SSEC/CIMSS
Seminar

DENIS BOTAMBEKOV

Department of Marine & Environmental Systems
Florida Institute of Technology

Statistical Evaluation of the National Hurricane Center's Tropical Cyclone Wind Speed Probability Forecast Product

The National Hurricane Center (NHC) uses the Wind Speed Probability Forecast Product (WPFP) to predict categorical wind speed likelihood due to tropical cyclones, and provides information to the general public, government agencies and private companies. In this work the product is evaluated for all threatening TC events in Atlantic basin over 2004-2009 seasons, and 150 International, U.S. and open ocean locations for which the NHC WPFP issues forecasts. The discernment of risk is associated with the interval probability (IP) forecasts for the three wind speed categories (≥ 34-kt, ≥50-kt, and ≥ 64-kt). A quantitative assessment of the WPFP is performed for the seven IP (0-12, 12-24, 24-36, 36-48, 48-72, 72-96, and 96-120 h).

The attribute diagram is used for direct evaluation of the WPFP forecasts. It has all benefits of a reliability diagram, but with the additional ability to assess forecast skill. Overall, the IP attribute diagram indicates that the WPFP over-forecasts for forecast probabilities less than 60%, nearly perfectly reliable for forecast probabilities between 61% and 80%, and under-forecasts for probabilities between 81% and 100%. For the most part the forecast system has positive skill.

Additional available methods for forecast verification are indirect measures of forecast skills, but they require the probabilities to be converted into binary (yes/no) forecasts with the use of a decision threshold. The decision thresholds, those that provide the greatest forecast skill, allow for evaluation of the performance of the forecasting system. Decision thresholds were selected using maximum values from the true skill statistic (TSS) and the Heidke skill score (HSS) methods. IP decision thresholds range from 1% to 37%. More statistically relevant decision thresholds are obtained from bootstrap resampling methods. These results suggest that decision thresholds should be determined from an average (or median) of the resampled data. Indirect evaluation of the NHC WPFP shows that it performs reasonably well. Both TSS and HSS IP statistics demonstrate significant forecast skill up to the 72h forecast interval.

WPFP null forecasts, i.e. less than 0.5% probability ("X"), can affect the value of the decision thresholds. Sensitivity of the threshold selection methods to null forecasts is tested by using a buffer zone to filter varying amounts of null forecasts. The sensitivity tests show that the TSS based thresholds are more sensitive to the inclusion of the null forecasts, while the HSS based thresholds are not. The use of a 400 km buffer zone is recommended for filtering null forecasts.

The WPFP forecasts are segregated into the following TC intensity categories to test for performance differences: intensifying, no-trend (no change) and weakening. The performance of these categories is different with statistical significance. However, the forecast system performs well for any stage of TC intensification, and the best for weakening TCs.

Monday, 27 June 2011
11:00 a.m.
Room AOSS 351