Objective Lightning Probability Forecast Tool
Phase II

Winnie Lambert
AMU/ENSCO, Inc.

Outline

- Task Overview
- Phase I Summary
- Phase II Goal
- Modifications and Effects
- Phase II Tool
  - Predictors
  - Performance
  - Automation
- Summary and Future Work
Overview

• 45 WS provides lightning probability for the day/week
  – Daily Weather Briefing at 7:00 am local time
  – Used for general daily Range operations planning
• Subjective analysis of model and observational data
• AMU-developed Objective Lightning Forecast Tool
  – Provide probability of lightning occurrence May–September
  – Accessed through GUI
• 45 WS requested an update to the tool:
  – Modify certain predictors and possibly improve performance
  – Create automated tool

Phase I Summary

• Pre-Phase I: Neumann-Pfeffer Thunderstorm Index (NPTI)
  – Developed over 30 years ago, tuned to KSC/CCAFS area
  – Official objective lightning forecasting tool
• NPTI performance worse than 1-day persistence
• Forecasters requested new lightning forecast tool
• New tool showed
  – 48% improvement over NPTI; 31-53% over persistence
  – Good reliability, accuracy measures, and skill scores
  – Ability to distinguish between lightning/non-lightning days
• Transitioned to operations before 2005 lightning season
Phase I Summary

- 5 equations output probability of CG occurrence
  - One equation for each month
  - Logistic regression: \[ y = \frac{e^{(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}} \]
- Each equation had 5-6 predictors
  - Common to all 5 equations: Daily climatology, flow regime, 1-day persistence
  - Common to 4 equations (Jun – Sep): Mean RH in 800–600 mb layer
- Created GUI to interface with complex equations

Phase II Goal

- 45 WS Request
  - Try new formulations of certain predictors
  - Develop an automated tool
- Predictor Modifications
  - Increased POR from 15 to 17 warm seasons (1989 – 2005)
  - New valid area for CG occurrence
  - Used new smoothing function for daily climatology
  - Changed calculation of flow regime
  - Determined optimal RH layer
- Automated Tool
  - Developed in MDDS by P. Wahner of CSR
  - GUI format similar to previous Excel tool
**Data Sources**

- **Cloud-to-Ground Lightning Surveillance System (CGLSS)**
  - Ground truth
  - Climatology

- **CCAFS (XMR) 1000 UTC sounding**
  - Data used for 7:00 am briefing
  - 11 parameters (e.g. LI, KI, etc.)
  - Flow regime in Phase II

- **Florida 1200 UTC soundings**
  - Flow regimes
  - Low-level wind dir at MIA – TBW – JAX

**Modifications**

**Valid Area**

- Previous valid area defined by rectangle surrounding all 5 n mi warning circles
- 45 WS planning forecast is for KSC/CCAFS circles
- Modification: only use CG strikes within KSC/CCAFS 5 n mi circles
Modifications
Valid Area

- Area reduced
- Spatial climatology shows steep gradient in CG occurrence
- Will change #strikes
- Might change #lightning days
- Recalculate probabilities:
  - Daily climatology
  - Flow regime climatology
  - Persistence

Modifications
Daily Climatology

- # of CG days for each date divided by # years (green curve)
- Smoothing technique: Center-weight Gaussian
  - Phase I (blue curve): ±7 days, scale = 3 days
  - Phase II (red curve): ±14 days, scale = 7 days
- Values decreased ~10% after change in valid area
Modifications
Flow Regime Probability

• Method from FSU study: identified six distinct flow regimes
• Flow regime determined by:
  – Average wind direction in 1000–700 mb
  – 1200 UTC MIA – TBW – JAX
• Lightning frequencies calculated for each flow regime in each month
• Modification
  – ISSUE: no flow regime 42% of days
  – SOLUTION: Used XMR 1000 UTC sounding as discriminator

SW-2 Regime

Modifications
Flow Regime Probability

• Using XMR sounding:
  – Reduced days in 'Other' and 'Missing' by over 70%
  – Increased number of days in SW-2, SE-1, NE and NW
• New values are ~10% lower than Phase I

<table>
<thead>
<tr>
<th>Flow Regimes</th>
<th># of Days Before</th>
<th># of Days After</th>
<th>Lightning Prob (%) Before</th>
<th>Lightning Prob (%) After</th>
</tr>
</thead>
<tbody>
<tr>
<td>SW-1</td>
<td>301</td>
<td>301</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>SW-2</td>
<td>256</td>
<td>606</td>
<td>72</td>
<td>57</td>
</tr>
<tr>
<td>SE-1</td>
<td>318</td>
<td>438</td>
<td>51</td>
<td>32</td>
</tr>
<tr>
<td>SE-2</td>
<td>248</td>
<td>248</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>NW</td>
<td>100</td>
<td>307</td>
<td>43</td>
<td>32</td>
</tr>
<tr>
<td>NE</td>
<td>114</td>
<td>317</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>Other</td>
<td>1077</td>
<td>326</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>Missing</td>
<td>187</td>
<td>58</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Modifications
Optimal Mid-Level RH Layer

- Mean 800 – 600 mb RH used as a predictor in NPTI
- Modification: Find mean RH layer most correlated with lightning occurrence
- Iterative technique
  - Bottom: 950 mb; Top: 450 mb
  - Calculate correlation of each layer to lightning occurrence
- Optimal layer: 825 – 525 mb

Phase II Equations

- Development data: 14 yrs
- Verification data: 3 yrs
- 14 candidate predictors
- 5 logistic regression equations
- Chose predictors that made > 0.5% reduction in variance

<table>
<thead>
<tr>
<th>Predictors for Each Month in Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>May</strong></td>
</tr>
<tr>
<td>K-index</td>
</tr>
<tr>
<td>Flow Regime</td>
</tr>
<tr>
<td>Vertical Totals</td>
</tr>
<tr>
<td>Daily Climatology Persistence</td>
</tr>
</tbody>
</table>
Performance

- Four tests using 3-yr verification set
- Brier Skill Score
  - Phase II improved skill over other methods
  - Overall 8% improvement over Phase 1, 56% over NPTI
- Reliability Diagram
  - Black line: perfect reliability
  - Phase I and II have "under-forecast" bias
    - Phase I: -5.9%
    - Phase II: -0.4%

% Improvement over Forecast Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence</td>
<td>28</td>
<td>41</td>
<td>37</td>
<td>47</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Daily Climo</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly Climo</td>
<td>29</td>
<td>27</td>
<td>30</td>
<td>25</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Flow Regime</td>
<td>16</td>
<td>11</td>
<td>18</td>
<td>18</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Phase-1 Eqns</td>
<td>0.2</td>
<td>5</td>
<td>19</td>
<td>-0.8</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

Reliability Diagram for the New Equations
May-September 1999-2005

Performance

- Lightning/non-lightning day distributions
  - Phase I and II distinguish non-lightning days well
  - Phase II better at distinguishing lightning days

Contingency Table Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>P-2 (0.47)</th>
<th>P-1 (0.35)</th>
</tr>
</thead>
<tbody>
<tr>
<td>POD</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>FAR</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>HR</td>
<td>0.74</td>
<td>0.71</td>
</tr>
<tr>
<td>CSI</td>
<td>0.52</td>
<td>0.49</td>
</tr>
<tr>
<td>HSS</td>
<td>0.47</td>
<td>0.40</td>
</tr>
<tr>
<td>KSS</td>
<td>0.47</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Contingency table statistics
- Yes/No cutoff 0.47 for Phase II, 0.35 for Phase I
- Both Phases better than persistence
- Phase II scores show best accuracy and skill
Automation

- Equations available through MIDDS GUI
  - Developed by Paul Wahner of CSR
  - Accesses date and parameters from the 1000 UTC XMR sounding.
  - Forecasters choose ‘Yes’/‘No’ for persistence and a flow regime for the day.
- The MIDDS GUI similar to the Phase I GUI

Summary and Future Work

- Phase II equations performed better than Phase I
- Transitioned for 2007 lightning season
- Phase III:
  - Extend forecast to include October
  - Create equations based on daily climatology instead of month

AMU Website: http://science.ksc.nasa.gov/amu