Continued Development of Tropical Cyclone Wind Probability Products

John A. Knaff – Presenting
CIRA/Colorado State University
and
Mark DeMaria
NOAA/NESDIS

Presentation at
60th Interdepartmental Hurricane Conference
Mobile, AL
22 March 2006
Outline

• Training Activities
• Verification Activities
• Results/insights from examination of hurricane warning break points
Training Activities

- Mark DeMaria coordinated with Rick Knabb to provide feedback on a TPC/NHC training session.

- Several cases rerun for 2004 and 2005
  - For Pablo Santos, Miami WFO for an experimental algorithm that uses the probabilities.
  - Web page with examples and a product description [http://rammb.cira.colostate.edu/projects/tc_wind_prob](http://rammb.cira.colostate.edu/projects/tc_wind_prob)
Experimental Tropical Cyclone Wind Probabilities

2004 Florida Hurricane Landfall Case Studies
and 2005 Katrina and Rita Landfall Case Studies

Case Studies
Charley 2004
Frances 2004
Jeanne 2004
Ivan 2004
Katrina 2005
Rita 2005

Previous Page
Verification: Current Status

• Developed:
  – Input data handling (GRIB, ATCF, …)
  – Statistical Methods
    • Scalar measures of skill, accuracy, confidence
    • Conditional measures
  – Methods to assess deterministic forecasts

• Remaining
  – Treatment of the OFCL forecast & wind radii through 5 days.
  – Integrating the pieces.
  – How to Interpret the statistics and optimize use.
Statistical methods: Probability Bias

• Mean Forecast Probabilities \( (F_i) \) minus the Mean Observed Frequencies \( (E_i) \) = 1 or 0

\[
Bias = \frac{\sum_{k=1}^{N} F_k}{\sum_{k=1}^{N} E_k}
\]

Determines if the probabilities over/under forecast the outcome.
Statistical Methods: Brier Score

• Mean of square of the Forecast Probabilities \((F_i)\) minus the Observed Frequencies \((E_i)\) = 1 or 0

Measures the Mean Square Errors (Accuracy) associated with a probabilistic forecast

\[
BS = \frac{1}{N} \sum_{k=1}^{n} (F_k - E_k)^2
\]
Statistical Methods: Brier Skill Score

• A scalar skill score comparing a given Brier Score with the Brier Score of a reference forecasts (OFCL, CLIPER etc.).

Assess relative accuracy (skill) of a probabilistic forecast with respect to a reference forecast.

\[ BSS = 1 - \frac{BS}{BS_{\text{ref}}} \]
Statistical methods: Discrimination Distance

• Distance between the mean forecast probability ($F$) for all event ($E$) and all non-events ($E'$).

Measures the ability of a forecast scheme to discriminate events.

\[
d = \left| \mu_F|_E - \mu_F|_{E'} \right|
\]
Statistical Methods: Conditional Distributions

Observed frequency of events

Bins of Forecast Probabilities
Statistical Methods: Relative Operating Characteristics

A series of 2x2 contingency tables, which are conditional on a range of forecast probabilities are constructed.

For instance a warning would be issued if the probability exceeded 1,2,3,4,…100 %

<table>
<thead>
<tr>
<th>Observation</th>
<th>Warning (W)</th>
<th>No Warning (W’)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event (E)</td>
<td>h</td>
<td>m</td>
<td>E</td>
</tr>
<tr>
<td>Nonevent (E’)</td>
<td>f</td>
<td>c</td>
<td>E’</td>
</tr>
<tr>
<td>Total</td>
<td>w</td>
<td>w’</td>
<td>N</td>
</tr>
</tbody>
</table>

The results of the contingency tables can be quantified in terms of hit rate \((hr) = h/(h+m)\) and false-alarm rate \((far)=f/(f+c)\)

A plot of far vs. hr can be created and a skill score created from the area under the curve.
Statistical Methods: ROC diagram & Skill Score

**ROC Skill Score**

\[ SS_{ROC} = 2A - 1 \]

where \( A \) is the area under the curve

Mason and Graham (1999)
Verification Procedure: Test Dataset

• Dataset
  – 5-day cumulative 64-kt wind probabilities were generated for 342 coastal break points (195 official breakpoints + 147 additional points)
  – These were analyzed when warnings were issued for the 14 storms to the right
  – N=128250 points

<table>
<thead>
<tr>
<th>Storm Name</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alex</td>
<td>2004</td>
</tr>
<tr>
<td>Charley</td>
<td>2004</td>
</tr>
<tr>
<td>Frances</td>
<td>2004</td>
</tr>
<tr>
<td>Gaston</td>
<td>2004</td>
</tr>
<tr>
<td>Ivan</td>
<td>2004</td>
</tr>
<tr>
<td>Jeanne</td>
<td>2004</td>
</tr>
<tr>
<td>Arlene</td>
<td>2005</td>
</tr>
<tr>
<td>Cindy</td>
<td>2005</td>
</tr>
<tr>
<td>Dennis</td>
<td>2005</td>
</tr>
<tr>
<td>Emily</td>
<td>2005</td>
</tr>
<tr>
<td>Katrina</td>
<td>2005</td>
</tr>
<tr>
<td>Ophelia</td>
<td>2005</td>
</tr>
<tr>
<td>Rita</td>
<td>2005</td>
</tr>
<tr>
<td>Wilma</td>
<td>2005</td>
</tr>
</tbody>
</table>
Verification Procedure: Summary Skill Measures

\[ \text{Bias} = 0.893 \text{ (under forecasts)} \]
\[ BS = 0.0248 \]
\[ BS_{OFCL} = 0.0346 \]
\[ BS_{zero} = 0.0392 \]

\[ BSS_{OFCL} = 28.30\% \]
\[ BSS_{zero} = 36.75\% \]
Verification Procedure: Conditional Distribution of Break Point Probabilities

Slight under forecast of probabilities for this dataset – due to Wilma
Verification Procedure: Summary

• The probabilities are skillful
  – Brier Skill Score 28% more accurate than the OFCL deterministic forecast
    • Note 50% of the OFCL forecasts verified
  – ROC Skill Score 88%
  – The discrimination distance $d=26\%$ is large
  – Probabilities slightly under forecast and are well calibrated for this limited dataset
Can the wind speed probabilities be used to decrease the area warned or increase lead time?
Probability Model for NHC Hurricane Warnings

NHC storm total hurricane warning lengths 1963-2005

NHC storm average hurricane warning lead times 1963-2005

Since 2000 warning areas have decreased and lead times increased.

60th IHC  22 March 2006    -    Mobile, AL
Discrimination Distance

\[ \mu_{F|E} = 27.8\%, \sigma_{F|E} = 29.2\% \]

\[ \mu_{F|E'} = 1.4\%, \sigma_{F|E'} = 3.4\% \]

\[ d = 26.4\% \]
Why are there so many low probabilities at the break points?

\[ \mu_{F|E} = 27.8\%, \sigma_{F|E} = 29.2\% \]
Distribution of Probabilities at the Ending Break Points

\[ \mu_{F_{\text{end}}|E} = 8.8 \]

\[ \sigma_{F_{\text{end}}|E} = 14.9 \]
Example: Hurricane Rita

Warnings are brought down too slowly in this case.
Summary of Warning Break Points

• Probabilities are useful in the watch/warning process.
  – Objectively assign of the warnings at fixed lead time?
  – Average at warnings = 28%
  – Average at end points of the Warnings = 9%

• It appears that warnings can be dropped sooner, thus decreasing the area warned area.
Future Plans

• Seasonal Verification Code
• See what we can learn from the verification.
• Report to this audience

Questions?