Artificial Neural Network verses Multiple Linear Regression: Forecasting Hurricane Intensity

Abstract

There are several different types of models used to forecast future hurricane intensity. The simplest, and least skillful, is the statistical model. It depends on climatology of past hurricanes for its predictions. Dynamical models do the opposite. They create their forecasts based solely on current and future atmospheric conditions and the physics that govern the storms. Dynamical models tend to work well. The final category is the statistical-dynamical model which does a bit of both using linear regression. SHIPS is a model that uses this approach and it is one of the top two most skilled intensity models. The goal of this project is to first create a simplified model based on SHIPS using the most important and skillful predictors and then use a model which uses the same predictors but uses artificial neural networks instead of regression. These models will be tested against each other to determine which has higher skill in predicting future hurricane intensity

Description:

The project will result in the creation of two separate hurricane intensity prediction models. One is based on SHIPS, a statistical model developed in 1994. It is evaluated and updated at the end of each hurricane season, and quickly morphed into a statistical-dynamical model by 1997. The model employs statistical regression on a number of different variables, assigning them each a weight to determine the impact each variable should have on the forecast. As of 2003, up to 21 variables were in use.

The first model will be statistical-dynamical in nature, like SHIPS, and will use the listed variables, most of which are also used in the SHIPS model. The following variables were chosen due to their high skill in predicting hurricane intensity:

1) Initial maximum winds (initial intensity)

2) Intensity change over last 12 hours

3) Atmospheric conditions change over last 12 hours

4) Maximum potential intensity

5) 850-200 hPa vertical shear

6) 500-300 hPa relative humidity

7) Sea surface temperature

8) Best-track storm forecast location

9) 200 hPa divergence

These variables will give the model the cyclone's intensity trend, including how the dynamic variables have changed over the past 12 hours to help determine the future trend, as well as using important dynamic variables for both the current time and forecast times at the storm's future location to create a 48-hour intensity forecast with 6-hour forecast intervals. This model will be programmed in Python and use multi-threading to improve its efficiency and will use multiple linear regression to create its forecasts.

The second model will use neural networks to make intensity predictions. Artificial neural networks act similarly to linear regressions like SHIPS uses, but are more flexible and are able to model nonlinear functions with more skill. As SHIPS makes use of many variables and predicting hurricane intensity is about as nonlinear as possible, theoretically a model using neural networks will have greater skill than one using linear regression. By training the neural network using data of past storms with the same variables and the following intensity changes a model can be created to use those same variables to predict future intensity. Current data will then be fed into the trained network which will then create a 48-hour intensity prediction with 6-hour forecast intervals. It will be programmed using a neural network specific language which will be determined after more research.

Both models will be tested against the SHIFOR model for skill, a model completely based on climatology which is used as the skill baseline against which all intensity forecasts are compared. The two models will also be tested against each other to determine which of the two approaches to forecasting tropical cyclone intensity is superior. If there is no clear victor in terms of skill, the models will be analyzed to find their relative strengths and weaknesses.

As most tropical cyclones in the North Atlantic and East Pacific ocean basins occur in the months of August-October, this project will last through the fall 2015 semester and become my computer science senior project. The results of the model testing will be presented at the fall Symposium.

Timeline:
June 1st-July 9th (6 weeks): Data collection and integration into models
July 13th - August 7th (4 weeks): Algorithm creation and calibration using past cyclones
August-December: Testing and calibration of models using current cyclones as part of the computer science senior project.

Budget :
Student stipend: $1500

Resources

Mark DeMaria and John Kaplan. (1994) A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic Basin. *Wea. Forecasting*, **9**, 209–220.

Mark DeMaria and John Kaplan. (1999) An Updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and Eastern North Pacific Basins. *Wea. Forecasting*, **14**, 326–337.

Mark DeMaria, Michelle Mainelli, Lynn K. Shay, John A. Knaff, and John Kaplan.(2005) Further Improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543.