HIWPP Test Program
February 9, 2016

Bonny Strong
Tasks

1. Ensemble Statistical Post-Processing
2. Visualization & Extraction via NEIS
3. Verification
4. Real-time IT Operations
Statistical Post-Processing

- Initial objective: combine a few high-resolution deterministic weather models to obtain a more skillful forecast
- At each grid point, produces model weighting that is inversely proportional to the MAE at that point in 30 days of training data
  - 2m temp
  - 10m winds
  - Z500
  - 6-hr precip
- Standard deviation of the weighted forecast error in the training period at each grid point (T2m, U/V10m, Z500).
- For precipitation, the probability of exceeding 1mm, 5mm and 10mm of precipitation.
Statistical Post-Processing

Additional research areas

• Evaluation of field alignment method, which allows quantification and removal of displacement and amplitude errors

• Extension of post-processing code to ensembles of 20 or members, permitting evaluation of GEFS and FIM ensemble research

• Evaluating the feasibility and benefit of methods developed to improve precipitation forecasts at a regional scale in the Sandy Supplemental Blender project, applied at a global scale within HIWPP
Verification - objectives

**Metrics available for evaluation of HIWPP hydrostatic models**
- Retrospectives
- Real-time

**Unified platform to view both:**
- EMC’s VSDB verification images
- GSD interactive web-based images

**Enhancements including some of:**
- Advanced precipitation verification
- Ensemble verification
- Multi-parameter score-cards
HIWPP Verification Available

- Running on dedicated hardware
- Publicly accessible

VSDB FILE-BASED VERIFICATION

00Z cycles - http://hiwpp.noaa.gov/verify/HiWPP_realtime_00/
12Z cycles - http://hiwpp.noaa.gov/verify/HiWPP_realtime_12/

These sites display verification statistics for HIWPP Global NWP model forecasts. Statistics are currently computed for GFS and FIM on the G2 grid (2.5x2.5 degree) over five regions, Global, Northern Hemisphere, Southern Hemisphere, Tropical, and Pacific North America. The 30-year NCEP/NCAR reanalysis is used for computing pattern and anomaly correlations.

INTERACTIVE UPPER-AIR VERIFICATION

Time Series - http://hiwpp.noaa.gov/verify/upper-air/time-series/Welcome.cgi
Vertical Profile - http://hiwpp.noaa.gov/verify/upper-air/profile/Welcome.cgi

These sites display verification statistics for HIWPP Global NWP model forecasts. Statistics are currently computed for GFS and FIM over several regions. Data are stored in 10 mb. increments in the vertical and displayed in 50 mb. increments. Time series and vertical profiles of the statistical comparisons can be viewed.

For more detailed information on the Global models in use, please see the model descriptions page.
Verification: EMC and GSD

HIWPP Global Model Experimental Forecast Performance Statistics

- The site displays verification statistics for EMC and GSD models.
- The regions referred to are: G2S, 60N-90N, 60S-90S, N60-60N, S0-60S, 320E-180E, 20N-70N, 60N-90N, 60S-90S.
- Pattern correlations for all NWP models and the NCEP/NCAR reanalysis are shown.
- Please see this presentation for more information.
Verification – MATS web interface

- Installing GSD’s interactive web application uses Java Applets:
  - Have been labeled a major security risk
  - No longer supported by Chrome and some other applications

MATS
Model Assessment Tool Suite
- Uses HTML5 in place of Java Applets
Verification – adding SYNOP observations
Verification - Ensembles

- Using VSDB package provided by Yuejian Zhu.
- To view output, we have a HIWPP installation of DTC’s METViewer software.
HIWPP has stimulated significant interactions and sharing of code between:

- NCEP/EMC
- DTC
- ESRL/GSD

Each has significant strengths, but also areas that need important development.

For model development and transition into operations, would be very helpful to share a common set of metrics.

Led to a vision of Unified Verification system, which has been proposed through NGGPS V&V team.
Real Time IT Operations

• Purchase of hardware
  • Including ~170 TB storage
• Managing transfer of data to collect model output in one place
• Open Data Initiative
Outcomes

• Improved transfer performance and resilience between major HPC locations
  • Installed and tuned GridFTP from Globus

• Large storage pool available for model data
• Distribution mechanism via THREDDS for sharing model data to public or collaborators
• Upgrades and improvements to serving data through web applications
## Open Data Initiative Users

### Total Registered Users

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>56</td>
</tr>
<tr>
<td>Commercial users</td>
<td>18</td>
</tr>
<tr>
<td>NOAA users</td>
<td>13</td>
</tr>
<tr>
<td>Academic users</td>
<td>12</td>
</tr>
<tr>
<td>Other gov’t agencies, CIs</td>
<td>13</td>
</tr>
</tbody>
</table>

### Usage 9/10/15 - 10/10/15

<table>
<thead>
<tr>
<th>Count</th>
<th>gbytes</th>
<th>Org name</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,644</td>
<td>1,119</td>
<td>Minnesota education cooperative</td>
</tr>
<tr>
<td>5,528</td>
<td>804</td>
<td>Meteociel</td>
</tr>
<tr>
<td>10,615</td>
<td>577</td>
<td>Weatherbell</td>
</tr>
<tr>
<td>30,295</td>
<td>74</td>
<td>Weather Company</td>
</tr>
<tr>
<td>52,082</td>
<td>2,577</td>
<td>Total</td>
</tr>
</tbody>
</table>
What Did We Learn?

Users:
• Early access to research model data was very highly welcomed by private enterprise
• Reliability was not a concern
• Most active users were companies with a web-based product that included weather information

Engagement with modelers:
• Users did not engage with modelers or provide any feedback about the models through the forum
• Some engagement did occur directly with some researchers

Support required:
• User support was not a significant issue for real-time data
• High-resolution visualization was difficult to support for bandwidth available to many users
Unexpected outcomes

Benefit for research collaboration:
• Infrastructure for public access was very useful for sharing dycore test output for researchers

Issues with volume of data:
• Issues with quantity of data produced by high resolution non-hydrostatic models had not been fully anticipated
• This is likely to be a significant issue in many aspects of model research in the near future
What Next?

• HIWPP project is in its final year and Open Data Initiative will conclude on Feb 8, 2016

• **But** – access to real-time research model data will carry on under NOAA’s Big Data Partnership

• Infrastructure will continue to be used to support NOAA research collaboration

• FIM and NAVGEM data are expected to remain available as long as:
  1) the model runs in real-time at NOAA/ESRL/GSD, and
  2) resources are available to support it