track, and 3) Predict the hurricane displacement.

2.1 Initial Conditions

It was assumed that during the first 24 hours of a tropical storm there were five consecutives observations, which were obtained every 6 hours. The first observations were used to estimate the initial conditions for the development of a multivariate time series model. Climatology and persistence variables were used to identify analog hurricanes. Similarities among tropical storms were identified in terms of the following variables: the Julian date, Eastward and Northward displacement, sea surface temperature, and *hurricane intensity and direction.

A competitive artificial neural network (ANN) and a classification algorithm were used to identify the analog hurricanes to the current storm. The analog was identified by implementing the following strategy. The fisrt step consists of selecting 15 neurons with a competitive ANN and using the Kohonen learning rule. The competitive ANN assigns a code to each hurricane based on similarities identified in each variable at a single point in time. The hurricanes that have the same code to the current storm indicate strong similarities and are named potential analogs. The potential analogs are selected from a file that contains 100 hurricanes. The second step consists of processing the first five observations of the potential analogs by a second ANN and using 5 neurons. The hurricanes that fall in the same category of the current storm define the analog hurricanes. The third step involves filtering the analogs hurricanes by a classification algorithm whose criterion consists of minimizing the Euclidian distance between the current storm and the analogs. Thus, the storm that exhibited the minimum distance was called the analog hurricane to the current storm.

The locations of the first five points of the current storm were used to generate the expected initial 12 points; and all together made 17 points, which were used to extract the upper air data. The NCEP data of the analog storm along the 17 initial points were used to estimate the upper air of the current storm. Gridded upper air data (2.5x2.5 degree on the horizontal) included the following variables: geopotential height, air temperature, relative humidity, and wind vector components. These variables were extracted at the following levels: 1000, 850, 700, 500, 400, 300, 250, 200, 150 and 100 mb. These data were centered on the storm position with 10 degrees measured into the axial directions to East, West, South, and North. Spatial interpolation algorithm was used to obtain estimation at one degree of resolution. Thus, the extracted number of gridded points at each time interval was 441, i.e., the covered area at each time interval was a square of 21x21 degrees.

The information provided in the ten pressure levels was combined to generate a deep layer (Neumann, 1988). The values of each variable at each level were multiplied by some coefficients that provide more relevance to information located in the lower levels. Thus, the deep layer at each time interval and the location of the storm during the first 17 points was defined as the initial conditions for the current tropical storm. The initial conditions have enough information to develop a time series model and perform prediction at 6, 12, 18 and 24 hours in advance.

2.2 Develop the upper-air time series

The upper air for the current tropical storm was estimated every six hours. For instance, assuming that the location of the storm was known at a point 18, the upper air at point 18 was estimated based on selecting the appropriate analog. It should be noted that an analog was identified every six hours. The strategy for estimating the upper air at each point in time included three tasks: 1) Use the last observation to estimate the potential analogs, 2) Use a classification algorithm to identify the analog, and 3) Estimate the upper air for the current storm position.

1) A competitive ANN with 5 neurons was also used to identify the analog based only on the last observation of the current storm. Six climatology and persistence variables (as mentioned in the previous section) were used to identify the analog. 2) Twelve synoptic variables were added to the persistent variables for each analog hurricane. A classification algorithm based on minimizing the Euclidian distances between the current storm and the analogs was used. The winner storm from this process was called the analog hurricane. 3) NCEP data was used to obtain upper air centered on the location of the storm and a gridded deep layer was developed as explained in the previous section. The described process was applied at every six hours, as soon as new information become available. Dimensionality reduction was

*Nazario D. Ramirez-Beltran, Department of Industrial Engineering, University of Puerto Rico, Mayaguez, PR 00680, nazario@ece.uprm.edu

accomplished by using the first 10 principal components, which represents more than 90% of the total variance of the considered variables. An algorithm was used to identify the optimal lags and the best three variables that best explained the hurricane displacement.

2.3 Predicting hurricane displacement

Univariate and bivariate time series models were identified depending on the length of the available information (Brockwell, and Davis, 2002). Typically, the bivariate time series model requires more observation than the univarite model to build a time series model. Thus, at the early stages of the storm only the univariate models were identified, one for modeling zonal and the other for meridional displacements. The bivariate model expressed simultaneously the two displacements.

The best three predictors and the Hook and Jeeves algorithm were used to identify the structure of the time series model and the hurricane displacement was predicted (Reklaities, 1983). A different prediction model was built at every 6 hours and predictions were computed for 6, 12, 18 and 24 hours in advanced.

3. RESULTS

Arbitrarily three hurricanes were selected to implement the proposed algorithm. Tables 1 and 2 show the prediction error for the three hurricanes at 12 and 24 hours, respectively.

Table 1 Prediction Evaluation at 12 hours

	Marilyn		Danielle		Kyle	
Models	Error (nm)	Cases	Error (nm)	Cases	Error (nm)	Cases
Official	38	39	31	39	33	56
Time Series Model	34	37	40.3	39	40.3	57

Table 1 Prediction Evaluation at 24 hours

	Marilyn		Danielle		Kyle	
Models	Error (nm)	Cases	Error (nm)	Cases	Error (nm)	Cases
Official	71	39	47	37	59	51
Time Series Model	81.6	35	64.8	37	69.2	52

Figure 1. Hurricane Danielle Predicted and Observed Displacement.

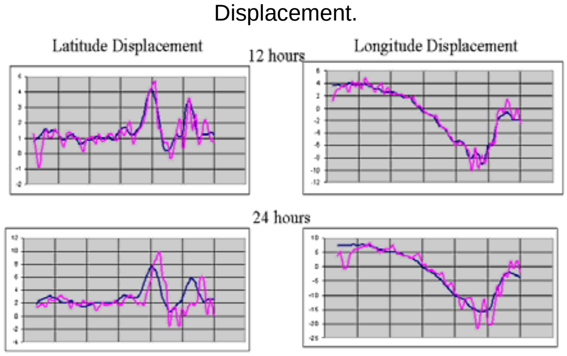
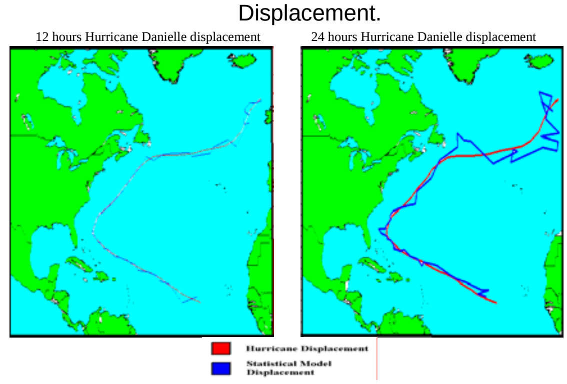


Figure 1 shows the zonal and meridional displacements of the Hurricane Danielle at 12 and 24 hours lead time. Continuous line indicates the observed values and the dot line represents the predicted values. Figure 2 shows the observed and predicted values of the track of Hurricane Danielle.

Figure 2. Hurricane Danielle Predicted and Observed Displacement.



4. CONCLUSIONES

Although, the prediction error of the proposed model is larger than the official prediction error, the proposed model generated prediction errors smaller than other available models, especially for the lead time of 6, 12, 18 and 24 hours. More exploration is required to claim that the proposed method is better than the existing prediction schemes. However, preliminary results show that the prediction scheme is a potential tool to increase the accuracy of predicting hurricane displacements.

A current study pretends to use observations from aircraft reconnaissance data and from the Advance Microwave Sounding Unit (AMSU) sensor to improve the upper air temperatures every twelve hours and consequently to improve algorithm prediction accuracy.

5. ACKNOWLEDGEMENT

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 want to appreciate and recognize the funding support from these institutions.

6. REFERENCES

Brockwell, P.J., and R.A. Davis. (2002). Introduction to Time Series and Forecasting. 2nd Ed., Springer-Verlag, New York.

Reklaitis, G.V., A. Ravindran, and K.M. Ragsdell, (1983). Engineering Optimization Methods and Applications, John Willey and Sons., New York.

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