Advancing Limited Area Model Verification and Diagnostics Capability through the Enhanced Model Evaluation Tools (METplus)

Tara Jensen, NCAR, Boulder, Colorado USA;
with M. Mittermaier, the Met Office, Exeter, UK
and the rest of the METplus Team

the 42nd EWGLAM and 27th SRNWP Meeting
Remote Meeting
28 September – 2 October 2020
Why MET and then METplus?

Universities and National Laboratories

Operational Centers

Forecasters

Comprehensive and unified verification tool - Make R2O more efficient - Provide a consistent set of metrics

Allows researchers and operational scientists to speak a “common verification” language

User support of unified package provides greater opportunity to train all on verification best practices

METplus

Verification Packages
AF verification, EMC point, EMC grid, EMC precip, EMC ensemble, EMC global, ESRL MATS, WPC FVS, SPC Verification, DTC MET, University individual researcher packages, etc…
Object Based and Spatial Methods

- Selected to be the major component of the verification capability for NOAA, Navy, Air Force, and NCAR in the US and the Met Office in the UK
- ~100 traditional statistics and diagnostics using both point and gridded datasets
- Applied to many spatial and temporal scales
- Applied to many applications

- Originally developed to replicate and extend existing verification systems
- Platform independent and extensible
- Outstanding helpdesk support
- Online and in-person tutorial
- Container Support
- 3500+ users; Int’l and US
METplus

Python Wrappers for Verification Components

- **Simplify set-up** and running of MET
- **Open up MET’s interface** to work with Python algorithms
- **Python-based aggregation** and analysis
- **Python-based plotting** and diagnostics
- **Optionally** load data into METdb database and **display plots through METviewer or METexpress or generate plots of scorecards** through the METviewer batch engine

www.github.com/DTCenter/METplus
**METplus** Operational Categorical Statistics Use Case

- Analysis
- Forecast

Grid-Stat → Statistics → MET\text{datadb}

- \text{METviewer}
- Stat Plots
- Scorecard Plots
- **Python Script Embedding - Script**
  - User writes a script to read or pre-process data
  - The script should define a dictionary named `attr` which defines:
    - array must be named `met_data`.
    - `valid` and initialization (`init`) times as strings in `YYYYMMDD[._HH[MMSS]]` format.
    - `lead` and accumulation (`accum`) times as strings in `[HH[MMSS]]` format.
    - `name`, `long_name`, `level`, and `units` as strings.
    - `grid dictionary defining the projection` and grid information in the same way as the gridded NetCDF files produced by MET.
  - Support for:
    - Numpy, Xarray, Pandas & Python 3.6.3/3.7

Before Python Embedding

MET calls Python which

Reads file

After Python Embedding

MET calls Python which

Analyzes file

Gridded NetCDF

Gridded Forecast Analysis Obs

Grid Stat

STAT ASCII NetCDF

Python Script

Gridded NetCDF

Obs File

Stat-Analysis

Python Script

Stat-Analysis

STAT ASCII NetCDF

Obs File
Philosophy: One Tool-Many Applications, Example: GridDiag

- Inventories the data
- Develops the PDF (one or two-dimensional)
- User configurable bins for PDF and percentiles
- Writes out – or holds in memory – bins or percentiles for use by other tools (GridStat, Point-Stat, MODE, MTD)

**S2S**
Correlate distributions of two fields

**CAM**
Normalize fields based on their sample climatology

**Space Weather**
Adapt measures to regional max values

**Updraft Helicity**

![Graphs and diagrams illustrating the applications of GridDiag and S2S](image)
Distance Maps

- Apply threshold of 10-meter VGRD $\geq 5$ m/s.
- For each grid point, compute minimum distance to nearest event.
- DMAP statistics are computed by comparing fcst and obs distance maps.

**Table 8.7: Format information for DMAP (Distance Map) output line type.**

<table>
<thead>
<tr>
<th>Column Number</th>
<th>DMAP Column Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>DMAP</td>
<td>Distance Map line type</td>
</tr>
<tr>
<td>25</td>
<td>TOTAL</td>
<td>Total number of matched pairs</td>
</tr>
<tr>
<td>26</td>
<td>FY</td>
<td>Number of forecast events</td>
</tr>
<tr>
<td>27</td>
<td>OY</td>
<td>Number of observation events</td>
</tr>
<tr>
<td>28</td>
<td>FRIAS</td>
<td>Frequency Bias</td>
</tr>
<tr>
<td>29</td>
<td>BADDELEY</td>
<td>Baddeley's $\Delta$ Metric</td>
</tr>
<tr>
<td>30</td>
<td>HAUSDORF</td>
<td>Hausdorff Distance</td>
</tr>
<tr>
<td>31</td>
<td>MED FO</td>
<td>Mean-error Distance from forecast to observation</td>
</tr>
<tr>
<td>32</td>
<td>MED OF</td>
<td>Mean-error Distance from observation to forecast</td>
</tr>
<tr>
<td>33</td>
<td>MED_MIN</td>
<td>Minimum of MED_FO and MED_OF</td>
</tr>
<tr>
<td>34</td>
<td>MED_MAX</td>
<td>Maximum of MED_FO and MED_OF</td>
</tr>
<tr>
<td>35</td>
<td>MED_MEAN</td>
<td>Mean of MED_FO and MED_OF</td>
</tr>
<tr>
<td>36</td>
<td>FOM FO</td>
<td>Pratt's Figure of Merit from forecast to observation</td>
</tr>
<tr>
<td>37</td>
<td>FOM OFF</td>
<td>Pratt's Figure of Merit from observation to forecast</td>
</tr>
<tr>
<td>38</td>
<td>FOM_MIN</td>
<td>Minimum of FOM_FO and FOM_OF</td>
</tr>
<tr>
<td>39</td>
<td>FOM_MAX</td>
<td>Maximum of FOM_FO and FOM_OF</td>
</tr>
<tr>
<td>40</td>
<td>FOM_MEAN</td>
<td>Mean of FOM_FO and FOM_OF</td>
</tr>
<tr>
<td>41</td>
<td>ZHU FO</td>
<td>Zhu's Measure from forecast to observation</td>
</tr>
<tr>
<td>42</td>
<td>ZHU OF</td>
<td>Zhu's Measure from observation to forecast</td>
</tr>
<tr>
<td>43</td>
<td>ZHU_MIN</td>
<td>Minimum of ZHU_FO and ZHU_OF</td>
</tr>
<tr>
<td>44</td>
<td>ZHU_MAX</td>
<td>Maximum of ZHU_FO and ZHU_OF</td>
</tr>
<tr>
<td>45</td>
<td>ZHU_MEAN</td>
<td>Mean of ZHU_FO and ZHU_OF</td>
</tr>
</tbody>
</table>

Collaboration with Eric Gilleland, NCAR/RAL.
Point-Stat HiRA

- **Collaboration with Marion Mittermaier and USAF**
- Enhance **Point-Stat** to apply the High Resolution Assessment (HiRA) verification logic to deterministic forecasts matched to point observations (Mittermaier, 2014)
- Process neighborhood values as an ensemble forecast (ECNT line type).
- Threshold, compute fractional coverage, and verify as a probability forecast (**PCT**, **PSTD**, **PRC**, and **PJC** line types).
- Allows for some spatial / temporal uncertainty by giving credit for being ‘close’.
- Allows for comparison of deterministic and ensemble forecasts via the same set of probabilistic statistics.
- Also allows for comparison of models with different grid resolutions via adjustment of neighborhood size.

**HiRA Ensemble:**
- Write ECNT using neighborhood ensemble.

**HiRA Probabilities:**
- 1x1 NBRHD = 1/1
- 3x3 NBRHD = 1/9
- 5x5 NBRHD = 4/25

**Model Forecast**
- White boxes = 0
- Colored boxes > 0

**Threshold Forecast**
- Blue boxes = event
cat_thresh = [ >0 ];

**HiRA Ensemble:**
- Write ECNT using neighborhood ensemble.

```
output_flag = {
  // Enable or disable
  flag       = TRUE;
  // Neighborhood sizes
  width      = [ 2, 3, 4, 5 ];
  // Probability thresholds
  cov_thresh = [ ==0.25 ];
  vld_thresh = 1.0;
  // Neighborhood shape
  shape      = SQUARE;
};

Model Forecast
White boxes = 0
Colored boxes > 0

Threshold Forecast
Blue boxes = event
cat_thresh = [ >0 ];
```

---

**HiRA Probabilities:**
- 1x1 NBRHD = 1/1
- 3x3 NBRHD = 1/9
- 5x5 NBRHD = 4/25
Neighborhood Probabilities

- **Collaboration with NOAA MMM**

- **Schwartz and Sobash (2017).**

- **Ensemble-Stat** pre-processes to compute the frequency of an event across all ensemble members.

- **Ensemble-Stat** pre-processes to compute simple neighborhood probabilities (NEP) and neighborhood maximum ensemble probabilities (NMEP).

```
nbrhd_prob = { ... }
nmep_smooth = { ... }
ensemble_flag = {
  ... 
  nep = TRUE;
  nmeap = TRUE;
  ...
}
```

![Maps showing Ensemble Frequency, NEP, and NMEP](images)
Surrogate Severe

- **Collaboration with NOAA/NSSL and NOAA/WPC**
- Enhanced regridding options to support new **MAXGAUSS** method:
  - Compute maximum value over defined neighborhood
  - Apply configurable Gaussian smoother to the result
- Also includes smoothing options to support new **GAUSSIAN** method

```
regrid_data_plane_HRRR_20190518_013_036_24h.nc G211 \
regrid_data_plane_HRRR_MXUPHL_24_MAXGAUSS.nc \
-field 'name="MXUPHL_24"; level="A24";' \
-method MAXGAUSS -width 27 \
-gaussian_radius 120 -gaussian_dx 81.271
```
Practically Perfect Prog and Data Thinning

- **Collaboration with NOAA/NSSL, NOAA/WPC, and NOAA/EMC**
- **Gridding Local Storm Reports reports (Practically Perfect Prog)**
- **Data thinning of satellite data like NOAA/GOES-16/17**
- **Point to grid (Point2Grid)** tool reads NetCDF output of point pre-processing tools and writes a gridded NetCDF output file.
- Gridded NetCDF output can be read by other MET tools.
## Use of Scorecards for Short-Range Forecasting

<table>
<thead>
<tr>
<th>Fraction Skill Score</th>
<th>Daily Domain</th>
<th>CONUS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12 hr</td>
<td>18 hr</td>
</tr>
<tr>
<td>Composite Reflectivity</td>
<td>&gt;=20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=40</td>
<td></td>
</tr>
<tr>
<td>CSI</td>
<td>&gt;=20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=40</td>
<td></td>
</tr>
<tr>
<td>3 hr Accumulated Precipitation</td>
<td>&gt;=0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=1.0</td>
<td></td>
</tr>
<tr>
<td>3 hr Accumulated Precipitation</td>
<td>&gt;=0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=1.0</td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>&gt;=20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=1.0</td>
<td></td>
</tr>
<tr>
<td>3 hr Accumulated Precipitation</td>
<td>&gt;=0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;=1.0</td>
<td></td>
</tr>
</tbody>
</table>

### RMSE

<table>
<thead>
<tr>
<th>ME</th>
<th>Daily Domain</th>
<th>CONUS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12 hr</td>
<td>18 hr</td>
</tr>
<tr>
<td>Temperature</td>
<td>sfc</td>
<td></td>
</tr>
<tr>
<td>Dew Point</td>
<td>sfc</td>
<td></td>
</tr>
<tr>
<td>Wind Speed</td>
<td>sfc</td>
<td></td>
</tr>
</tbody>
</table>

### ME

<table>
<thead>
<tr>
<th>ME</th>
<th>Daily Domain</th>
<th>CONUS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12 hr</td>
<td>18 hr</td>
</tr>
<tr>
<td>Temperature</td>
<td>sfc</td>
<td></td>
</tr>
<tr>
<td>Dew Point</td>
<td>sfc</td>
<td></td>
</tr>
<tr>
<td>Wind Speed</td>
<td>sfc</td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- ▲ NCEP is better than HRRR3FullYear at the 99% significance level
- ▲ NCEP is better than HRRR3FullYear at the 95% significance level
- ▲ No statistically significant difference between NCEP and HRRR3FullYear
- ▲ NCEP is worse than HRRR3FullYear at the 95% significance level
- ▲ Not statistically relevant
TC Metrics

- **Track Error**: great-circle distance between the forecast location and the actual location of the storm center (nmi)
- **Along-track Error**: indicator of whether a forecasting system is moving a storm too slowly/quickly
- **Cross-track Error**: indicates displacement to the right/left of the observed track
- **Intensity Error**: Difference between forecast and actual intensity (kts)
  - Raw intensity errors (bias) vs. absolute intensity errors (magnitude of error)

Graphics courtesy of NCAR TCMT
TC-Genesis

- Collaboration with Dan Halperin, Embry-Riddle Aeronautical University
- Compare forecast of TC-Genesis to actual BEST track and CARQ genesis events
- Configurable options to control genesis definition and matching, both spatially and in time
- Writes contingency table counts and statistics
- Aggregate results using STAT-Analysis tool

Figure: Tropical cyclogenesis verification for the NH for 2016. Symbols represent the Best Track (black), hits (green), late Genesis (blue) and false alarms (red).

```
tc_gen -config TCGenConfig_dev
-genesis suite[1-4]/2016*/genesis.fort.66*
-track nhc_atcf/*/2016.dat
-v 3 -log run_tc_gen.log
```
Changing The Perspective: TC-RMW

- Radius of Maximum Wind tool (spherical coordinates)
  - Implement DIA-Post tool from Hurricane Research Division (HRD) to convert model output to storm-relative coordinates.
  - Centered on the lat/lon of interest and computes an azimuthal average over multiple heights and radii.
  - Optionally, normalize by the radius of maximum winds.
  - Tangential wind and radial wind are derived.
  - Write the output to a NetCDF file.
Typically, a given mission will have 3–4 passes through the center of the storm. Due to the X-band nature of the **Tail Doppler Radar (TDR)**, reliable observations only extend outward ~50 km from the aircraft, limiting the azimuthal coverage of observations. An example of the coverage the TDR provides for a single pass is shown.

**METplus** will read these in via Python Embedding and give model developers the ability to use field project datasets for evaluations and process-oriented studies.

Image courtesy of Michael S. Fischer, Robert F. Rogers, Paul D. Reasor at NOAA/AOML/HRD
Growing METplus Community

**Research Institutions**
- NCAR (RAL, MMM, CGD)
- Naval Research Lab
- SBU, U of Illinois Urbana-Champaign, U of Wisc Madison, OU, UND, NC State, Purdue, Albany, etc…
- Labs NOAA Research (GSL, NSSL, PSL, ARL, GFDL)
- NASA Coordinated Community Modeling Center (for Space Wx)
- Army Research Lab

**Operational Centers adopted portions or all of METplus**
- NOAA EMC, SPC, WPC, OPC, SWPC, NHC, CPC, MDL
- Air Force Operational Center
- The Met Office (agreement executed)
- Other UM partners (i.e. Australia BoM, S. African WS) also considering contributing
- Fleet Numerical is considering once Naval Research Lab has transitioned
- Central Weather Bureau (Taiwan) is considering adoption
Repositories, Online Support, Tutorials

Developing YouTube-style Training Videos
User Support

- Helpdesk tickets at met_help@ucar.edu
- up by 10% with “back-and-forths” up even more
- Preparing for transition to Forums
Next 2-5 Years of METplus Development

**Coupled System Metrics:**
- Metrics Workshop
- Sea-ice output evaluation
- PBL diagnostics
- MJO and teleconnection diagnostics
- Hurricane track diagnostics
- Ensemble evaluation
- DA
- Space

**Process-oriented diagnostic:**
- Cold surface temperature, SST, and PBL biases
- Gravity wave drag and stratospheric coupling
- Sources of cloud cover and precipitation biases
- ENSO and teleconnections

**EMC priorities for DTC development:**
- Sea Ice
- PBL and Aerosols/Air Quality
- Ensemble
- Non-Severe LAM
- S2S and coupled model evaluation (LSM, Marine and Cryo, Hydro, Stratosphere, Upper Atmos/Geo-space)

**Other High Priority Projects:**
- US Climate Prediction Center R2O for S2S
- Marine/Cryosphere coupled metrics for use by EMC and OPC
- Stratospheric Coupling, General Circulation Index diagnostics
- Space Weather R2O
- Cloud Verification / DA diagnostics / Ensemble diagnostics
- Generalization of File Format Support including verification on the native model domain
Thank You for Your Attention

- Tara Jensen, METplus PM, jensen@ucar.edu
- Marion Mittermaier, EGWLAM contact, marion.mittermaier@metoffice.gov.uk
- User’s Page: https://dtcenter.org/community-code/metplus
- Find us on GitHub: https://github.com/DTCenter