2C.6

A NEW PRODUCT FOR ESTIMATING THE PROBABILITY OF TROPICAL CYCLONE FORMATION

Mark DeMaria
NOAA/NESDIS, Fort Collins, Colorado

Charles W. Anderson
Department of Computer Science, Colorado State University, Fort Collins, CO

John A. Knaff and Bernadette H. Connell
Colorado State University/CIRA, Fort Collins, CO

1. INTRODUCTION

Beginning in 2003, the National Hurricane Center (NHC) extended their official tropical cyclone (TC) track and intensity forecasts from 3 to 5 days, based upon the need for increased lead times. With these longer forecasts, it becomes even more likely that a storm could form after the start and have an impact before the end of the forecast period. In addition, the prediction of TC formation contained within the NHC Tropical Weather Outlook product is a relatively subjective forecast. For these reasons, new objective methods for predicting TC formation are needed.

As the resolution of global forecast models continues to increase, they have begun to show some promise in predicting TC formation. However, global models sometimes over-predict TC genesis (Beven 1999).

Hennon and Hobgood (2003) presented a method to estimate the probability of Atlantic tropical cyclogenesis using a linear discriminant analysis technique, with input parameters describing the environment of tropical cloud clusters identified by global analyses and satellite imagery. A three-year sample (1998-2000) was included in their study. In this paper, the approach of Hennon and Hobgood is generalized by including east Pacific storms, a longer data sample, additional discriminators, and a quantitative use of satellite imagery to provide a purely objective estimate of TC formation probability.

2. DATA

The purpose of this study is to provide a new product for estimating the probability of TC formation in the NHC area of responsibility. The analysis domain extends from 140^{o}W to 10^{o}W and from the equator to 45^{o}N, which includes the formation locations of all Atlantic and eastern North Pacific TCs. This domain is divided into 26 by 9 sub-regions each covering 5^o by 5^o lat/lon areas. The product will be developed from input at 12 hr intervals from May 1st to December 15th. This time period includes ½ month before the start of the east Pacific hurricane season and ½ month after the end of the Atlantic and east Pacific hurricane seasons. The initial version of the product will estimate the formation probably within the next 24 hr after each analysis date and time.

The data used in the product development includes three time periods. Background climatological probability estimates are based upon the NHC best track from the Atlantic and east Pacific for the period 1949-2002. Formation is defined as the first point in the best track file, excluding extratropical and subtropical cases. The "unnamed" depressions since 1989 were included because the operational procedures for those systems have been fairly systematic since that time.

To determine storm environmental parameters, NCAR/NCEP reanalysis fields (2.5^o by 2.5^o resolution) for the period 1980-1999 and NCEP operational analyses for 2000-2002 were used. For additional environmental and tropical disturbance parameters, GOES-8 channel 3 (water vapor) imagery for the period 1995-2002 were employed. The full disk sector covers the entire analysis domain and is available every 3 hours. These images were re-mapped to a 16 km Mercator projection. The Levitus climatological (monthly) sea surface temperatures (SSTs) are also used, as described below.

3. PRODUCT DEVELOPMENT

To illustrate the difficulty of predicting TC formation, there are 229 days from May 1 to Dec 15, and the analysis domain includes 234 5^o by 5^o sub-regions. During 1949-2002 there was an average of 24 formations per year over this domain. Thus, in any given sub-region, there is less than a 1 in 2200 chance (0.044%) of a TC formation within the following 24 hours.

The formation probability development includes three steps. First, a 24 hr climatological formation probability is determined from the 1949-2002 best track data. Next, a set of "selection" rules are developed to eliminate cases where formation is extremely unlikely using climatological, synoptic and disturbance parameters. Finally, a linear discriminant analysis technique is used to estimate the formation probability, where the developmental sample includes all sub-region points for May 1-Dec. 15 from 1995-2002 for which satellite data and NCEP analyses are both available. Only those points retained after the application of the selection rules are included in the discriminant analysis.

**Corresponding author address*: Mark DeMaria, NOAA/NESDIS/CIRA, CSU, West Laporte Avenue, 80523, e-mail: Mark.DeMaria@noaa.gov

Table 1. Summary of selection rules used to identify sub-region areas where formation is unlikely

Parameter	Source	Elimination Criterion
Latitude	Domain definition	< 5^oN
max climo SST	Levitus SST	< 21 ^oC
%sub-region over land	land mask	100%
Pre-existing storm	NHC best track	All sub-regions with a TC
850-200 hPa vertical shear	NCEP analyses	> 65 kt
850 hPa circulation	NCEP analyses	< - 5 kt
Vertical instability	NCEP analyses	< -8^oC
%pixels < -40^oC	GOES imagery	0%
average WV brightness T	GOES imagery	> -23^oC

The 24 h climatological formation probability (CPROB) was determined by counting the number of formations in each sub-region for each month, and then dividing by the number of years (54) in the best-track sample. Due to the fairly small size of the sub-regions, the maximum CPROB value is only 1.9%, which occurs in July for the east Pacific sub-region centered at 12.5^oN, 102.5^oW. In the Atlantic basin, the maximum CPROB value of 0.65% occurs in August for the sub-region centered at 12.5^oN and 22.5^oW.

The next step is to eliminate as many sub-region areas as possible using the selection rules listed in Table 1. The first three rules are based upon the long-term climatology. The next rule eliminates all points that already have a storm in them. Rules 5-7 use parameters calculated from the 1980-2002 sample of NCEP analyses. The circulation is the average wind component tangent to the sub-region boundary. The vertical instability is determined by lifting a parcel from the surface to 200 hPa assuming the parcel θ_e is conserved, where the soundings are from the NCEP analyses averaged over each sub-region. The vertical instability is the pressure-weighted average of the difference between the parcel θ_e and the saturation θ_e of the environment. The parameters for the last two rules are determined from the three GOES images closest in time to each analysis date and time. The pixel percentage measures the amount of persistent deep convection in each sub-region. The average brightness temperature only includes pixels warmer than -40^oC and measures the mid-level moisture in the disturbance environment.

After the selection rules are applied, there were about 313,000 sub-region points left in the May 1st-Dec 15^{th} 1995-2002 12-h interval sample. Formation occurred at 427 of these points within the next 24 hours. This sample is used as input to a linear discriminant analysis routine, with the parameters listed in Table 1 as input, supplemented with CPROB.

Table 2. Normalized parameter weights from the discriminant analysis of tropical cyclone formation.

Parameter	E. Pacific value	Atlantic value
CPROB	1.52	1.75
circulation	1.19	1.35
%Pixels < -40^oC	1.04	0.84
distance to nearest storm	-0.81	-0.70
vertical shear	-0.25	-0.36

The discriminant analysis provides the weights for the linear combination of the input variables that best differentiates between the formation and non-formation cases. For the discriminant analysis, the Atlantic and east Pacific cases were treated separately. Before the discriminant analysis was applied, the input variables were scaled by subtracting the sample mean, and then dividing by standard deviation. With this scaling the linear discriminant weights can be compared for each input variable.

Table 2 shows the discriminant weights for each basin. In both basins, only CPROB, circulation, %pixel count, distance to nearest storm and vertical shear had a significant effect on the discrimination. The parameter weights for each basin are similar which suggests that this method may have application to other basins.

Once the discriminant weights are found, the value of the discriminant function is determined for each sub-region at each time and date. Then, the discriminant function values for the formation cases are sorted and divided into 10 equally sized groups. The non-formation cases are also put into these groups based upon the discriminant functions values that evenly divide the formation cases. After this partition, the formation probability for each group is calculated. For the group with the largest values of the discriminant function, the probability is 14.3% (6.6%) for the east Pacific (Atlantic), which is about a factor of 8 (10) greater than the largest probabilities based solely on climatology.

4. FUTURE PLANS

The data from the 2003 hurricane season will be used as an independent sample for evaluation of the formation probability product. Plans are underway to make the formation probability an operational NESDIS product that is updated 4 times per day.

REFERENCES

Beven, J.L., 1999: The boguscane – A serious problem with the NCEP medium range forecast model in the Tropics. Preprints, *23rd Conf. On Hurr. and Trop.l Meteor.*, Dallas, TX, Amer. Meteor. Soc., 845-848.

Hennon, C.C., and J.S. Hobgood, 2003: Forecasting tropical cyclogenesis over the Atlantic basin using large-scale data. *Mon. Wea. Rev*., **131**, 2927-2940.