NCAR

METplus

DTC Developmental Testbed Center

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

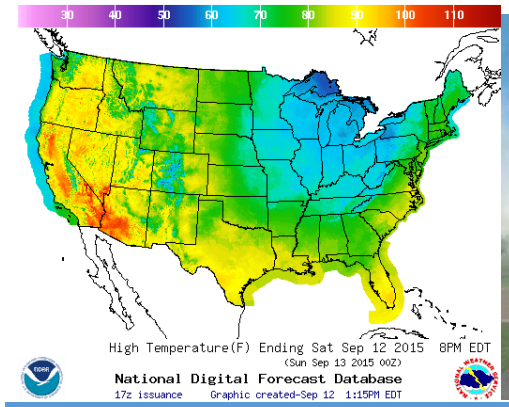
28 September – 2 October 2020

National Center for Atmospheric Research

Why MET and then METplus?

NCAR

Forecasters



Operational Centers

Universities and National Laboratories

Comprehensive and unified verification tool - Make R20 more efficient - Provide a consistent set of metrics

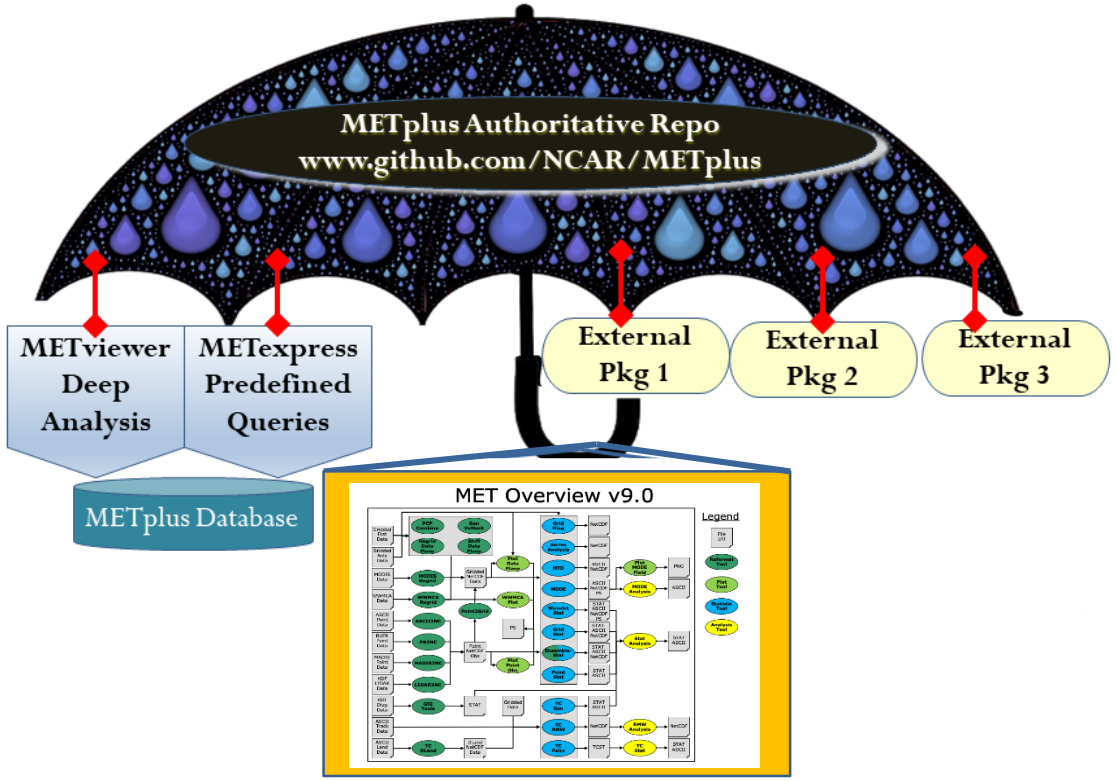
Allows researchers and operational scientists to speak a "common verification" language

METplus

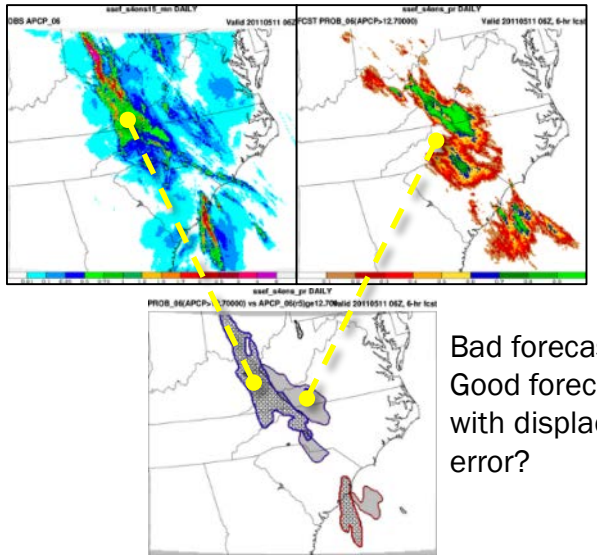
User support of unified package provides greater opportunity to train all on verification best practices

METplus

- 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

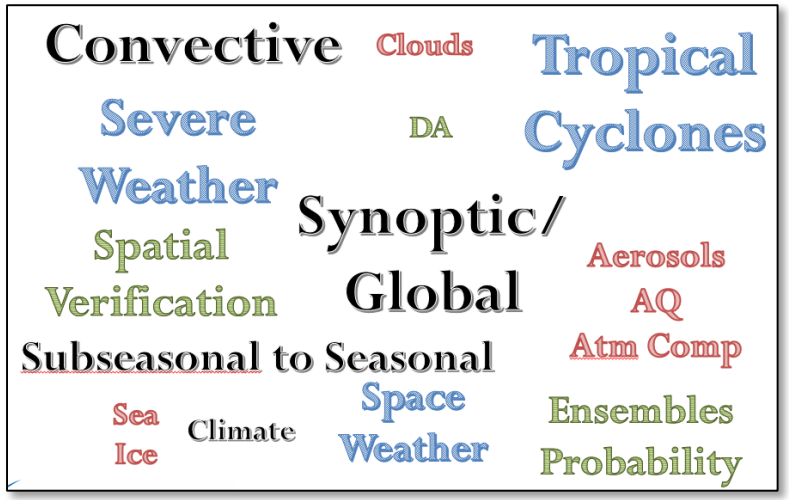


Object Based and Spatial Methods



Bad forecast or Good forecast with displacement error?

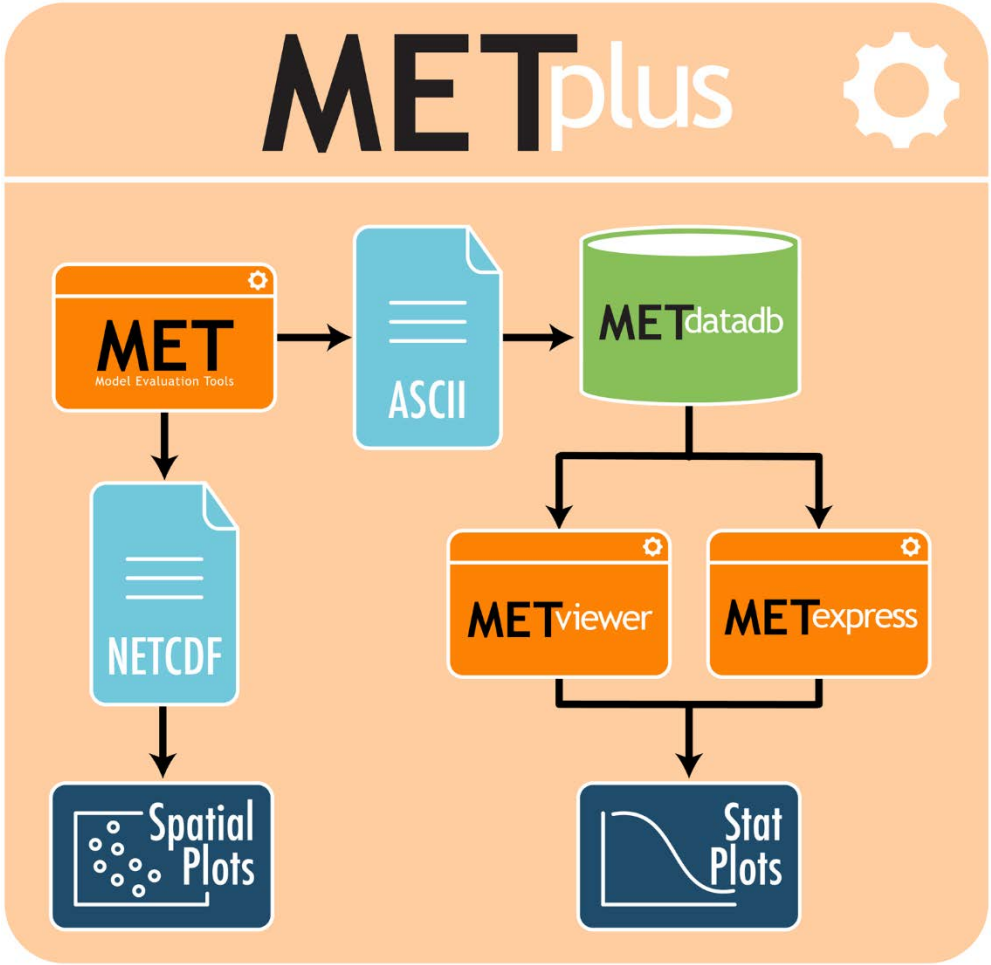
- Originally developed to replicate and extend existing verification systems
- Platform independent and extensible
- Outstanding helpdesk support
- Online and in-person tutorial
- Container Support
- 3500+ users; Int'l and US



METplus

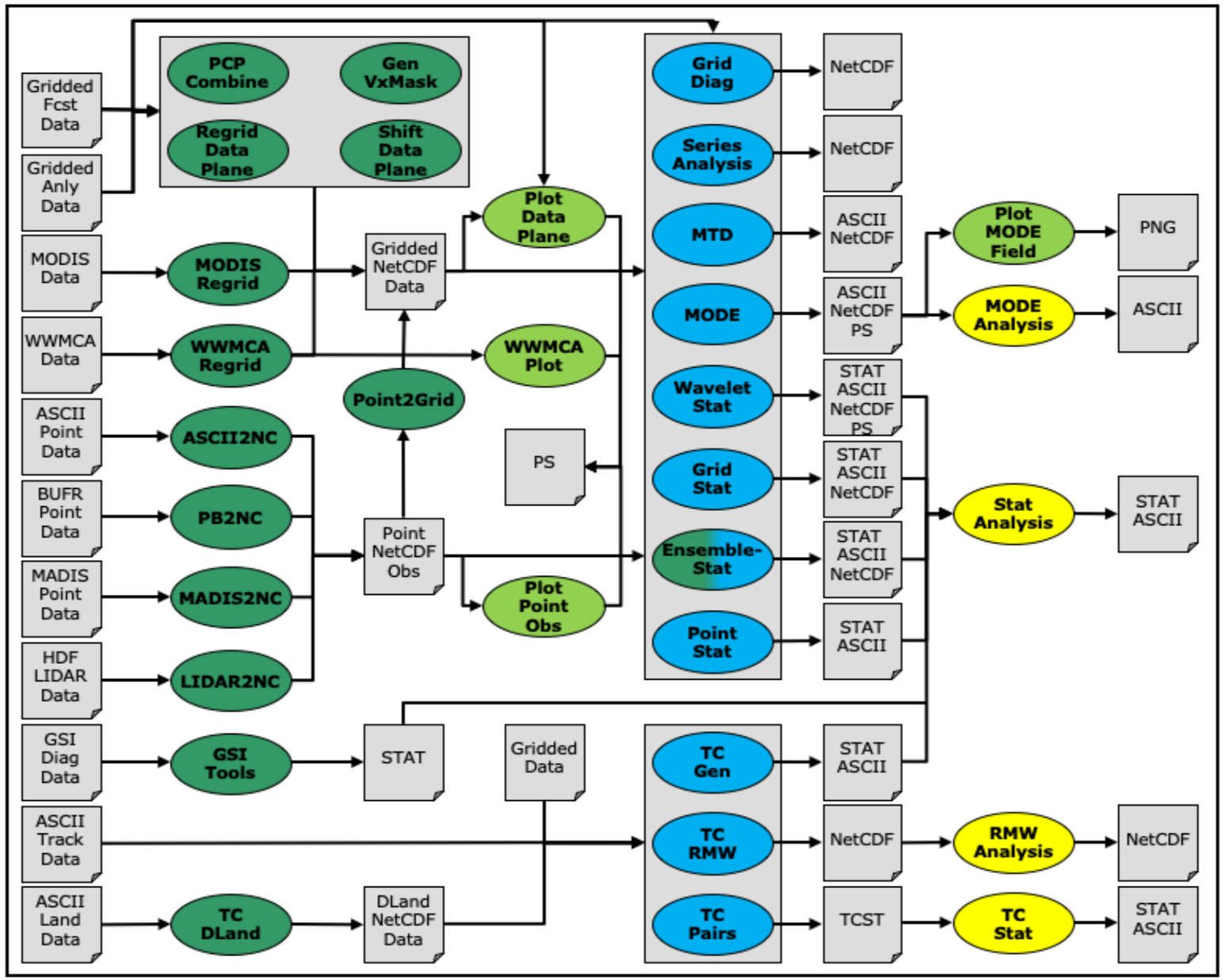
Python Wrappers for Verification Components

- **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



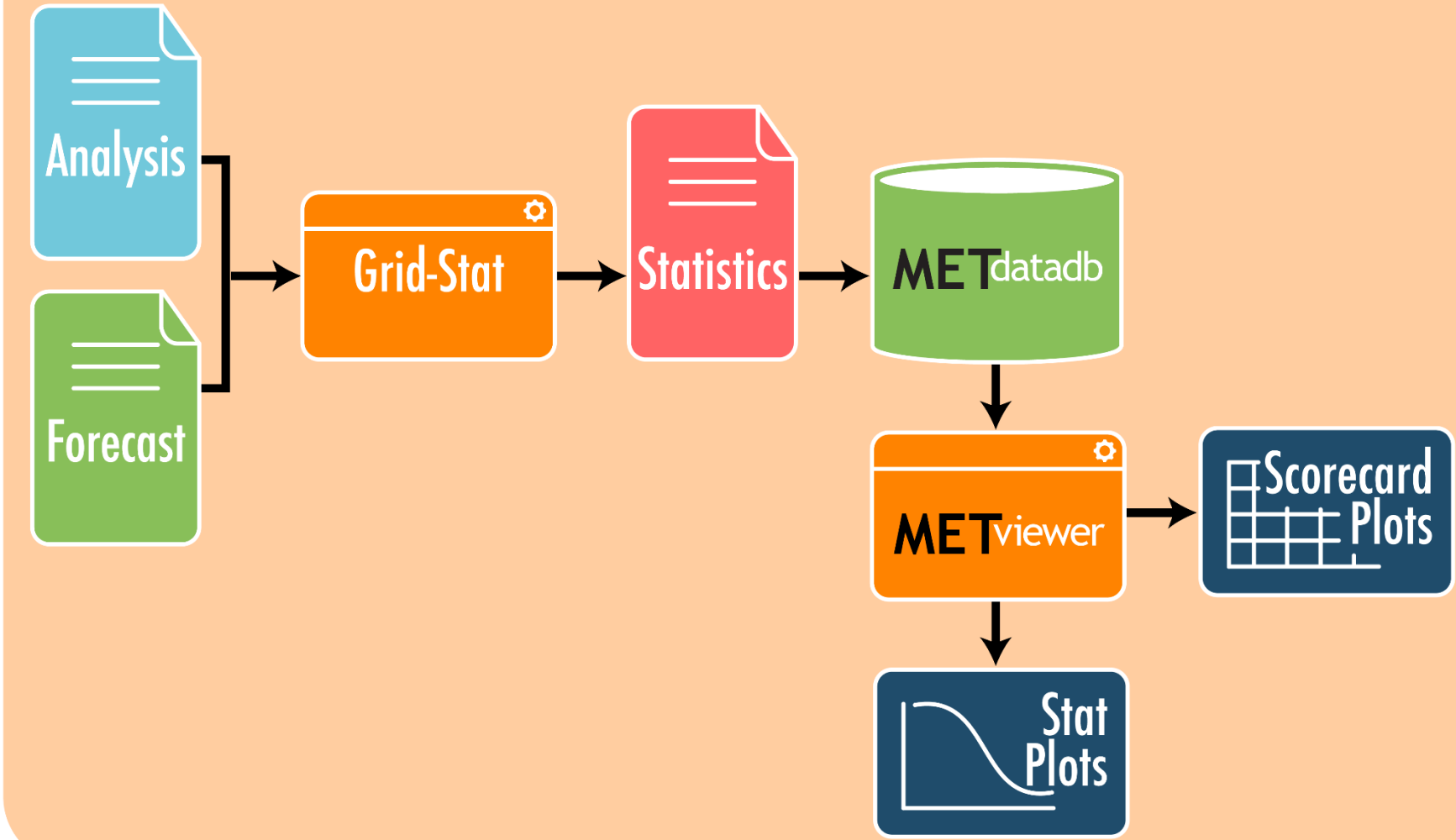
www.github.com/DTCenter/METplus

MET Overview v9.1

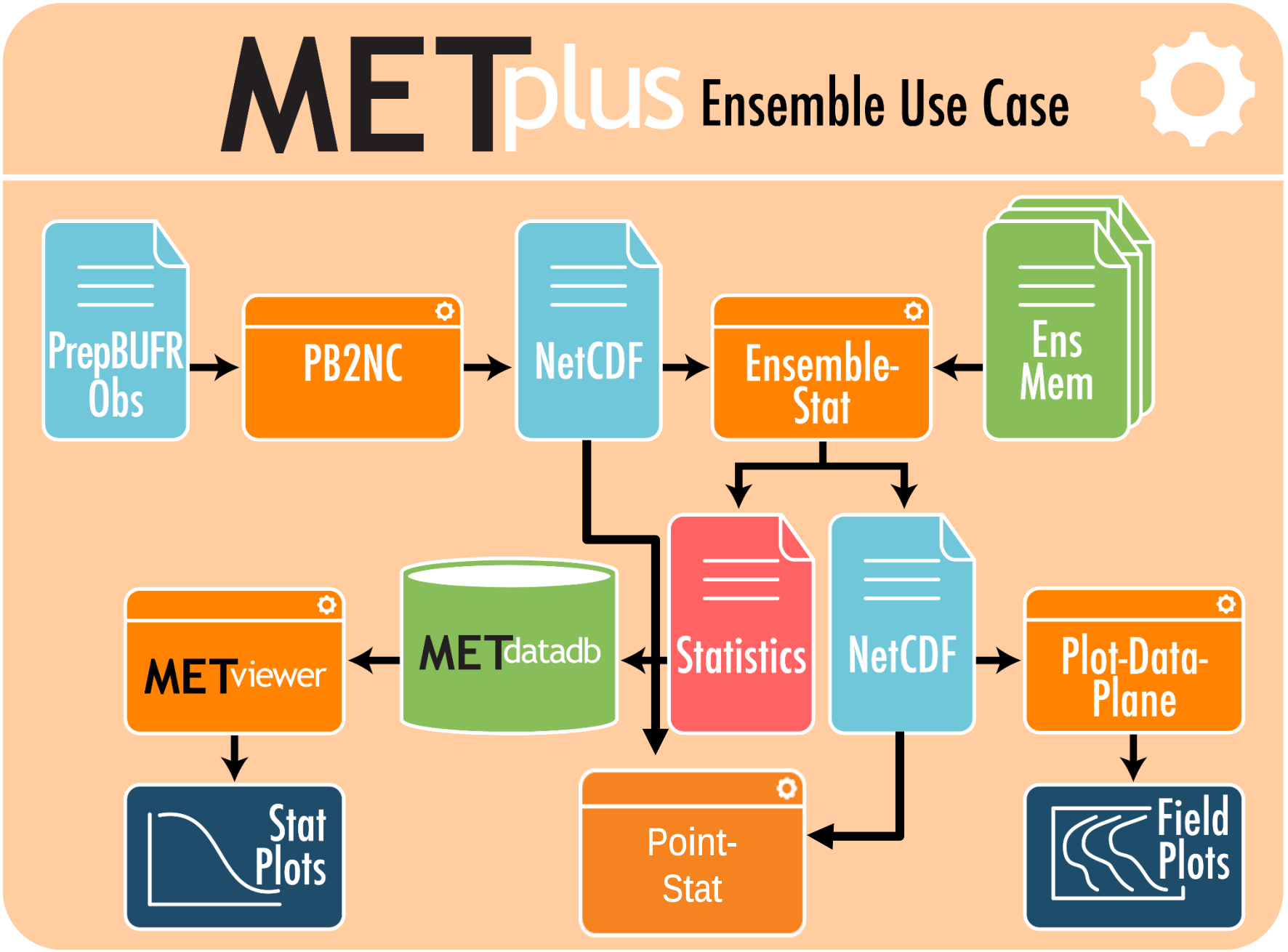


Legend

METplus Operational Categorical Statistics Use Case

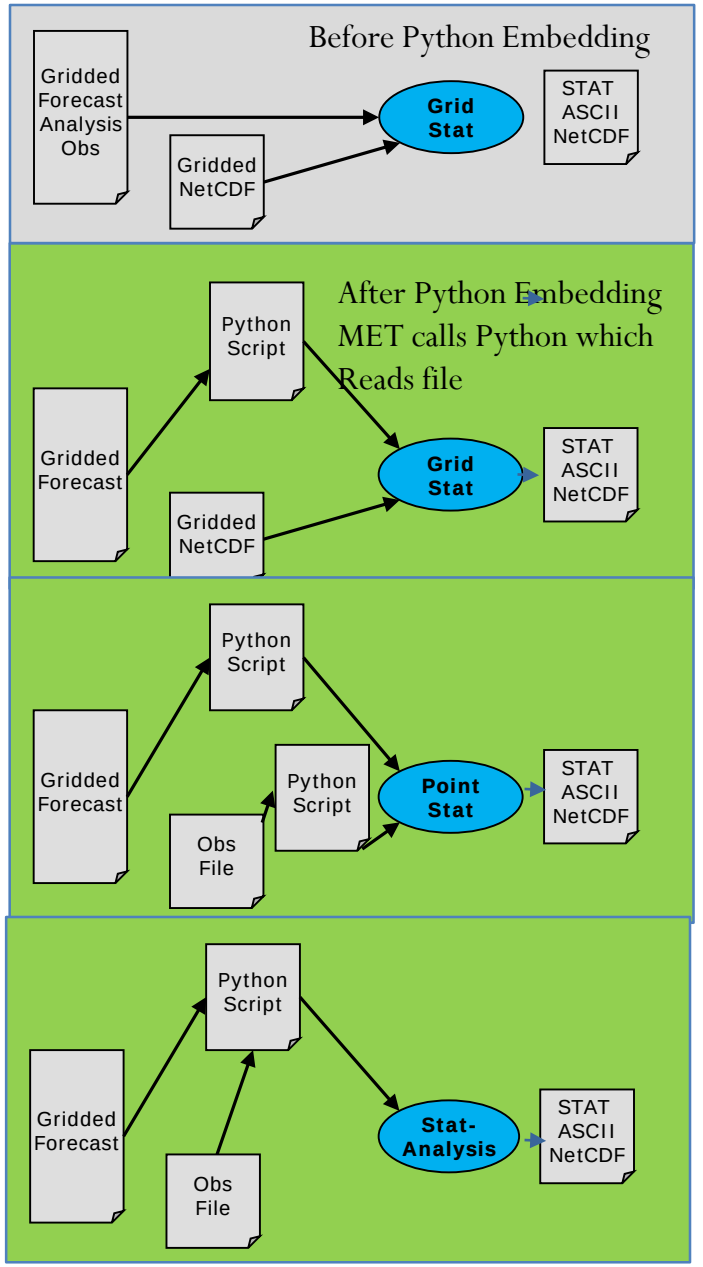


METplus Ensemble Use Case

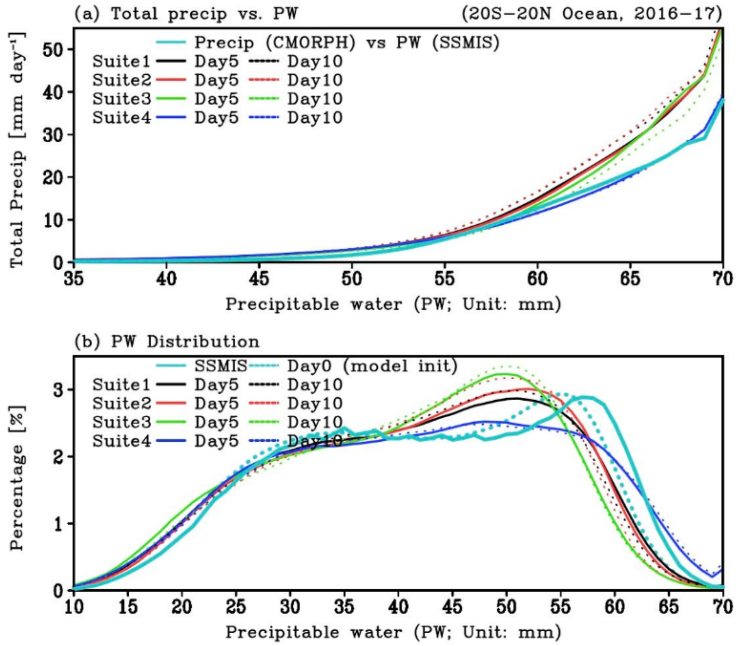


Providing More Flexibility to Support New Applications

- **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



Philosophy: One Tool-Many Applications, Example: GridDiag



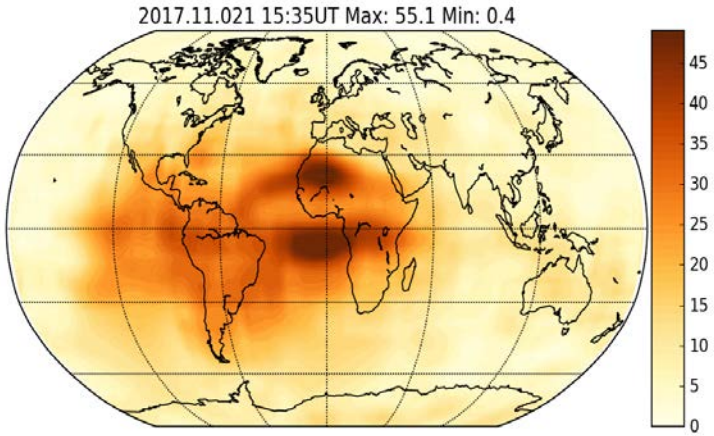
S2S

Correlate distributions of two fields

- 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 (Grid-Stat, Point-Stat, MODE, MTD)

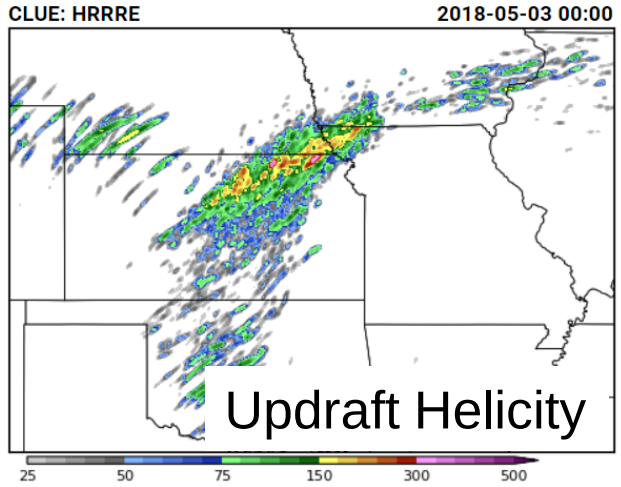
Space Weather

Adapt measures to regional max values



CAM

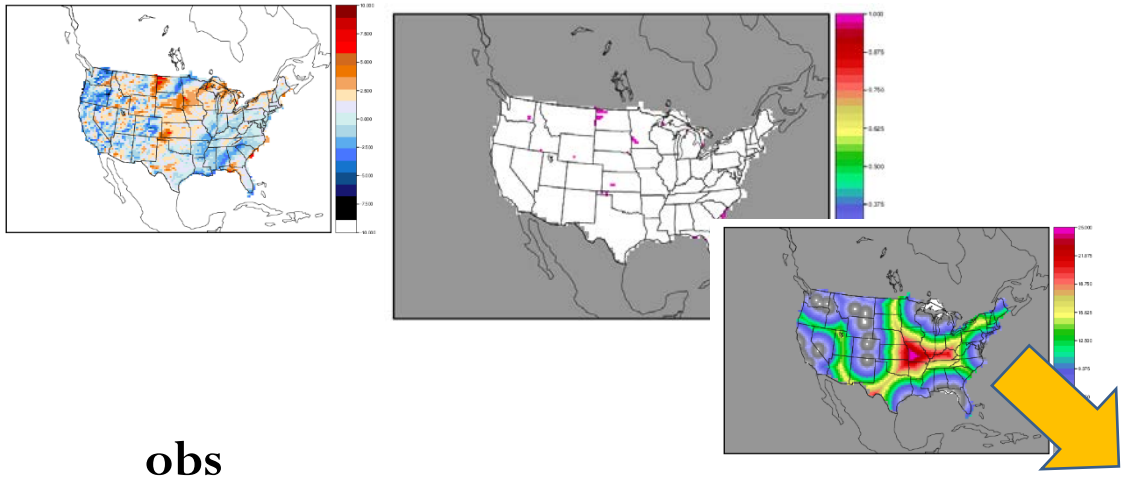
Normalize fields based on their sample climatology



Distance Maps

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

fcst



obs

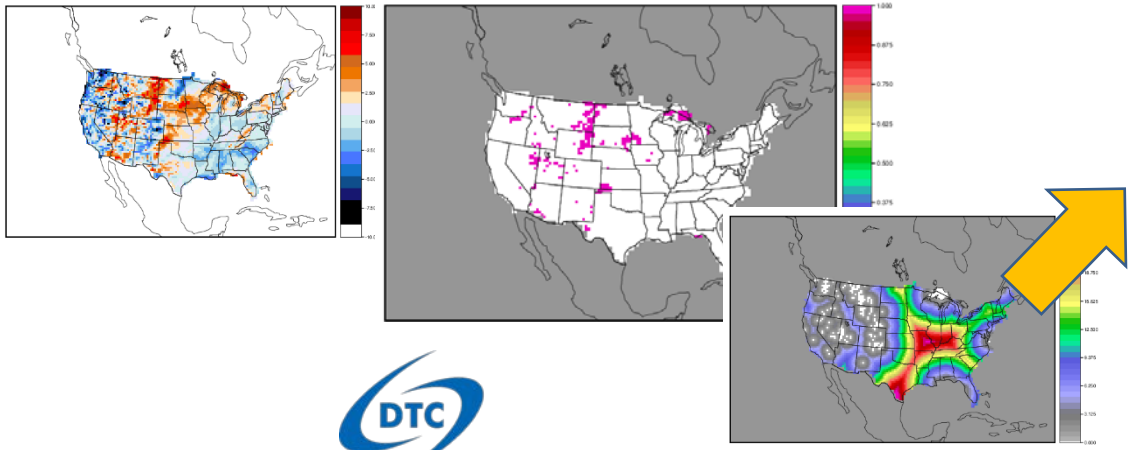


Table 8.7: Format information for DMAP (Distance Map) output line type.

DMAP OUTPUT FORMAT		
Column Number	**DMAP Column Name**	**Description**
24	DMAP	Distance Map line type
25	TOTAL	Total number of matched pairs
26	FY	Number of forecast events
27	OY	Number of observation events
28	FBIAS	Frequency Bias
29	BADDELEY	Baddeley's Δ Metric
30	HAUSDORFF	Hausdorff Distance
31	MED_FO	Mean-error Distance from forecast to observation
32	MED_OF	Mean-error Distance from observation to forecast
33	MED_MIN	Minimum of MED_FO and MED_OF
34	MED_MAX	Maximum of MED_FO and MED_OF
35	MED_MEAN	Mean of MED_FO and MED_OF
36	FOM_FO	Pratt's Figure of Merit from forecast to observation
37	FOM_OF	Pratt's Figure of Merit from observation to forecast
38	FOM_MIN	Minimum of FOM_FO and FOM_OF
39	FOM_MAX	Maximum of FOM_FO and FOM_OF
40	FOM_MEAN	Mean of FOM_FO and FOM_OF
41	ZHU_FO	Zhu's Measure from forecast to observation
42	ZHU_OF	Zhu's Measure from observation to forecast
43	ZHU_MIN	Minimum of ZHU_FO and ZHU_OF
44	ZHU_MAX	Maximum of ZHU_FO and ZHU_OF
45	ZHU_MEAN	Mean of ZHU_FO and ZHU_OF

Collaboration with Eric Gilleland, NCAR/RAL.

DTC Developmental Testbed Center

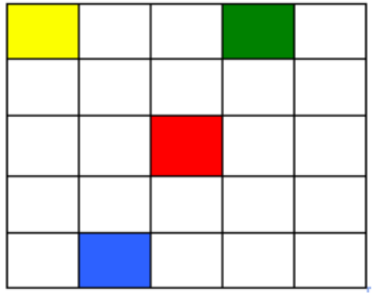
Point-Stat HiRA

- *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.

Model Forecast

White boxes = 0

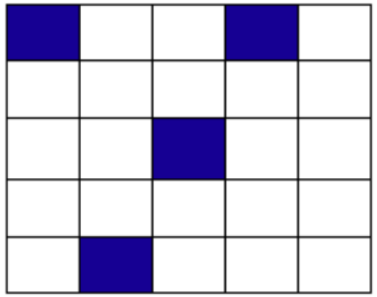
Colored boxes > 0



Threshold Forecast

Blue boxes = event

cat_thresh = [>0];



HiRA Ensemble:

- Write ECNT using neighborhood ensemble.

HiRA Probabilities:

- 1x1 NBRHD = 1/1
- 3x3 NBRHD = 1/9
- 5x5 NBRHD = 4/25

```
hira = {
 // 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;
};
```

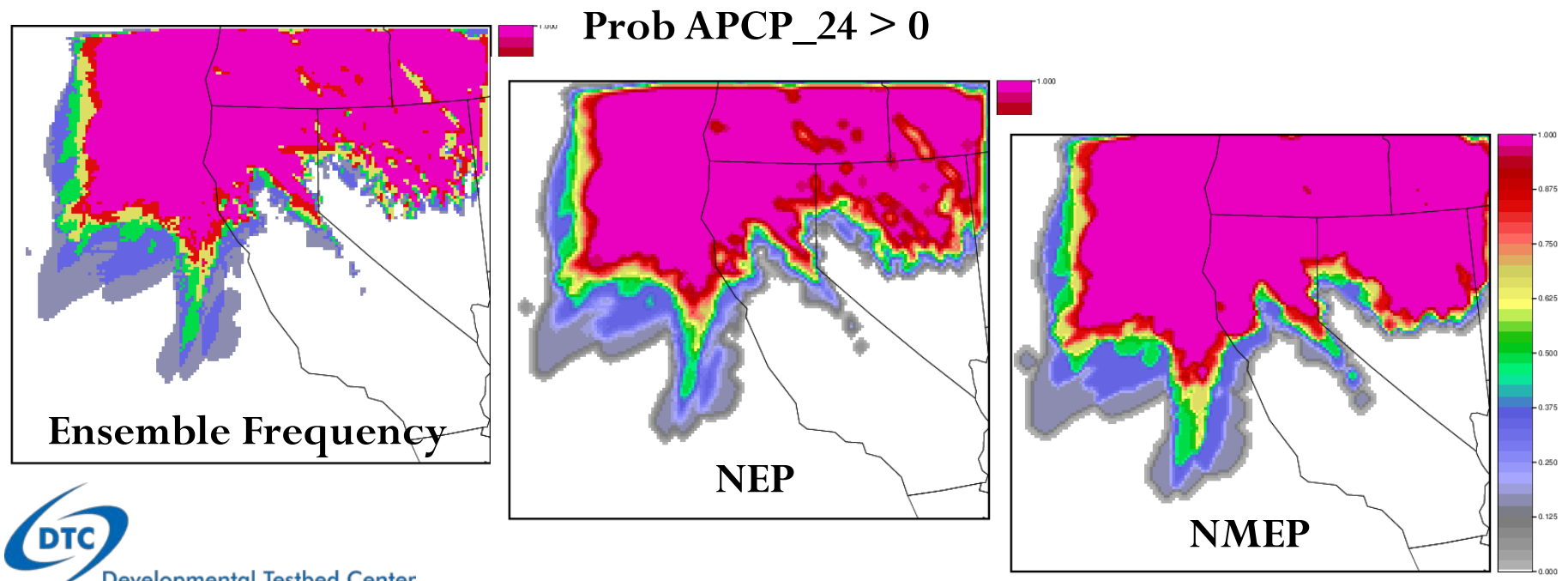
```
output_flag = {
 …
 pct    = STAT;
 pstd   = STAT;
 pjc    = STAT;
 prc    = STAT;
 ecnt   = STAT;
 …
};
```

DTC Developmental Testbed Center

Neighborhood Probabilities

- *Collaboration with NOAA MMM*
- **Schwartz and Sobash (2017).**
- **Ensemble-Stat** pre-processes to computes frequency of 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;
   nmep = TRUE;
   ...
}
```

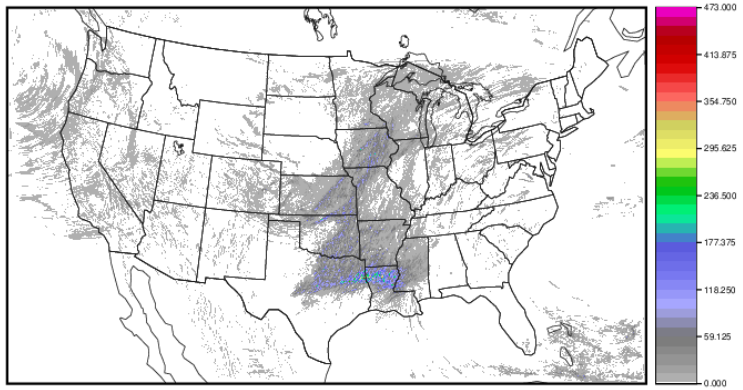


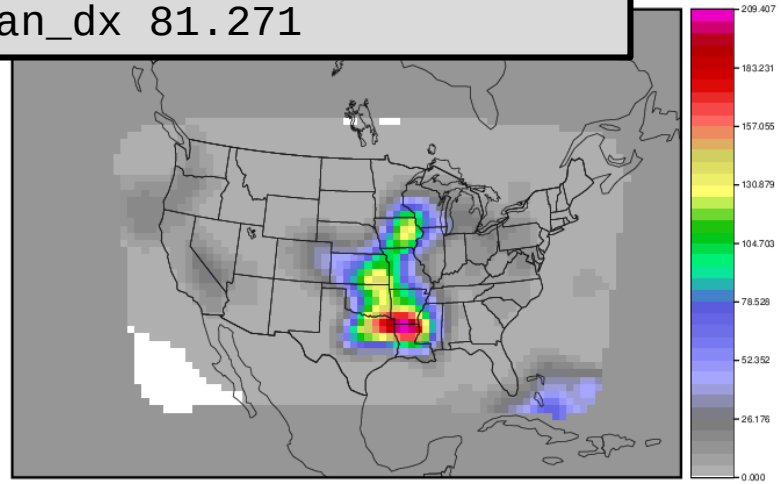
DTC Developmental Testbed Center

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

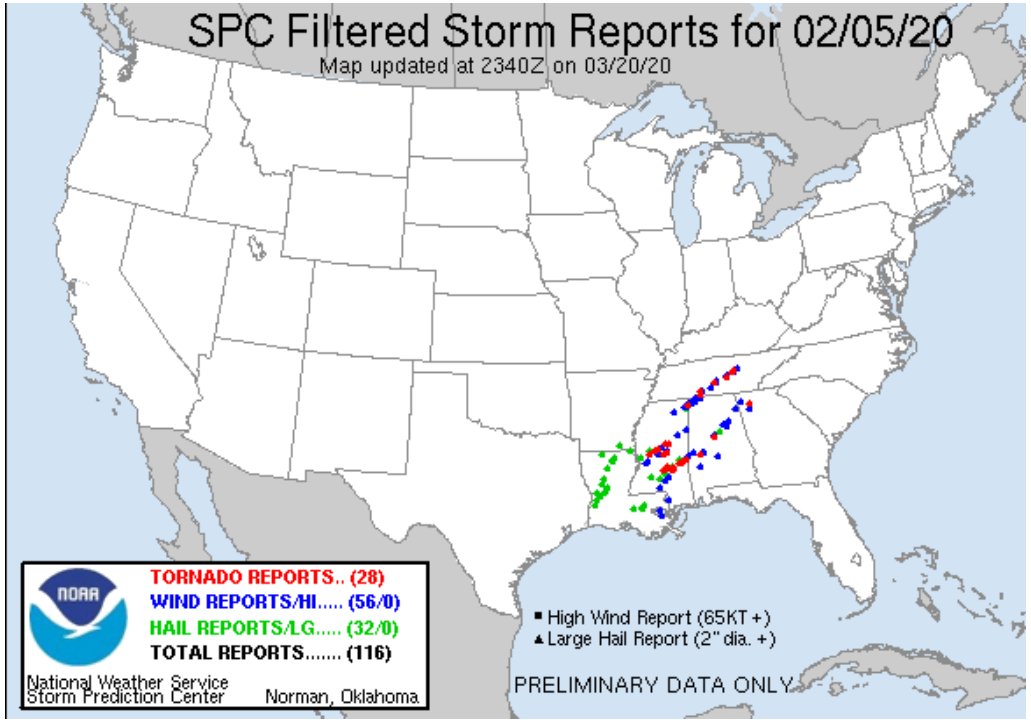
```
regrid_data_plane HRRR_20190518_013_036_24h.nc G211 \
  regrid_data_plane_HRRR_MXUPHL_24_MAXGAUSS.nc \
  -field 'name="MXUPHL_24"; level="A24";' \
  -method MAXGAUSS -width 27 \
  -gaussian_radius 120 -gaussian_dx 81.271
```

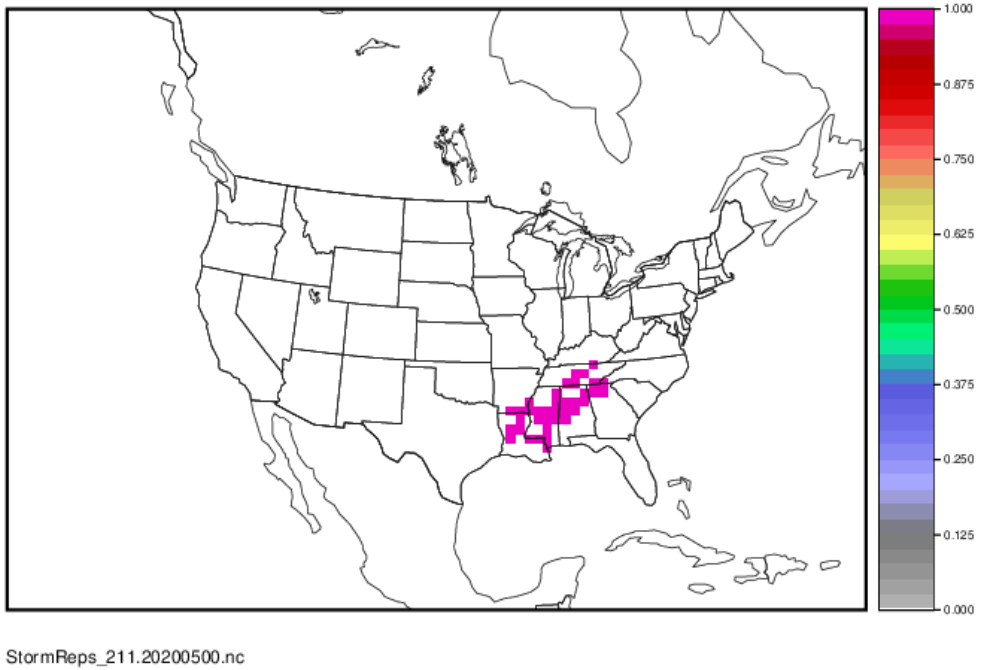




Practically Perfect Prog and Data Thinning

- *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.





StormReps_211.20200500.nc

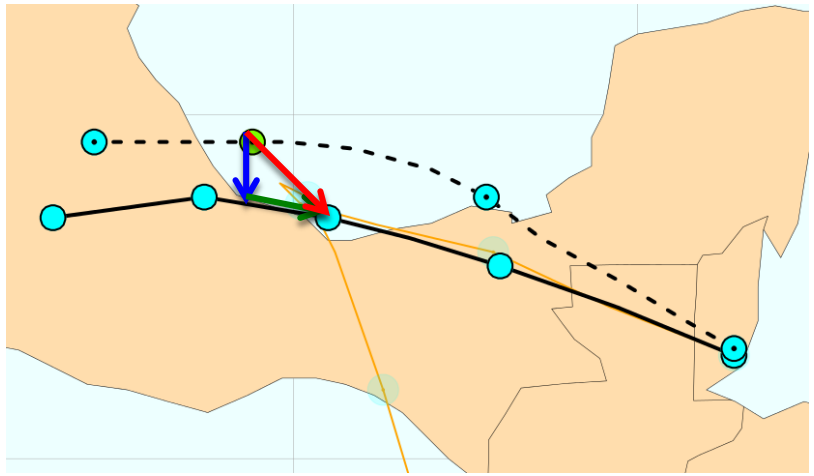
Use of Scorecards for Short-Range Forecasting

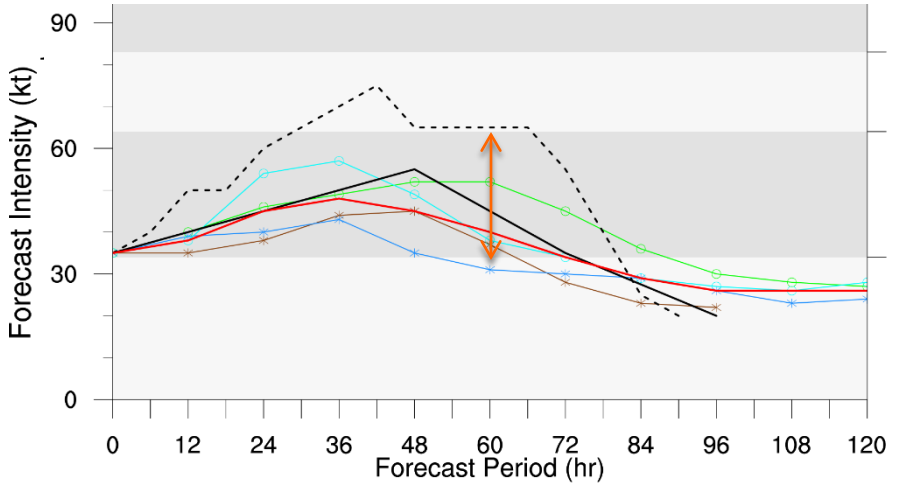
				Daily Domain								CONUS							
				12 hr	18 hr	21 hr	24 hr	27 hr	30 hr	33 hr	36 hr	12 hr	18 hr	21 hr	24 hr	27 hr	30 hr	33 hr	36 hr
Fraction Skill Score	Composite Reflectivity	49	>=20																
			>=30							▲									
			>=40																
	3 hr Accumulated Precipitation	49	>=0.25																
			>=0.5																▲
			>=1.0																
CSI	Composite Reflectivity	1	>=20																
			>=30							▲									
			>=40							▲	▲						▲	▲	
	3 hr Accumulated Precipitation	1	>=0.25							▲									
			>=0.5							▲									▲
			>=1.0																
Bias	Composite Reflectivity	1	>=20						▼							▲	▲	▲	
			>=30									▼	▼						
			>=40								▼	▼	▼	▼	▼	▼			▼
	3 hr Accumulated Precipitation	1	>=0.25		▼				▼			▲	▼	▼	▼				
			>=0.5					▼	▼				▲				▼	▼	
			>=1.0						▼							▼	▼	▼	
RMSE	Temperature	1	sfc										▲	▲		▼			
	Dew Point	1	sfc									▼							
	Wind Speed	1	sfc	▼								▼	▼			▲	▲		
ME	Temperature	1	sfc	▼	▲	▼	▼	▼	▼	▼	▼	▼	▼	▼		▼	▼	▲	▼
	Dew Point	1	sfc					▲	▲	▼	▼	▼		▼	▼			▲	
	Wind Speed	1	sfc	▼	▼			▼	▼	▼	▼	▼	▼	▼					

▲	NSSLfv3 is better than HRRRv3_cluegrid at the 99% significance level
	NSSLfv3 is better than HRRRv3_cluegrid at the 95% significance level
	No statistically significant difference between NSSLfv3 and HRRRv3_cluegrid
	NSSLfv3 is worse than HRRRv3_cluegrid at the 95% significance level
▼	NSSLfv3 is worse than HRRRv3_cluegrid at the 99% 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

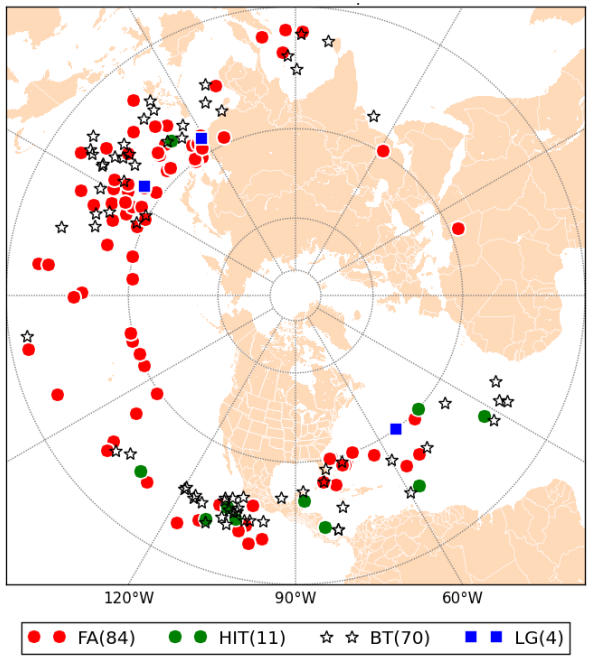
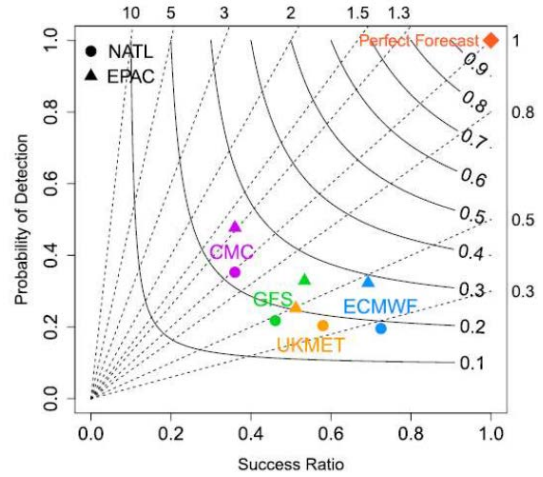


Figure: Tropical cyclogenesis verification for the NH for 2016. Symbols represent the Best Track (black), hits (green), late Genesis (blue) and false alarms (red).

Halperin, 2017

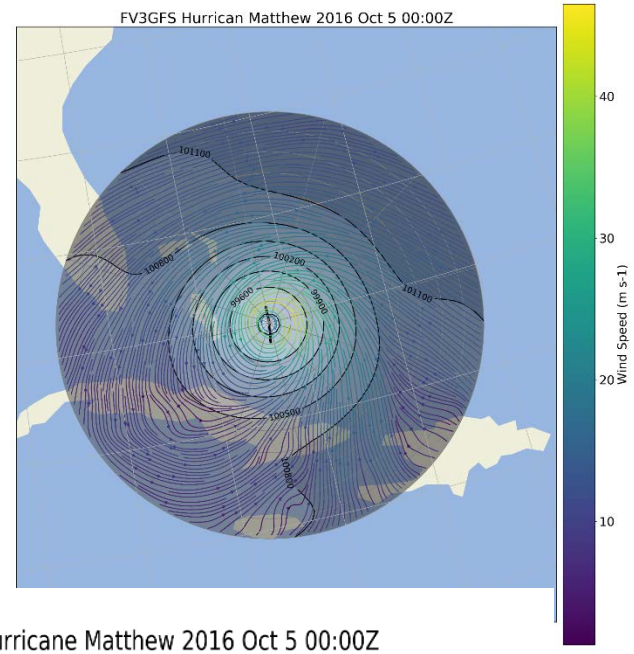
```
tc_gen -config TCGenConfig_dev \
-genesis suite[1-4]/2016*/genesis.fort.66* \
-track nhc_atcf/*/*2016.dat \
-v 3 -log run_tc_gen.log
```

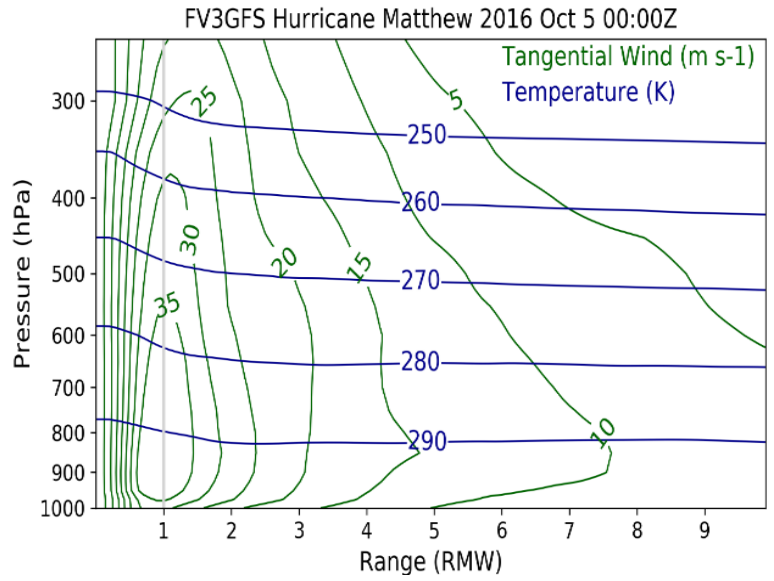


Developmental Testbed Center

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





Use of TDR for TC Development

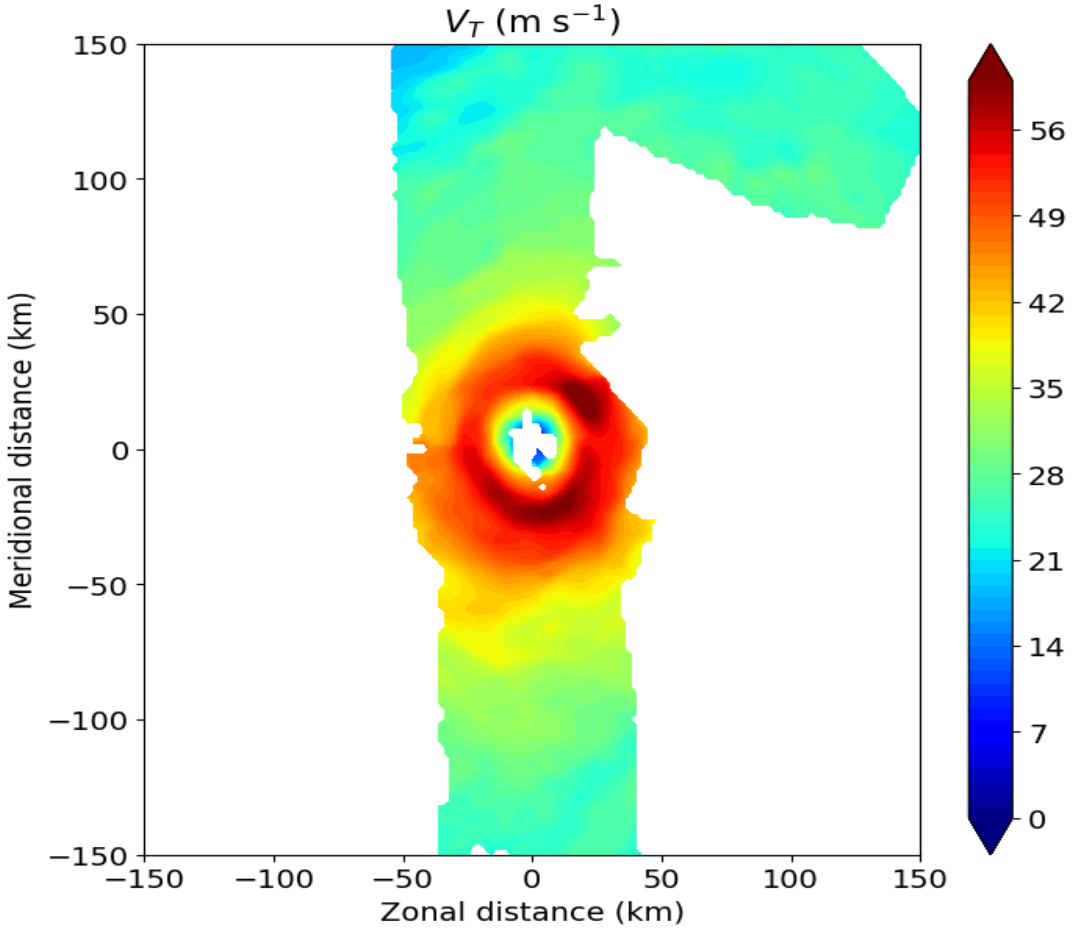


Image courtesy of Michael S. Fischer, Robert F. Rogers, Paul D. Reasor at NOAA/AOML/HRD

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

Growing METplus Community

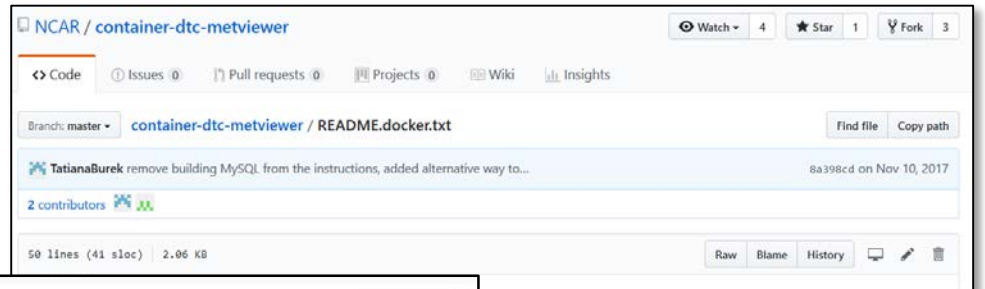
Research Institutions

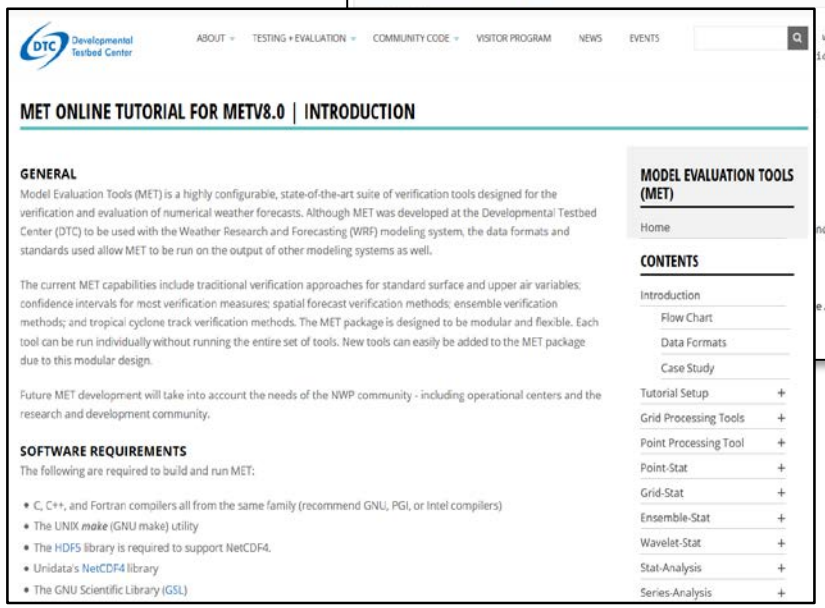
- NCAR (RAL, MMM, CGD)
- Naval Research Lab
- SBU, U of Illinois Urbana-Champaign, U of Wisc Madison, OU, UND, NC State, Purdue, Albany, etc…
- Labs NOAA Research (GSL, NSSL, PSL, ARL, GFDL)
- NASA Coordinated Community Modeling Center (for Space Wx)
- Army Research Lab

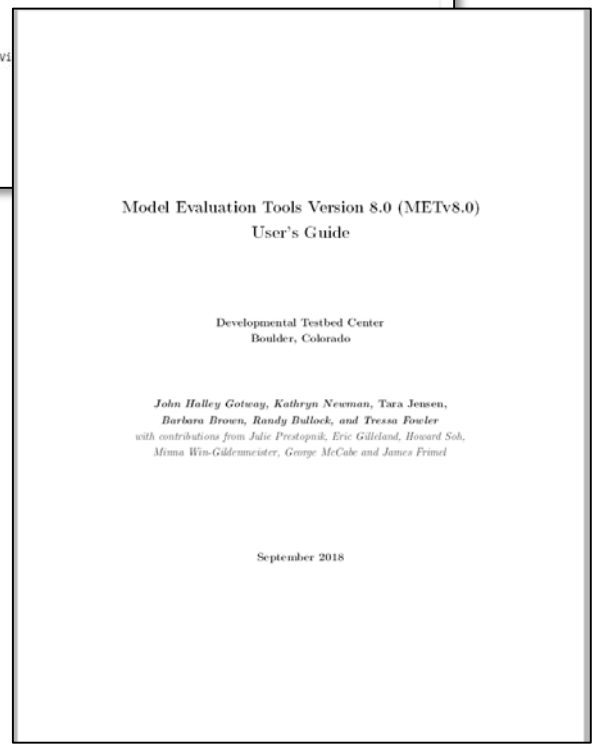
Operational Centers adopted portions or all of METplus

- 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



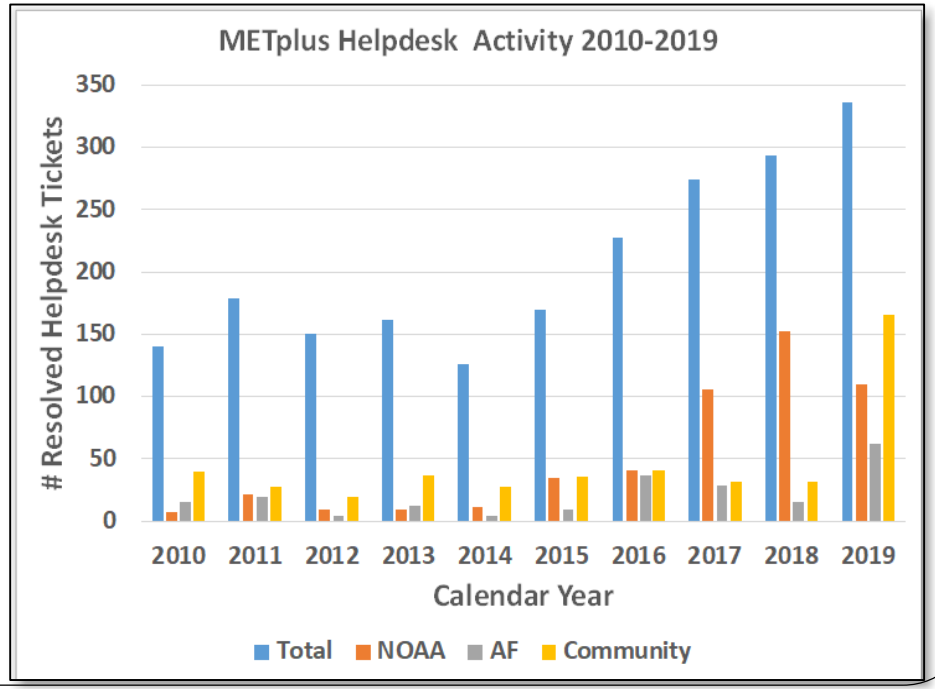




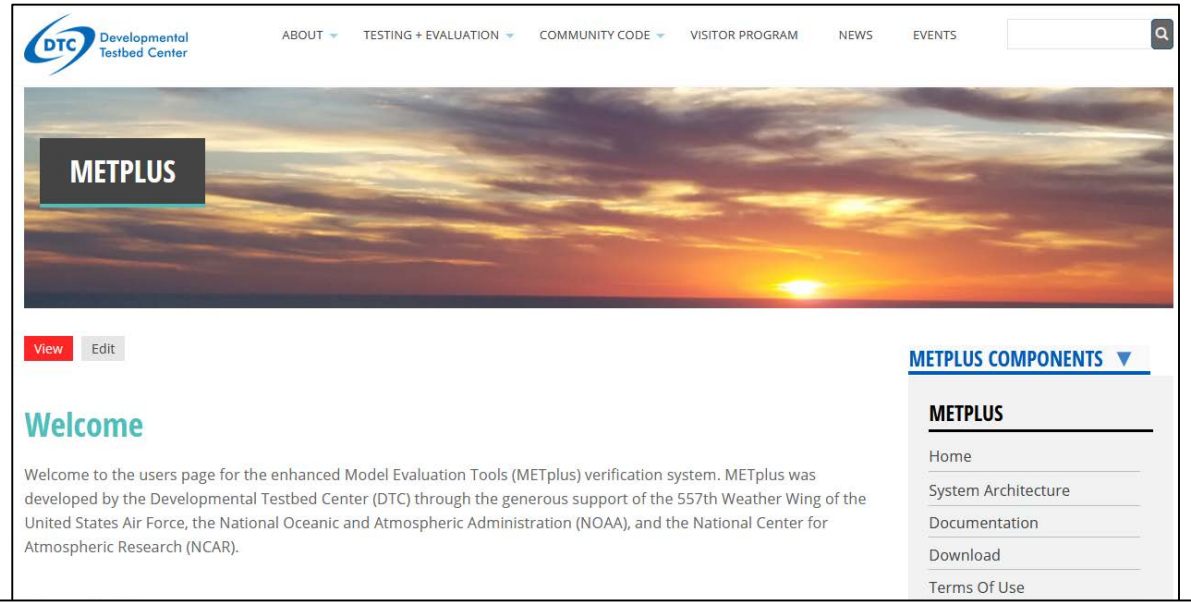
GitHub

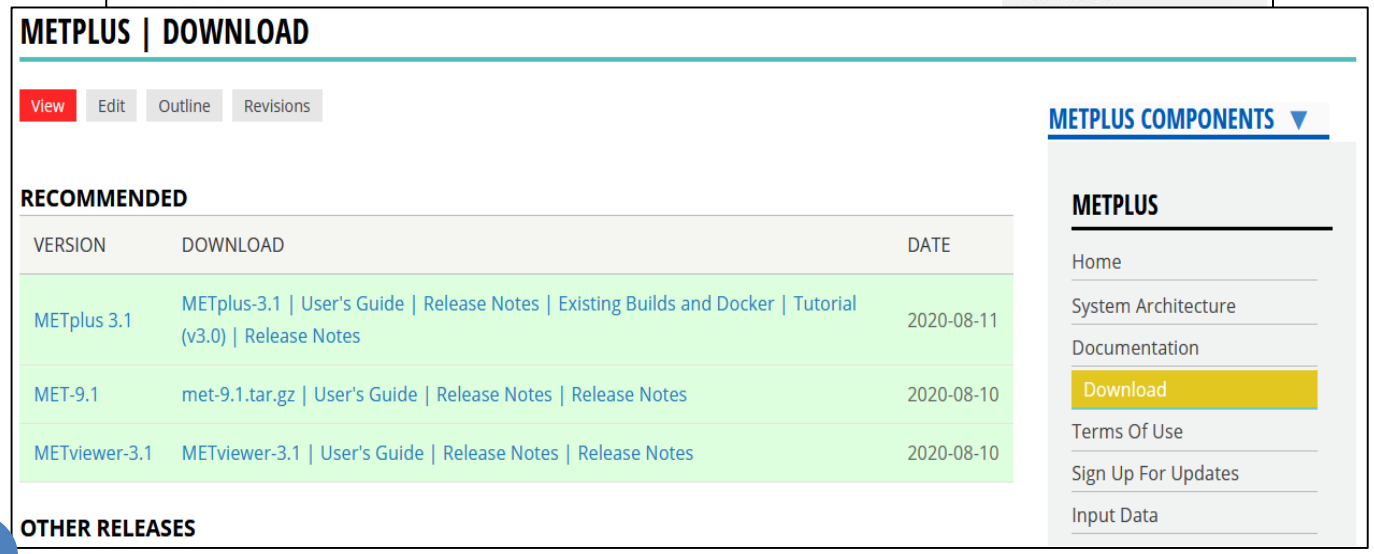
docker

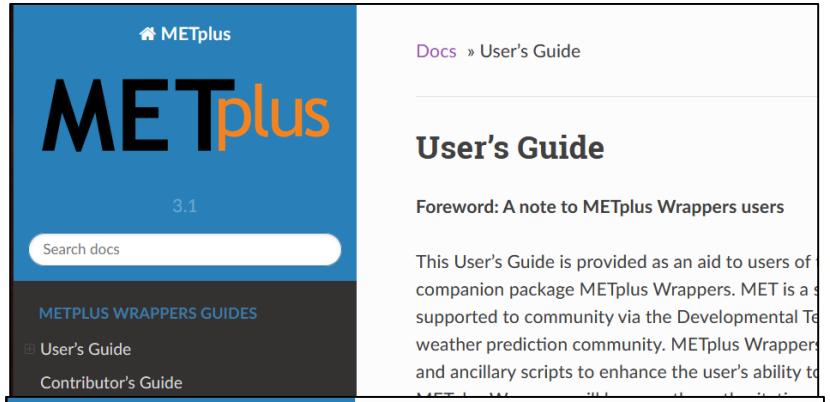
Developing YouTube-style Training Videos

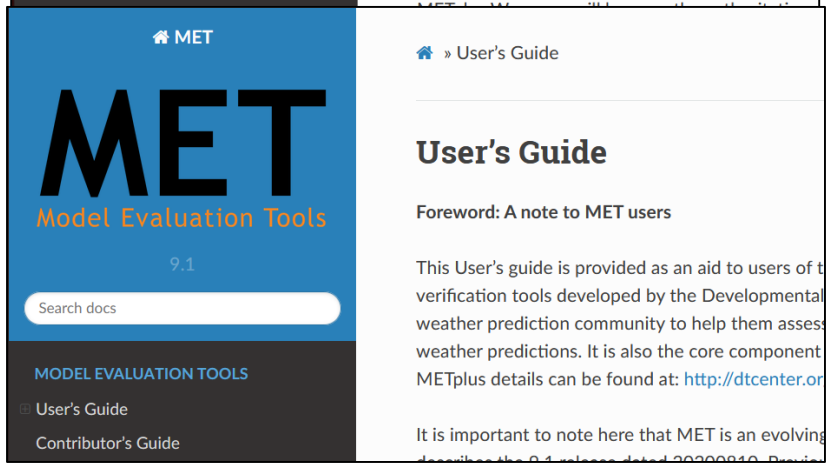


User Support









- 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

METplus