METplus

Fostering International Collaboration Through a Unified Verification, Validation, and Diagnostics Framework: METplus

Tara Jensen[1], Marion Mittermaier[2], Jason Levit[3], Elizabeth Satterfield[4], Evan Kuchera[5], Louisa Nance[1]

[1]National Center for Atmospheric Research & Developmental Testbed Center

[2]The Met Office

[3]NOAA/Environmental Modeling Center

[4]Naval Research Laboratory

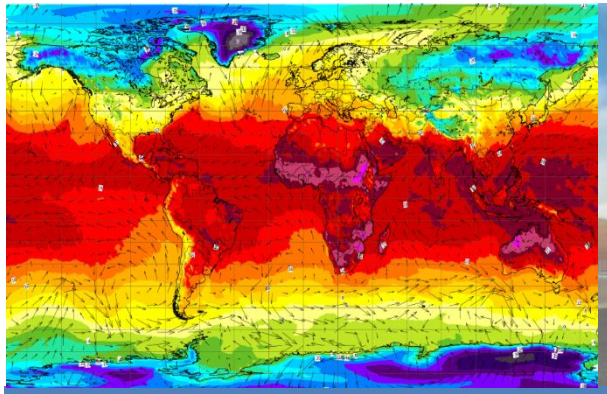
[5]Air Force 557th Weather Wing

2020 IVMW-Online-19 November 2020

DTC Developmental Testbed Center

Why MET and then METplus?

Forecasters



Operational Centers

Universities and National Laboratories

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

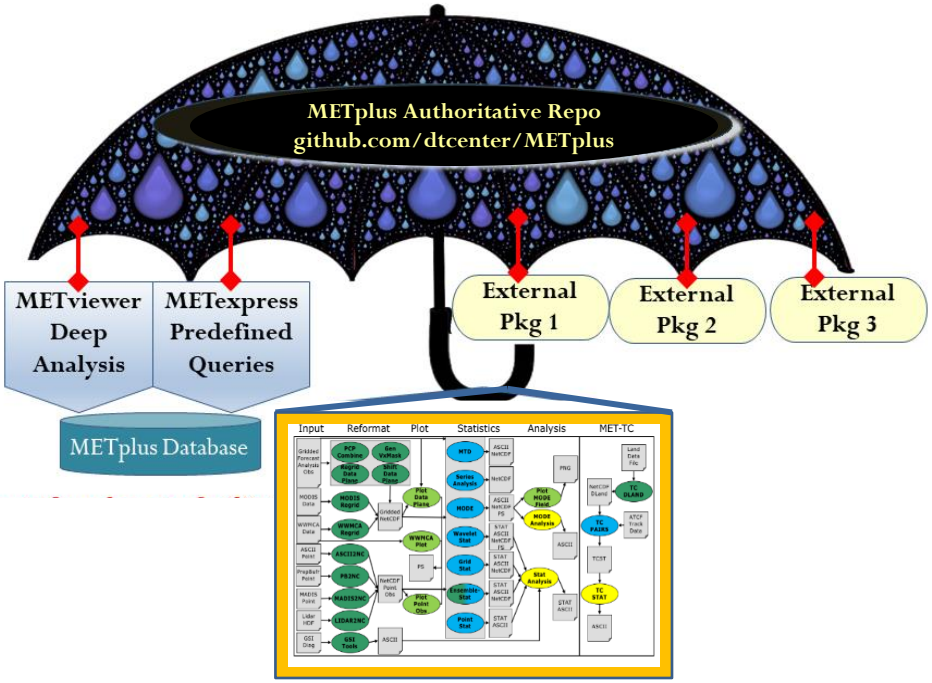
Allows researchers and operational scientists to speak a "common verification" language

METplus

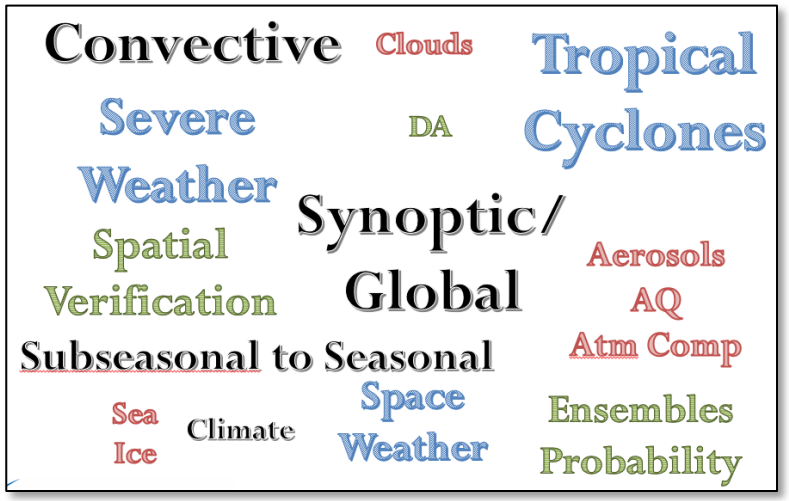
User support of unified package provides greater opportunity to train all on verification best practices

What is METplus?

- Over 100 traditional statistics and diagnostic methods for both point and gridded datasets
- 15 interpolation methods
- Applied to many spatial and temporal scales
- Developed to allow for easy sharing of config files for reproducible results
- 3500+ users; US and Int'l
  - 65% International
  - 35% US
  - Primarily Government/Academic, some Private/Commercial



METplus Examples/Use-Case In Development



Goal: A suite of Python wrappers around

- Model Evaluation Tools (MET - core)
- Database and display systems (core)
- Plotting
  - METviewer/METexpress User Interfaces
  - METviewer Batch Engine
  - Python plotting scripts
- Communication between MET & python algorithms
- Using manage_externals to connect repos

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…
- NOAA Research Labs (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
- Fleet Numerical is considering once Naval Research Lab has transitioned
- **The Met Office (agreement executed)**
- **Other UM partners (i.e. Australian BoM, S. African WS, India's NCMRWF) also considering contributing**
- **Shanghai Met Bureau**
- **Argentina's NMS**
- **Central Weather Bureau (Taiwan) is considering adoption**

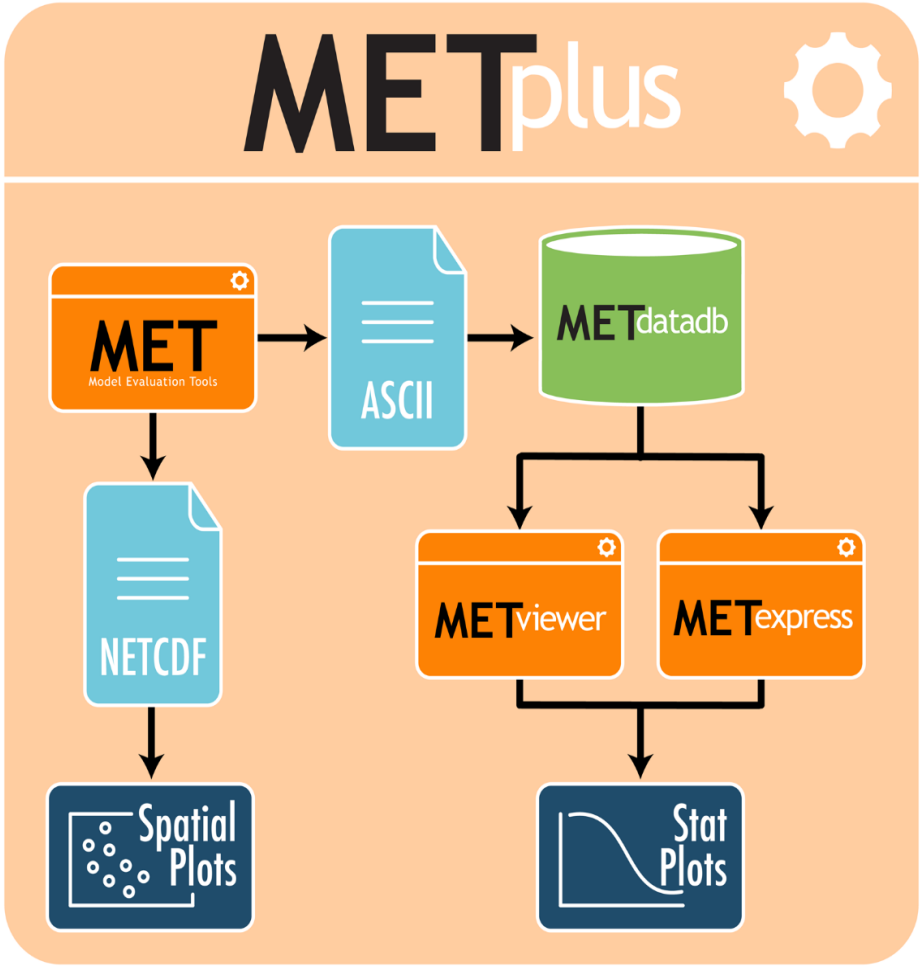
METplus

Python Wrappers for Verification Components

Python wrappers around:

- MET
  - Statistical engine
- METviewer
  - GUI and batch engine
- METexpress
  - Simplified plotting interface
- METdatadb
  - Verification database loader and other utilities
- METcalcpy
  - Aggregation, event equalization, pairwise differencing, confidence intervals
  - Other calculations related to diagnostics
- METplotpy
  - Plotting for METviewer and maybe portions of METexpress
  - Plotting of diagnostic fields

- Python Embedding:
  - Communication between MET & Python algorithms

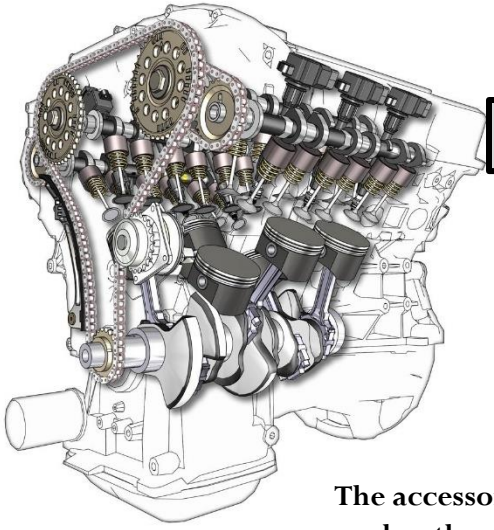


www.github.com/DTCenter/METplus

METplus Analogy

METplus

Drives the process and keeps everything together



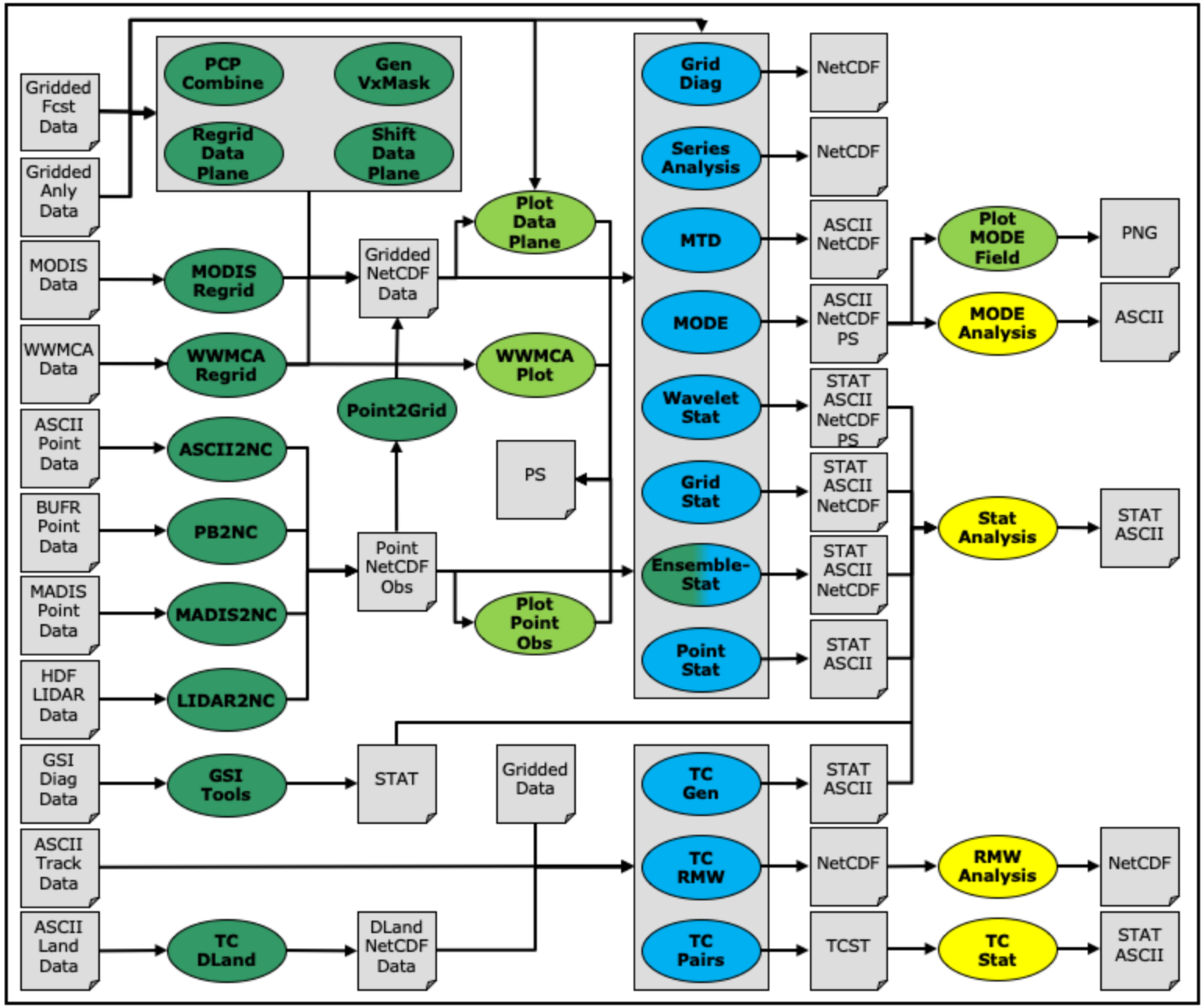
MET

Statistical engine

METviewer
METexpress
METcalcpy
METplotpy

The accessories that makes the process better

MET Overview v9.1

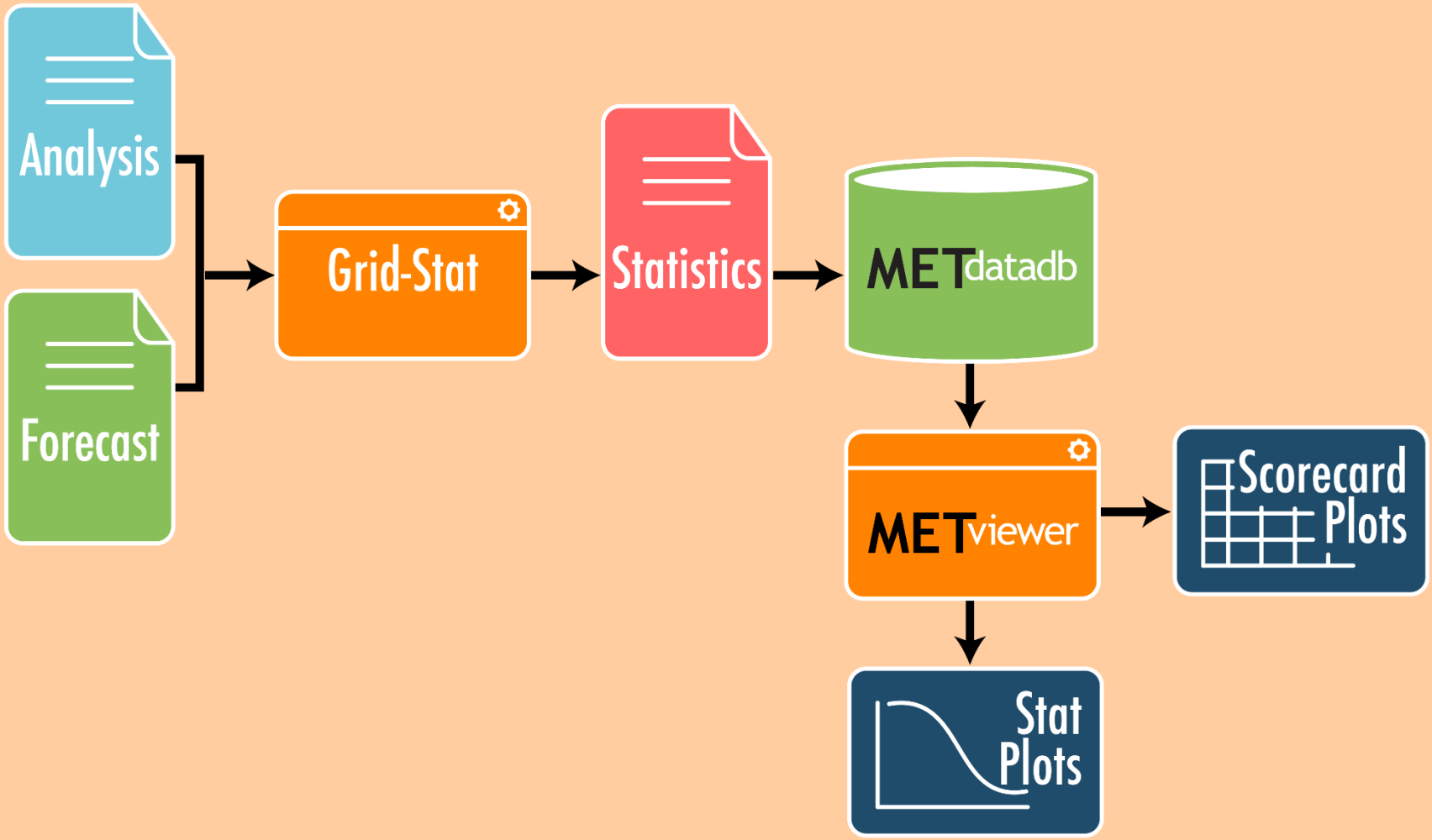


Legend

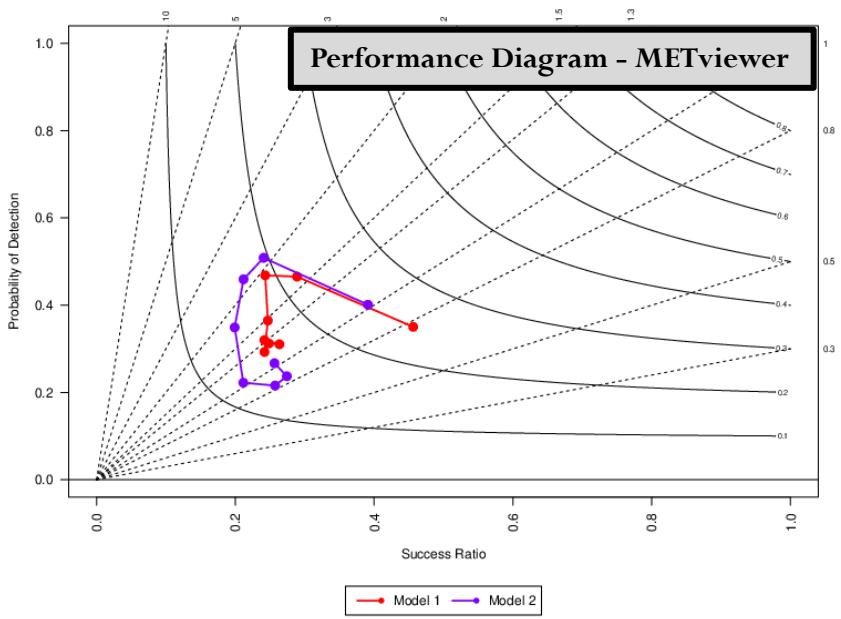
MET Tool Categorization

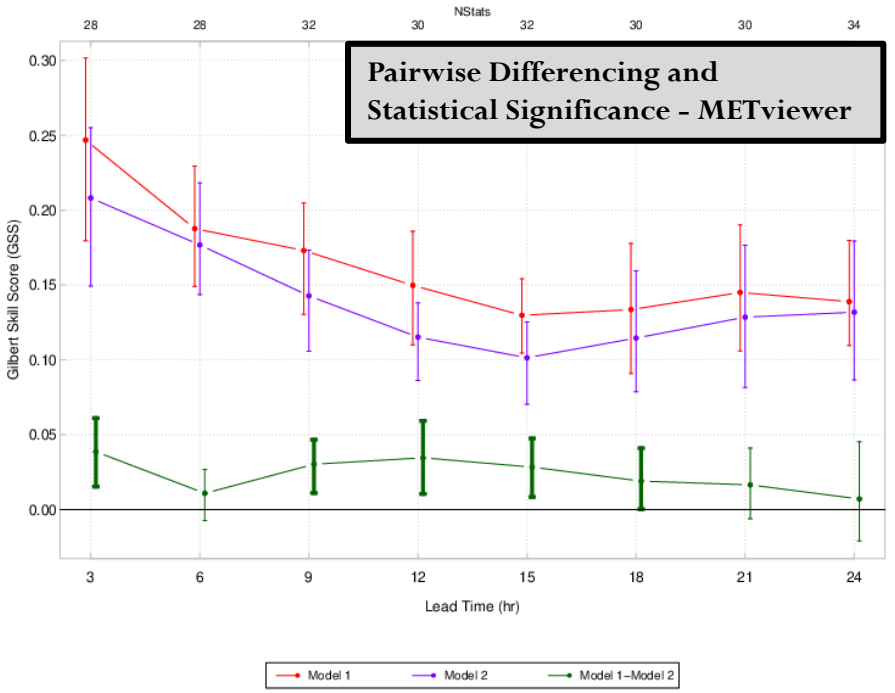
Traditional	
Grid-Stat, Point-Stat, Series-Analysis *Contingency table statistics (CTS)* *Continuous statistics* *Probability forecast statistics* *Confidence intervals*	**Ensemble-Stat** *CRPS, CRPSS* *Rank prob., Prob. Integral Transform (PIT), and Relative Position histograms* *Spread/Skill* *Ignorance* *Confidence intervals*
Spatial	
MODE *Location differences* *Geometric attribute differences* *Intersection area* *Intensity distributions & differences* *CTS measures*	**MODE-TD** *Time and location differences* *Volume differences* *Velocity differences* *Intersection volume* *Intensity distributions & differences*
Wavelet-Stat *MSE by scale* *Energy by scale* *Intensity-scale skill score*	**Grid-Stat and Point-Stat** *FSS, HiRA* *Distance Measures: MED, Baddeley, Hausdorff, Zhu, etc.*
Tropical Cyclones and Diagnostics	
MET-TC *Track error (along, cross, total)* *Intensity errors (pressure, wind)* *Rapid intensification/weakening errors* *CTS measures of TC genesis* **TC-GEN** *CTS measures of TC genesis*	**Grid-Diag** *Distributions of fields for use in contour plots* **TC-RMW** *Radius of maximum wind errors and metrics*

METplus Operational Categorical Statistics Use Case



Operational Use Case Example



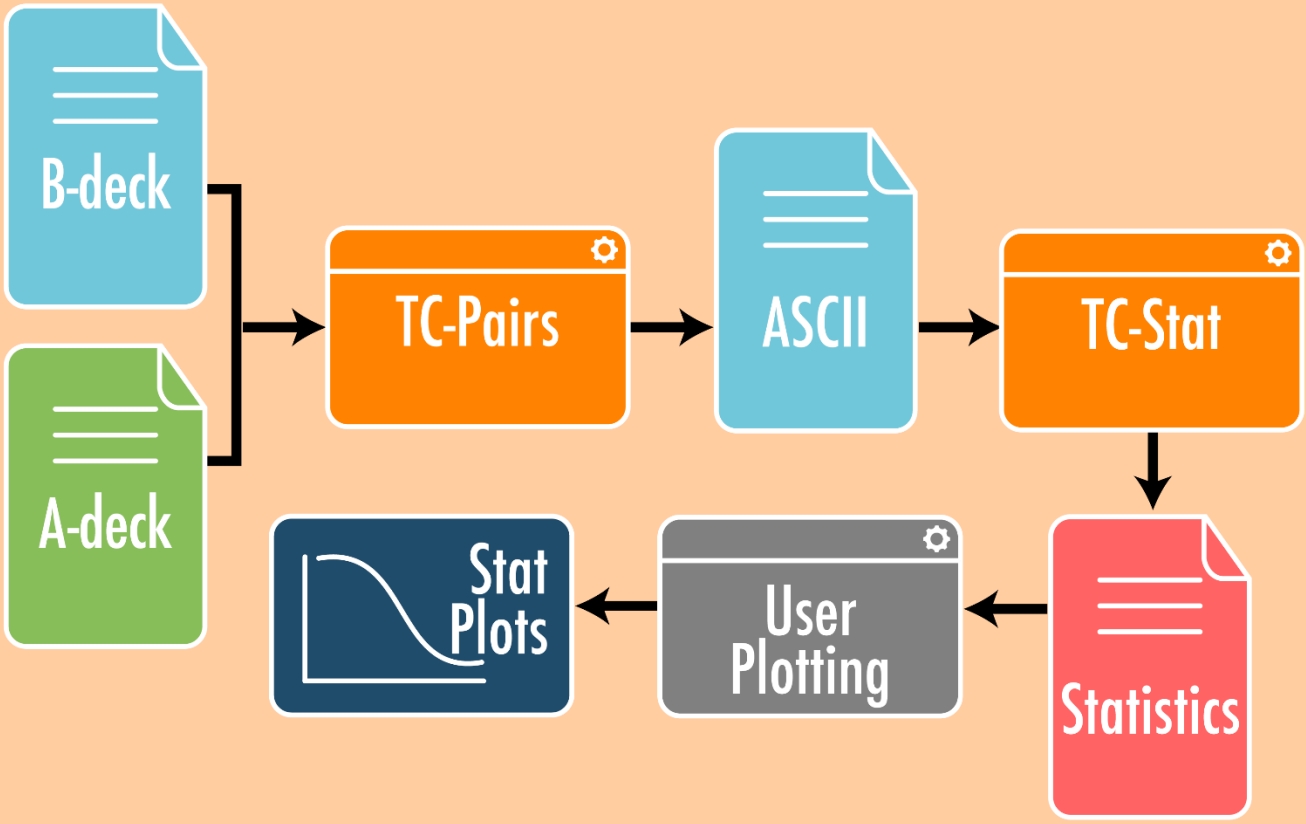


			CONUS				EAST				WEST			
			60000	120000	180000	240000	60000	120000	180000	240000	60000	120000	180000	240000
CSI	APCP_03	>=0.254		▲	▲						▲	▲		
		>=2.540		▲								▲		
		>=25.400	▲	▲		▲	▲	▲		▲	▲	▲	▲	▲
	APCP_06	>=0.254			▲						▲	▲	▲	
		>=2.540		▲	▲	▲		▲	▲	▲		▲		
		>=25.400												
PODY	APCP_03	>=0.254			▲	▲		▲	▲	▲	▼		▲	
		>=2.540	▼		▲		▼		▲	▲				
		>=25.400	▲	▲		▲	▲	▲		▲	▲	▲	▲	▲
	APCP_06	>=0.254	▼		▲	▲		▲	▲	▲	▼	▼	▲	▲
		>=2.540	▼		▲	▲	▼	▲	▲	▲		▼	▲	
		>=25.400												
FAR	APCP_03	>=0.254		▲		▼					▲	▲		
		>=2.540	▲	▲							▲	▲		
		>=25.400	▲			▲	▲			▲	▲		▲	▲
	APCP_06	>=0.254				▼		▼		▼	▲	▲		▼
		>=2.540	▲	▲			▲	▲			▲	▲		
		>=25.400												
FBIAS	APCP_03	>=0.254			▼	▼		▼	▼	▼	▲	▲	▼	▼
		>=2.540	▲	▲	▼	▼	▲		▼	▼		▲	▼	
		>=25.400				▲	▲					▲		
	APCP_06	>=0.254			▼	▼		▼	▼	▼	▲	▲	▼	▼
		>=2.540	▲	▲	▼	▼	▲		▼	▼	▼	▲	▼	▲
		>=25.400				▲				▲				

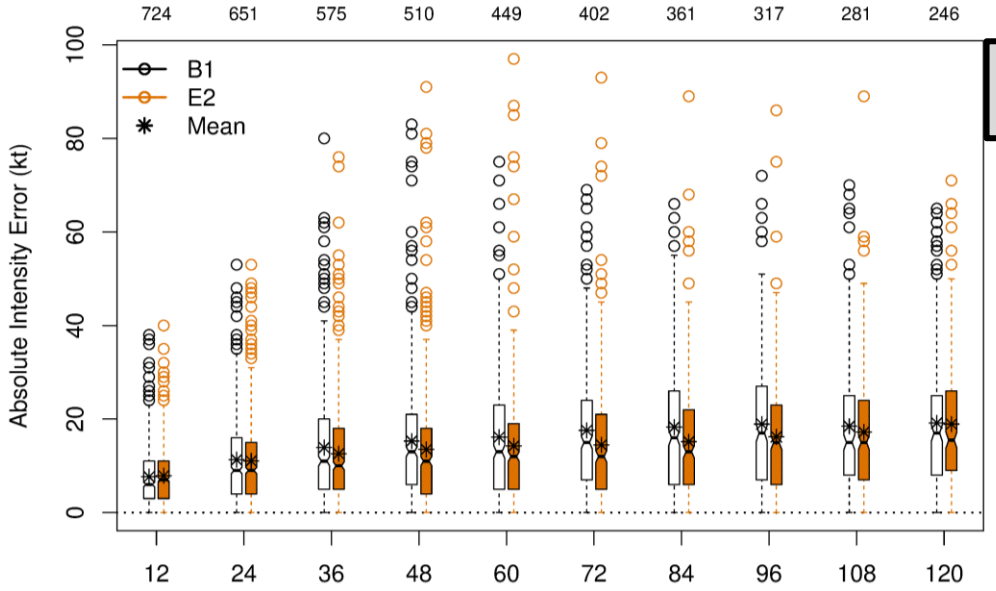
▲	Model 1 is better than Model 2 at the 99.9% significance level
▲	Model 1 is better than Model 2 at the 99% significance level
	Model 1 is better than Model 2 at the 95% significance level
	No statistically significant difference between Model 1 and Model 2
	Model 1 is worse than Model 2 at the 95% significance level
▼	Model 1 is worse than Model 2 at the 99% significance level
▼	Model 1 is worse than Model 2 at the 99.9% significance level
	Not statistically relevant
	Statistic for symbols: DIFF_SIG

Scorecard - METviewer batch engine

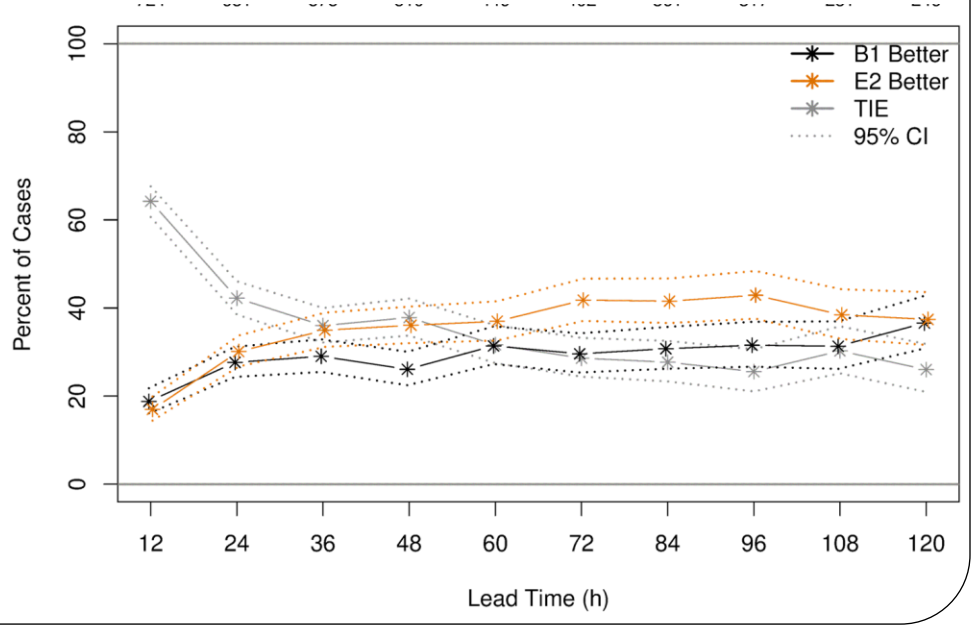
METplus TC Use Case



TC Use Case Example



Intensity Error Box Plots
Plot_TCMPR, R-statistics script

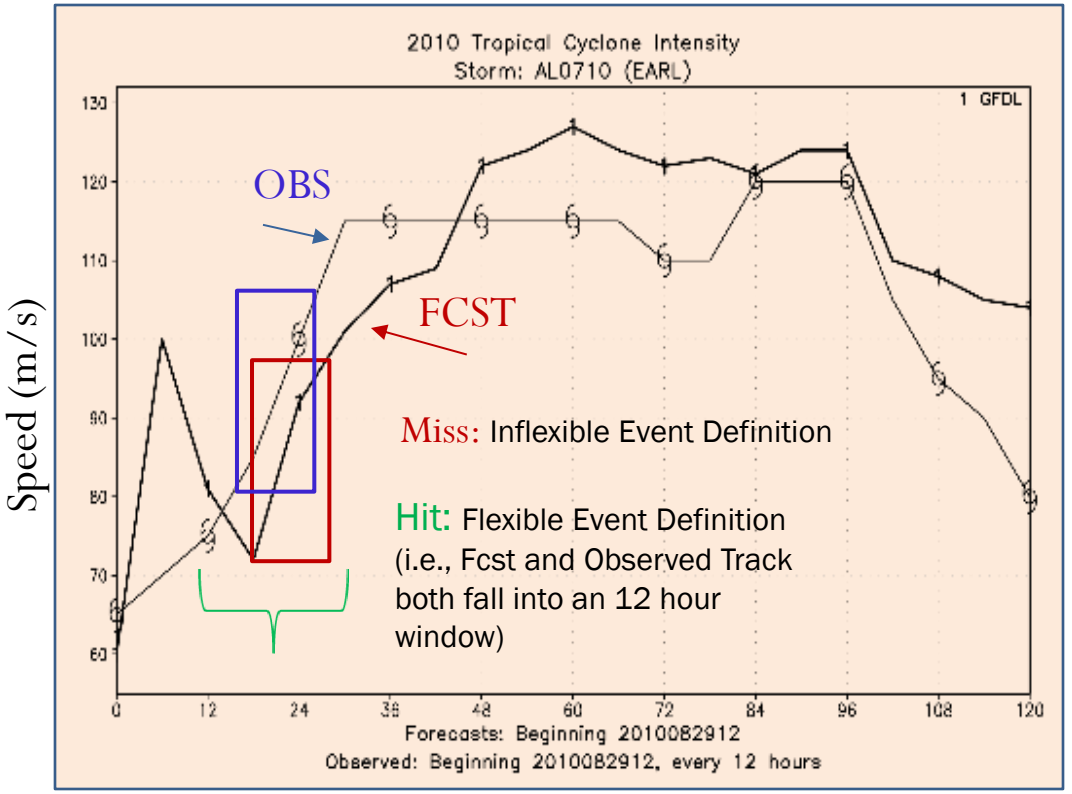


Frequency of Superior Performance
Plot_TCMPR, R-statistics script

Rapid Intensification (RI) Events

- Original support for tropical cyclone verification then expanded for other applications

Events identified by automated algorithm or user defined criteria of DI / DT

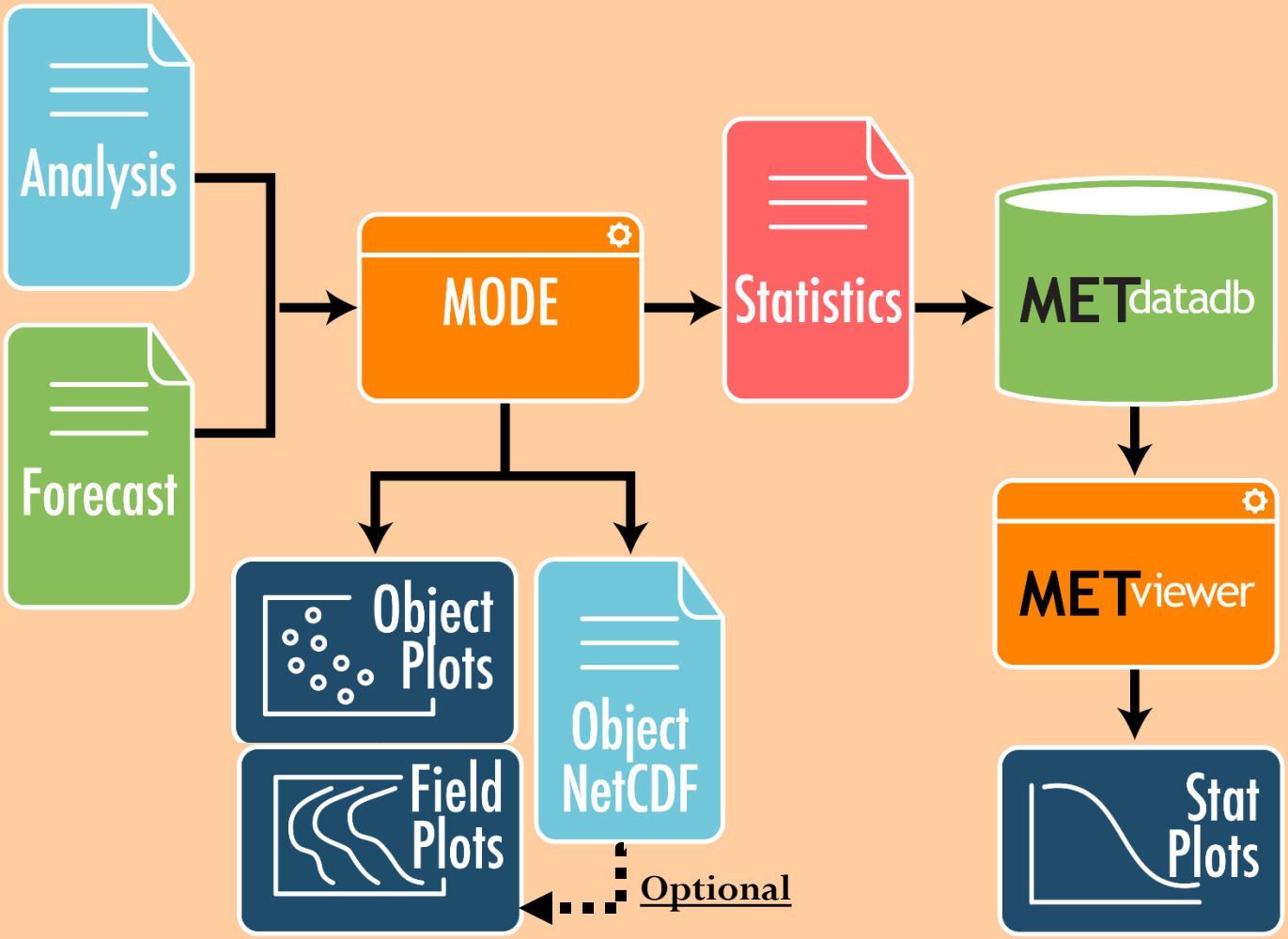


Also applicable to other intensity change event

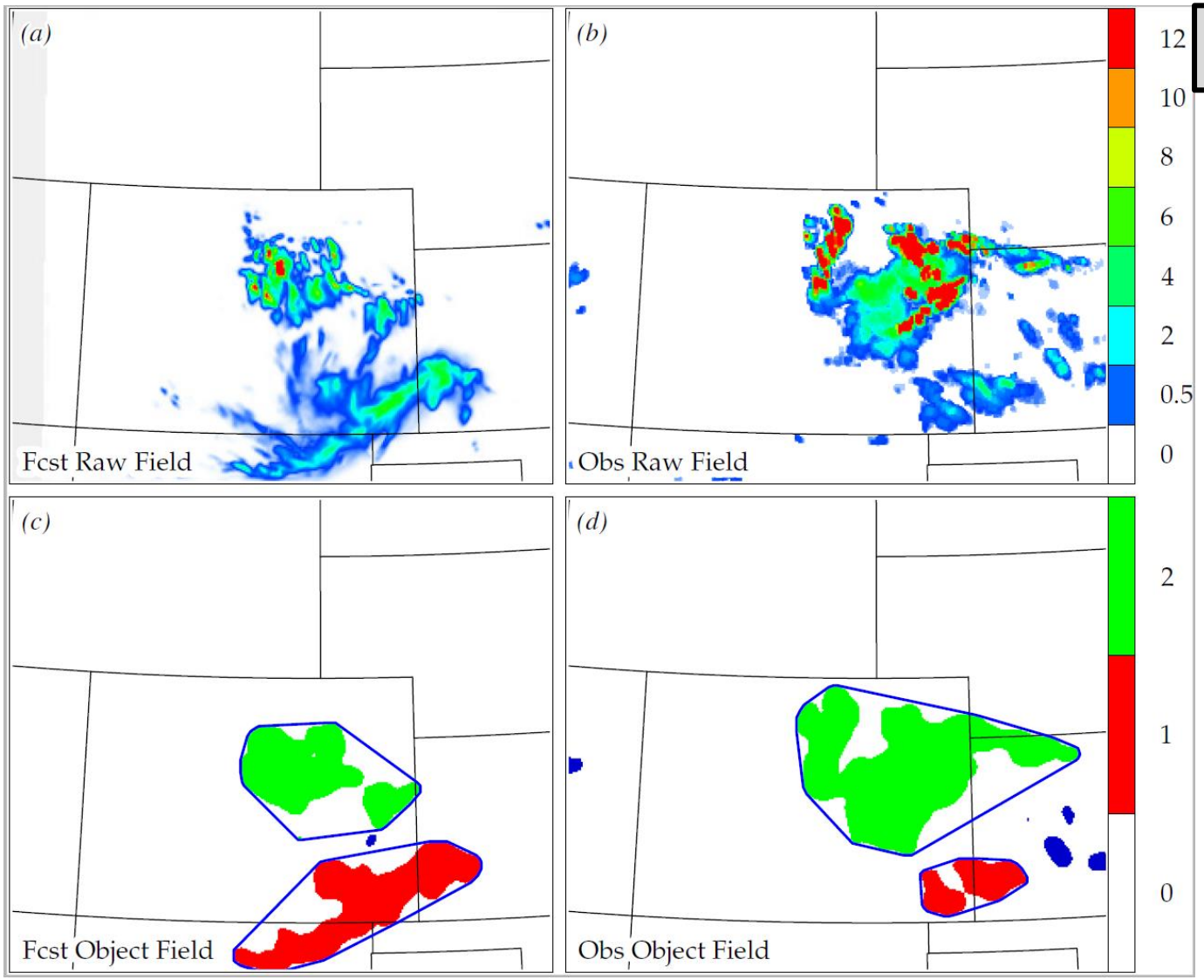
- ***Renewable Energy Ramps***
- ***Solar Wind during Coronial Mass Ejections***
- ***Extreme Temperature Changes***

Categorical statistics for RI or Rapid Weakening events (RW) can be computed

METplus MODE Use Case

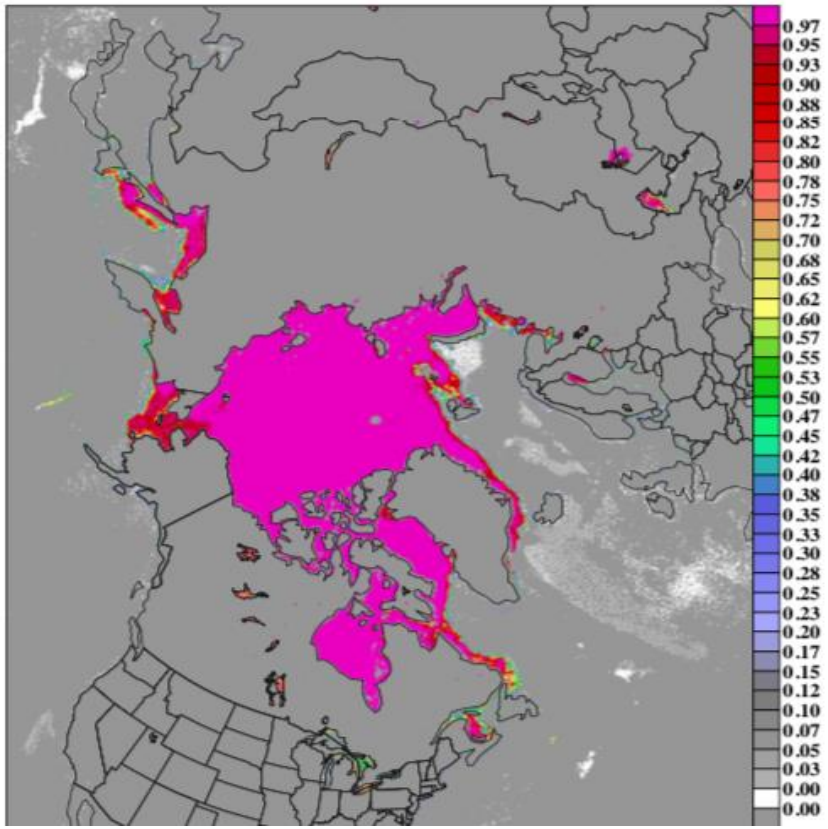


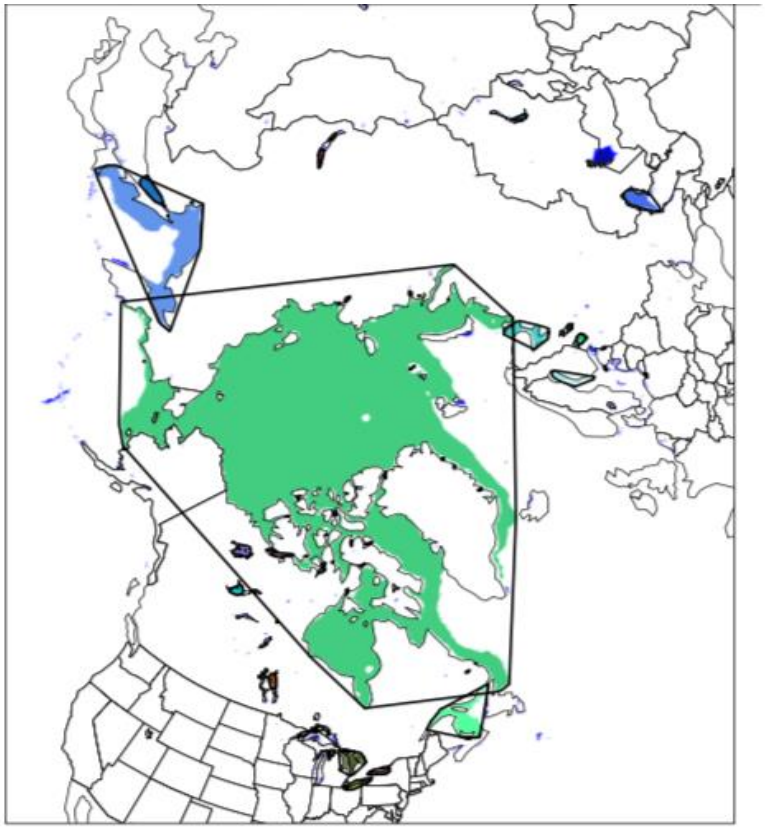
MODE Use Case Example - Precipitation



MODE postscript output

New Research – MODE for Sea Ice Fraction

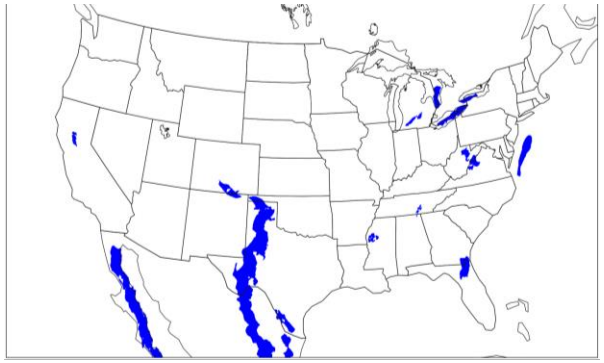




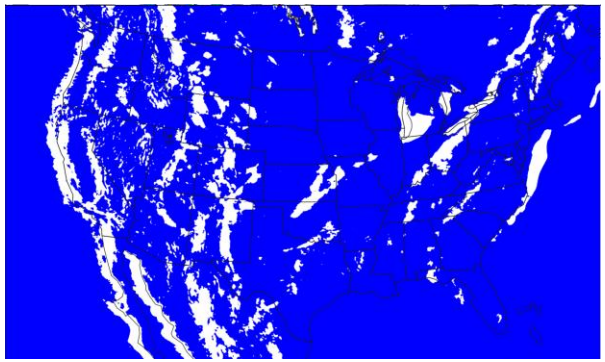
MODE postscript output

New Research – Multivariate MODE

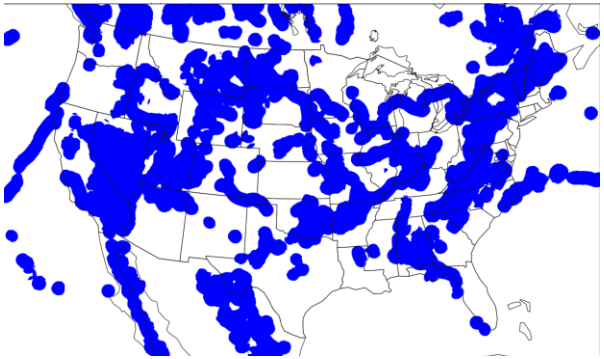
Specific Humidity Gradient

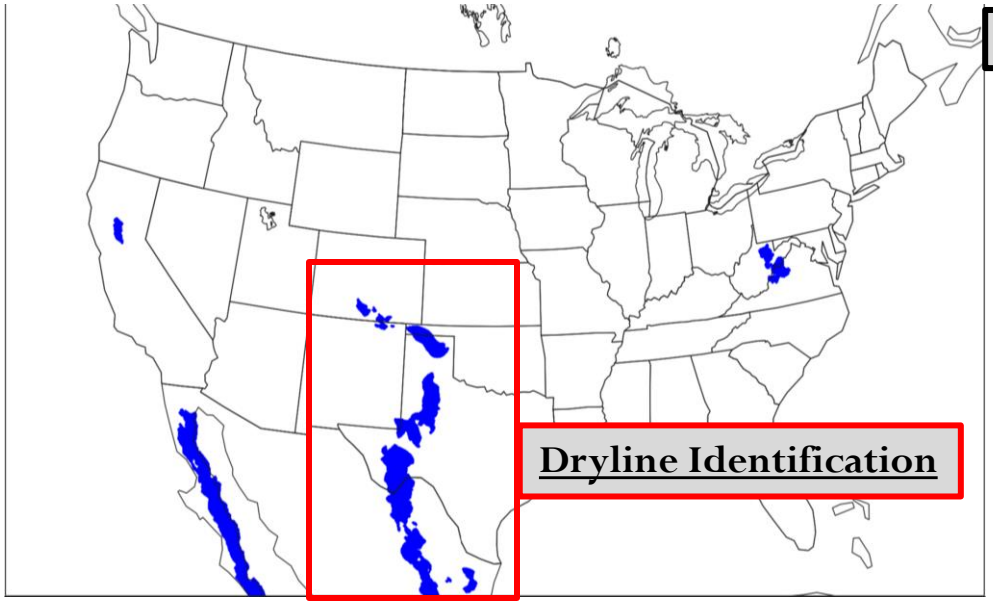


Temperature Gradient



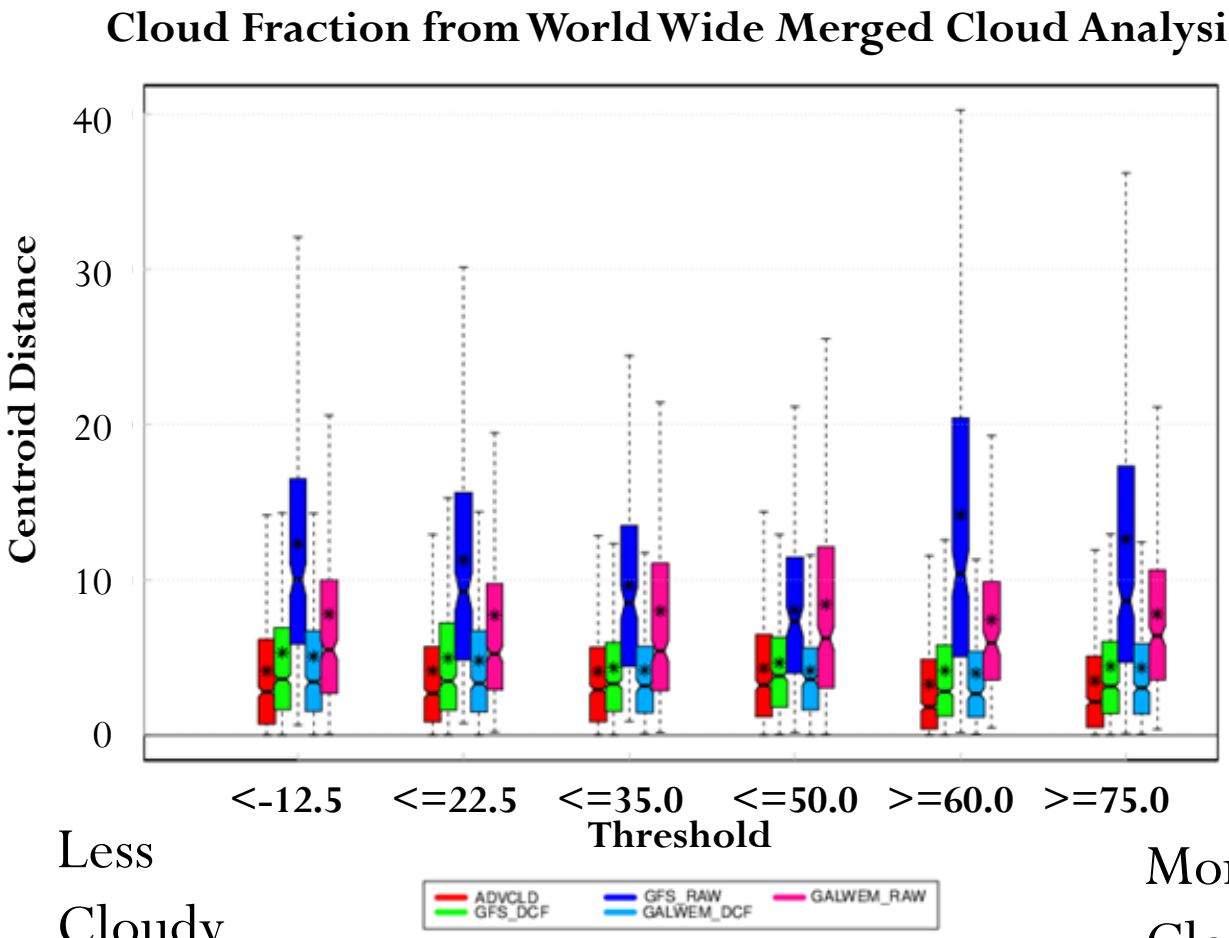
10-m wind direction shift





Contributed by Lindsay Blank

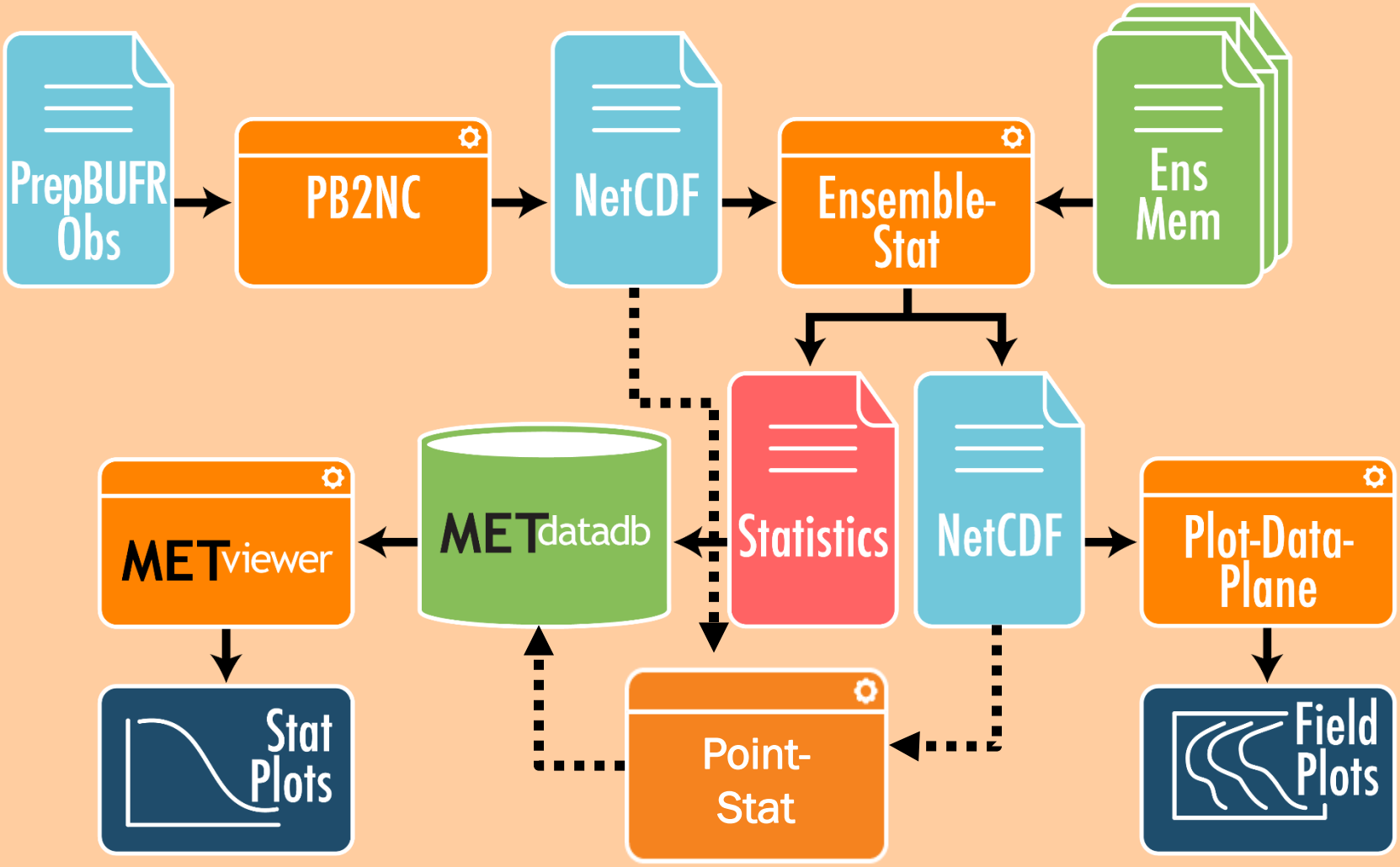
MODE Attributes Summarized



Less Cloudy

More Cloudy

METplus Ensemble Use Case



MET Ensemble and Probability Evaluation

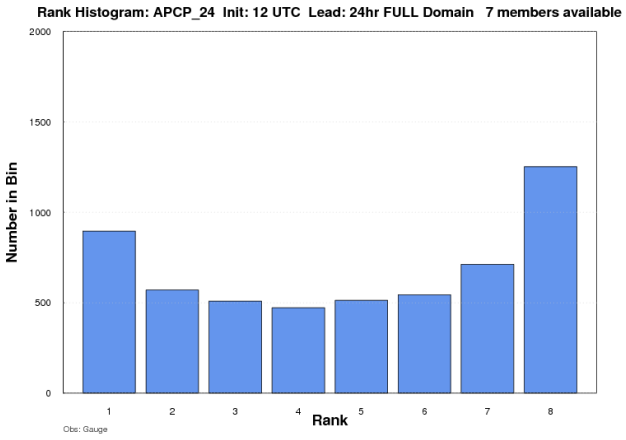
Ensemble Measures (Ensemble-Stat)

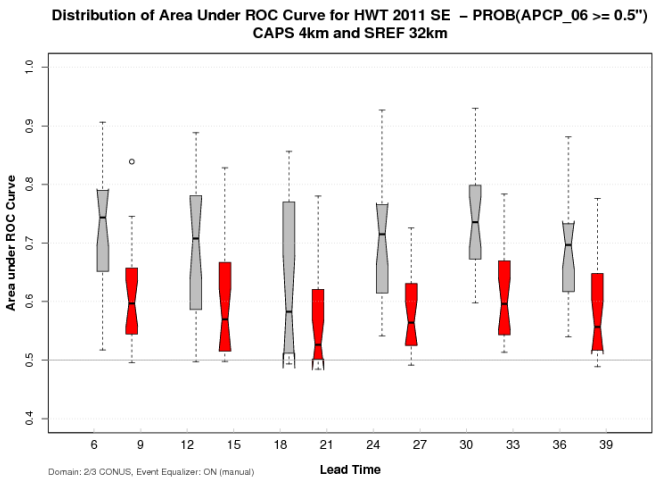
- CRPS, CRPSS, RPS, RPSS
- Ignorance Score
- Spread-Skill
- Rank Histogram, PIT, RELP

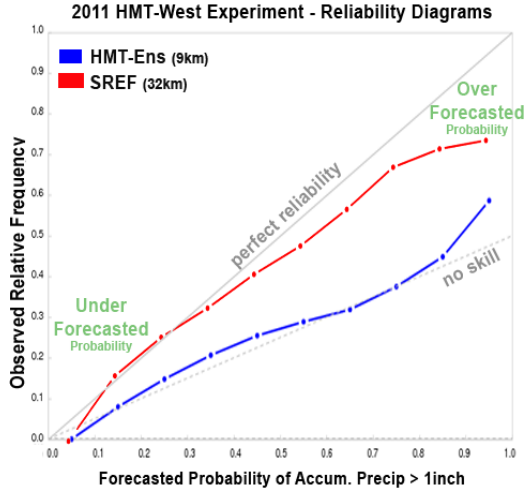
Probability Measures (Grid and Point-Stat)

- Brier Score + Decomposition, Brier Skill Score
- ROC and Area Under ROC
- Reliability

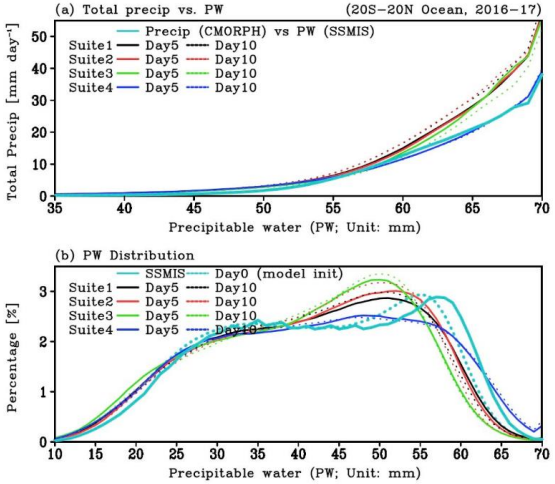
All METviewer







Philosophy: One Tool - Many Applications – GridDiag Tool



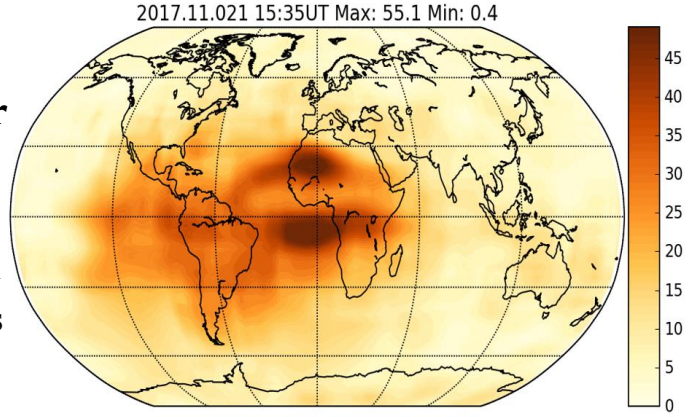
S2S

Correlate distributions of two fields

- Inventories the data
- Develops the PDF (one or two-dimensional)
- User configurable bins for PDF and percentiles
- Writes out netCDF files with bins or percentiles for use by other tools (Grid-Stat, Point-Stat, MODE, MTD)

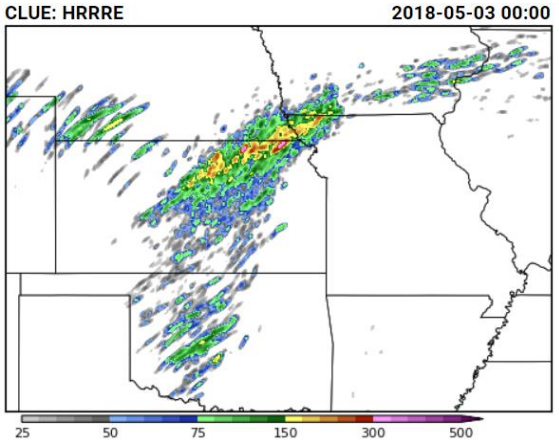
Space Weather

Adapt measures to regional max values



CAM

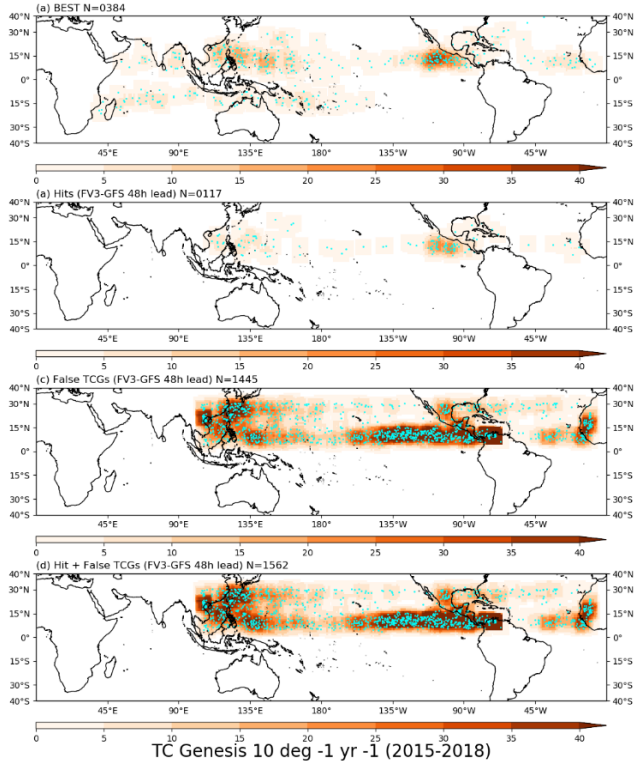
Normalize fields based on their sample climatology



Updraft Helicity

One Tools Many Apps: TC-Genesis

- *Collaboration with Dan Halperin, Embry-Riddle Aeronautical University*
- Compare forecast of TC-Genesis to actual BEST track and CARQ genesis events
- Writes contingency table counts and statistics; netCDF files of genesis events



Li et al., 2016

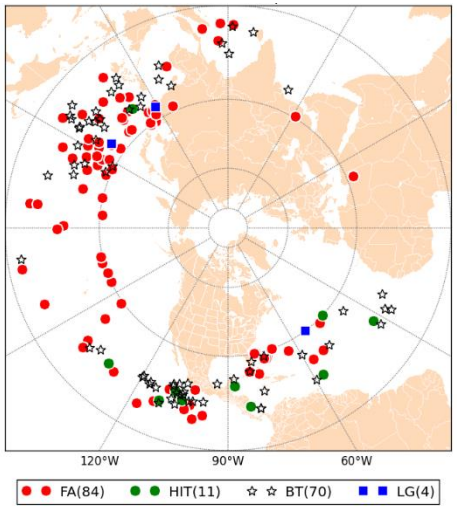
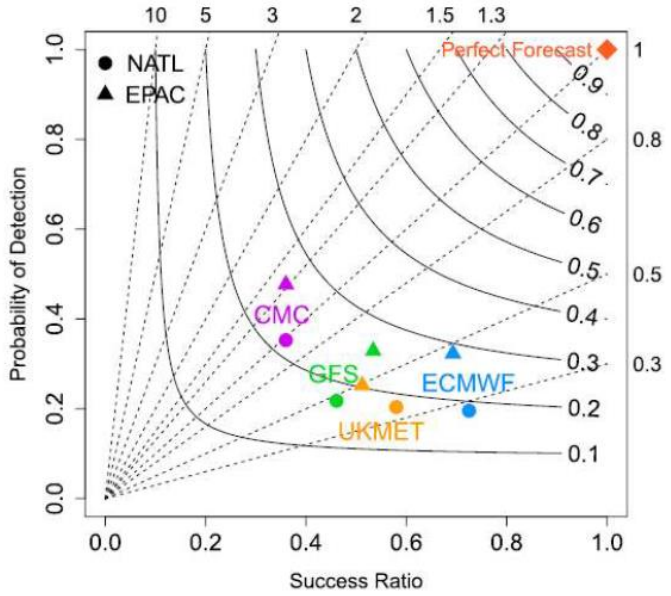


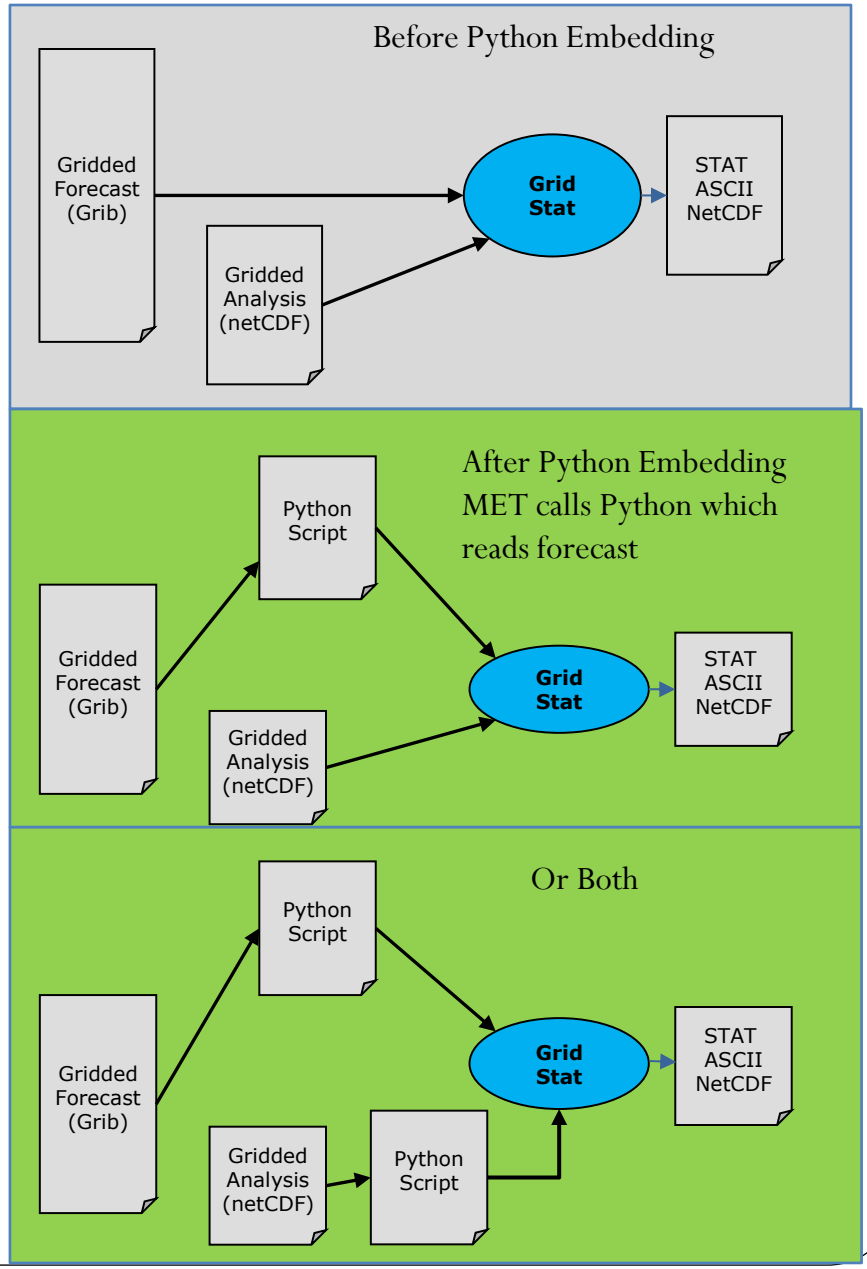
Figure: Tropical cyclogenesis verification for the NH for 2016. Symbols represent the Best Track (black), hits (green), late Genesis (blue) and false alarms (red).

Halperin et al., 2017



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
- Support for:
  - Python 3.6.3/3.7
  - Numpy, Xarray, Pandas arrays
  - Gridded fields through all tools
  - Point Observations through ASCII2NC and PB2NC tools



Use of Python Embedding - TDR for Evaluating TCs

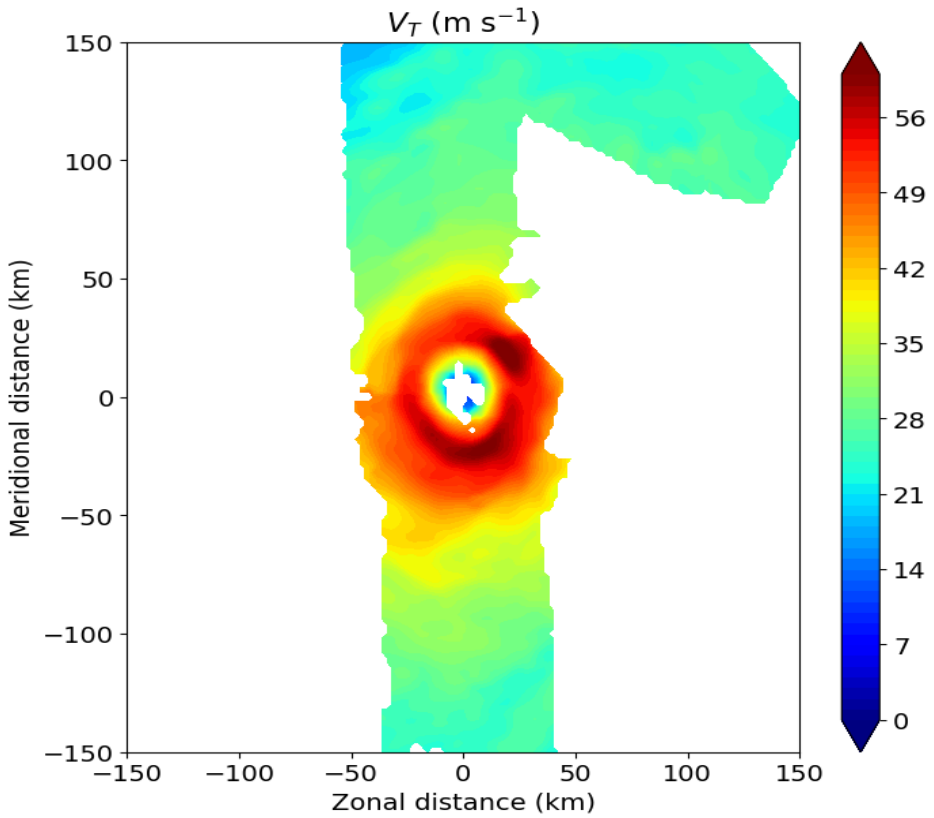


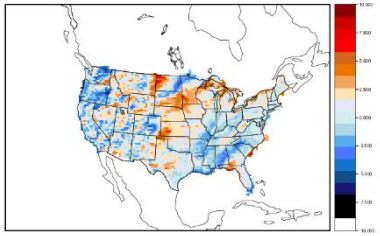
Image courtesy of Michael S. Fischer, Robert F. Rogers, Paul D. Reasor at NOAA/AOML/HRD

Typically, a given mission will have 3–4 passes through the center of the storm. Due to the X-band nature of the 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

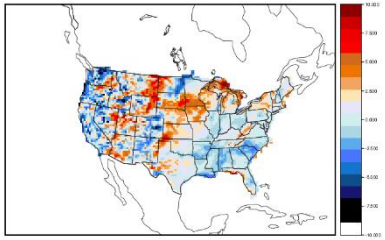
Distance Maps

- 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



Plot-Data-Plane Output

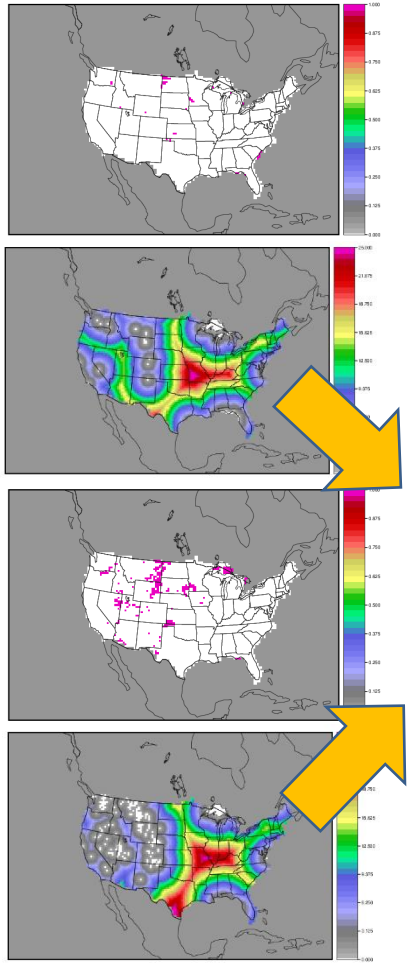


Table 8.7: Format information for DMAP (Distance Map) output line type.

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.

Next 2+ Years of METplus Development

NOAA Unified Forecast System

Coupled System Metrics:
- Metrics Workshop
- Sea-ice output evaluation
- PBL diagnostics
- Land-Surface diagnostics
- Hurricane track diagnostics
- Ensemble evaluation

Process-oriented diagnostic:
- Cold surface temperature, SST, and PBL biases
- Gravity wave drag and upward transport of adiabatic heating
- Sources of cloud cover and precipitation biases
- MJO, ENSO and other teleconnections

Additional Projects:

NOAA/OAR: S2S Weeks 2-4

NOAA/OAR ENSO diagnostics

NOAA/NWS: S2S Physics diagnostics

NOAA/NWS: Marine/Cryosphere coupled metrics

NOAA/NWS: Stratosphere diagnostics

NOAA/NWS: S2S Tropical diagnostics

NOAA/NESDIS: Space Weather system

Air Force: Cloud Verification / LSM / New Obs Data Types

NRL: Cloud Verification / DA diagnostics / Ensemble methods

DOE: LSM diagnostics

Met Office: Generalization of File Format Support / Optimization / Ensemble methods /SEEPS / Vx on Native Grid

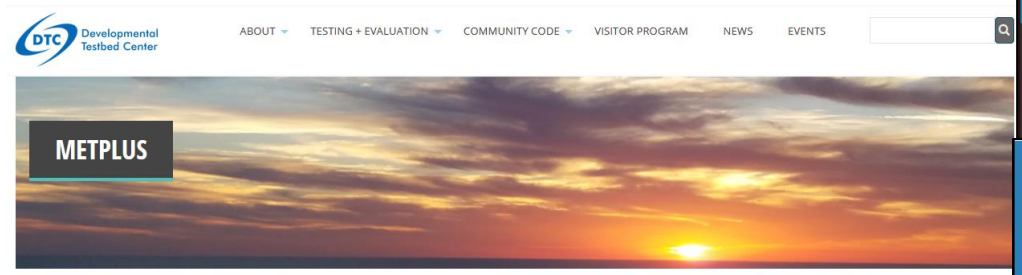
NCAR: Vx on Native Grid / G_β (E. Gilleland's talk 17 Nov 00UTC)

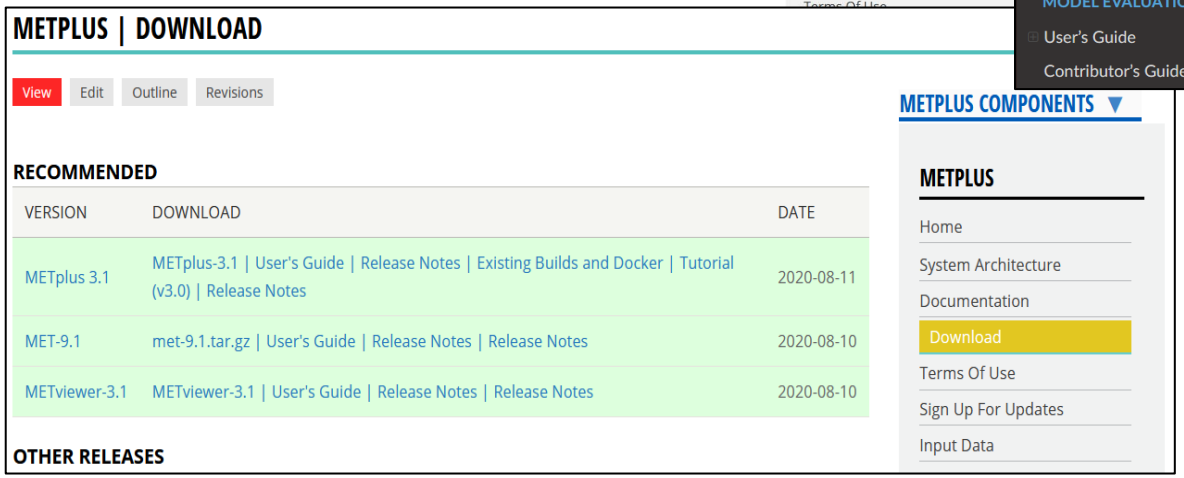
International Collaboration

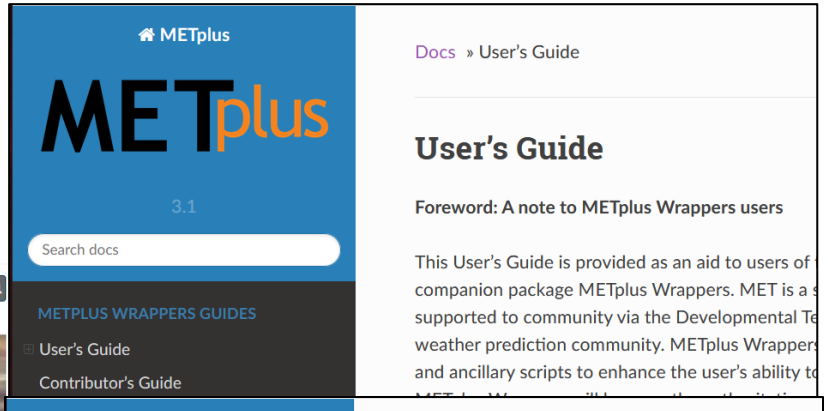
- Weekly to Monthly meetings between core contributors: NCAR, NOAA, Met Office, NRL – starting discussions on governance next month
- Growing online training video library
- METplus contributors guide
- Open repositories
  - Pull Requests
  - Continuous Integration
  - Cybersecurity Screening
  - Sample Data

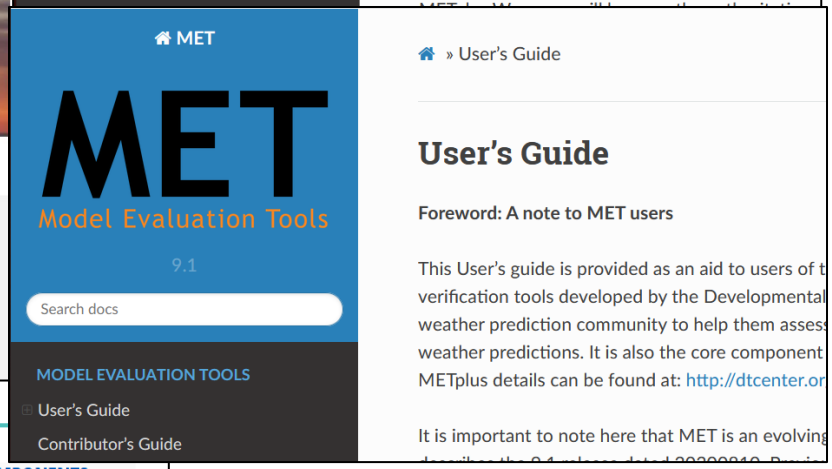
METplus:	https://github.com/dtcenter/METplus
MET:	https://github.com/dtcenter/MET
METviewer :	https://github.com/dtcenter/METviewer
METexpress:	https://github.com/dtcenter/METviewer
METcalcpy:	https://github.com/dtcenter/METcalcpy
METplotpy:	https://github.com/dtcenter/METplotpy

User Support









- Preparing to transition to from Help-Desk to Forums

- https://dtcenter.org/community-code/metplus

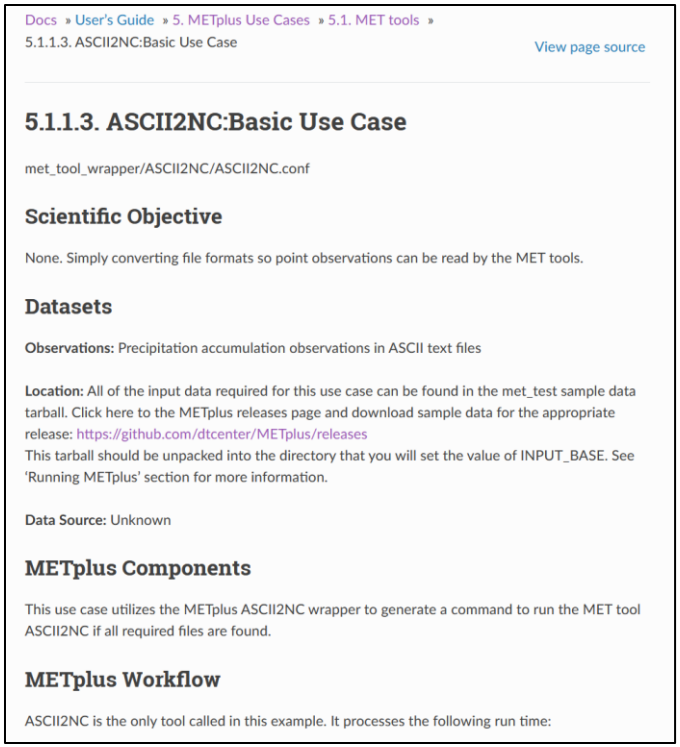
Training – Use Cases

https://dtcenter.github.io/METplus/Users_Guide/quicksearch.html

7. METplus Quick Search for Use Cases

7.1. Use Cases by MET Tool:

- ASCII2NC
- CyclonePlotter
- EnsembleStat
- GenVxMask
- GridStat
- GridDiag
- MODE
- MTD
- PB2NC
- PCPCombine
- Point2Grid
- PointStat
- RegridDataPlane
- SeriesAnalysis
- StatAnalysis
- TCMPRPlotter
- TCGen
- TCPairs
- TCRMW
- TCStat



Docs » User's Guide » 5. METplus Use Cases » 5.1. MET tools »
5.1.1.3. ASCII2NC:Basic Use Case

View page source

5.1.1.3. ASCII2NC:Basic Use Case

met_tool_wrapper/ASCII2NC/ASCII2NC.conf

Scientific Objective

None. Simply converting file formats so point observations can be read by the MET tools.

Datasets

Observations: Precipitation accumulation observations in ASCII text files

Location: All of the input data required for this use case can be found in the met_test sample data tarball. Click here to the METplus releases page and download sample data for the appropriate release: https://github.com/dtcenter/METplus/releases
This tarball should be unpacked into the directory that you will set the value of INPUT_BASE. See 'Running METplus' section for more information.

Data Source: Unknown

METplus Components

This use case utilizes the METplus ASCII2NC wrapper to generate a command to run the MET tool ASCII2NC if all required files are found.

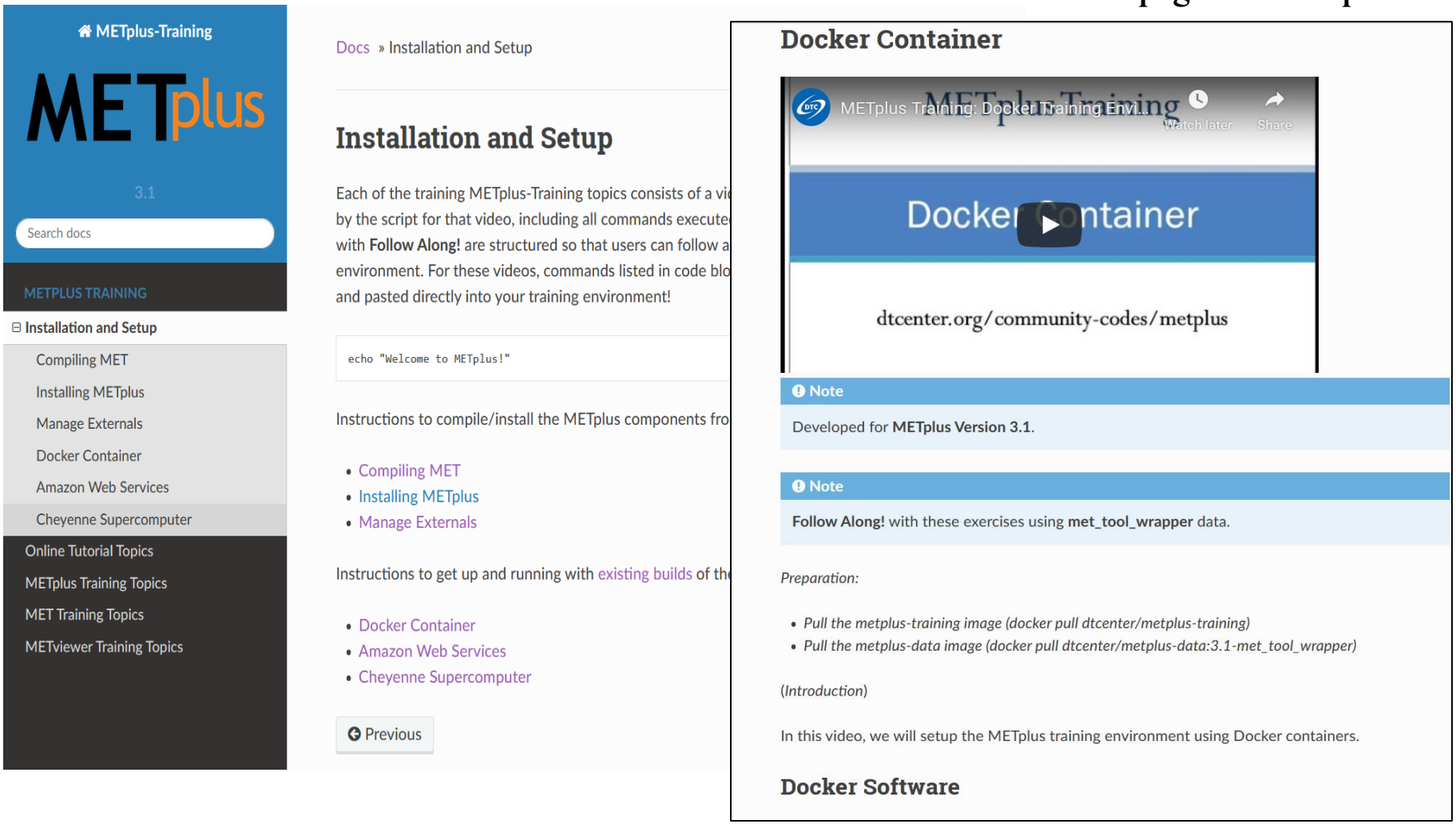
METplus Workflow

ASCII2NC is the only tool called in this example. It processes the following run time:

Training - Video

https://dtcenter.github.io/METplus-Training/index.html

Video embedded in page with script



Thank You for Your Attention

- Tara Jensen, NCAR, jensen@ucar.edu
- https://dtcenter.org/community-code/metplus

Much of this presentation taken from: Brown, B., T. Jensen, and Co-authors, 2020: The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bulletin of the American Meteorological Society*, in press, DOI 10.1175/BAMS-D-19-0093.1

2021 DTC UFS EVALUATION METRICS WORKSHOP

FEB 22 - 24 2021 **https://dtcenter.org/events/2021/2021-dtc-ufs-evaluation-metrics-workshop**

DTC Visitor Program: https://dtcenter.org/visitor-program

NSF NCAR DTC Developmental Testbed Center NOAA