The Model Evaluation Tools: Advanced Tools for Forecast Evaluation

Tara Jensen, John Halley Gotway, Randy Bullock, Tressa Fowler and Barb Brown
NCAR/Research Applications Laboratory, Boulder, CO
Developmental Testbed Center, Boulder, CO

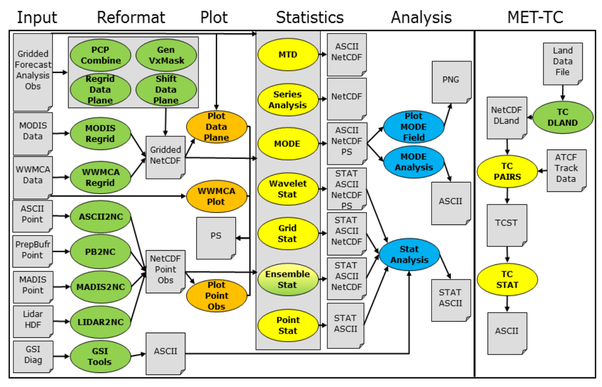
Developmental Testbed Center

http://www.dtcenter.org/met/users met_help@ucar.edu

The Model Evaluation Tools (MET) is a comprehensive numerical weather prediction (NWP) verification package supported to the community by the Developmental Testbed Center (DTC). It provides traditional verification statistics (e.g. RMSE, bias, skill scores), advanced spatial verification methods, and methods for ensemble and probabilistic forecasts. MET also includes pre-processing and aggregation tools, interpolation methods, and confidence intervals. **MET-TC** verifies tropical cyclone track, intensity etc. using ATCF A- and B-decks. MET-TC creates consensus forecasts and identifies rapid intensification and rapid weakening.

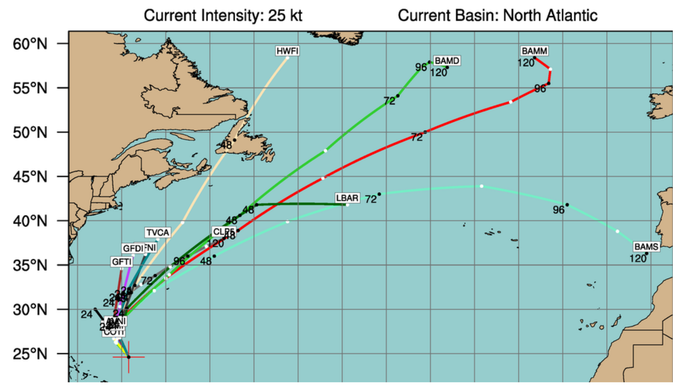
MET: A verification toolkit designed for flexible yet systematic evaluation

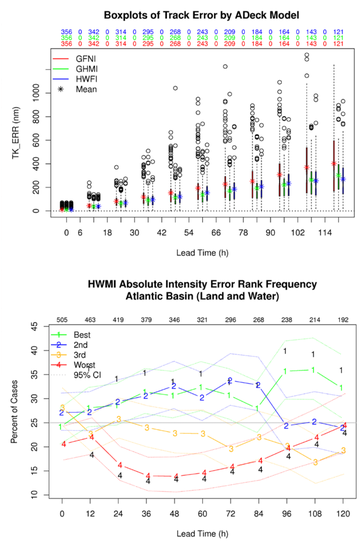
- Over 50 traditional statistics using both point and gridded datasets
- Multiple interpolation methods
- Computation of confidence intervals
- Object based, neighborhood and scale decomposition methods
- Able to read in GRIB1, GRIB2 and CF-compliant NetCDF
- Applied to many spatial (1km – global) and temporal scales (minutes to decades)



Preprocessing Data Inspection — Statistics — Analysis

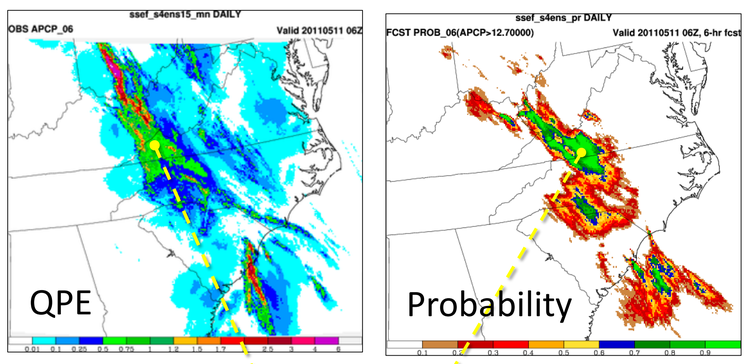
MET-TC for Tropical Cyclones



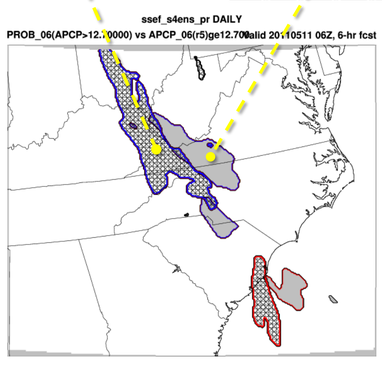


- Based on NHC system – reads ATCF format
- Produces matched pairs between forecast and best track
- Allows for filtering by storm etc.
- Identifies Rapid Intensification/Weakening events

MODE: Method for Object-based Diagnostics



- Identifies Features
- Matches them between fields
- Calculates Attributes



Bad forecast or Good forecast with displacement error?

Output from all tools to support geographical representation of errors

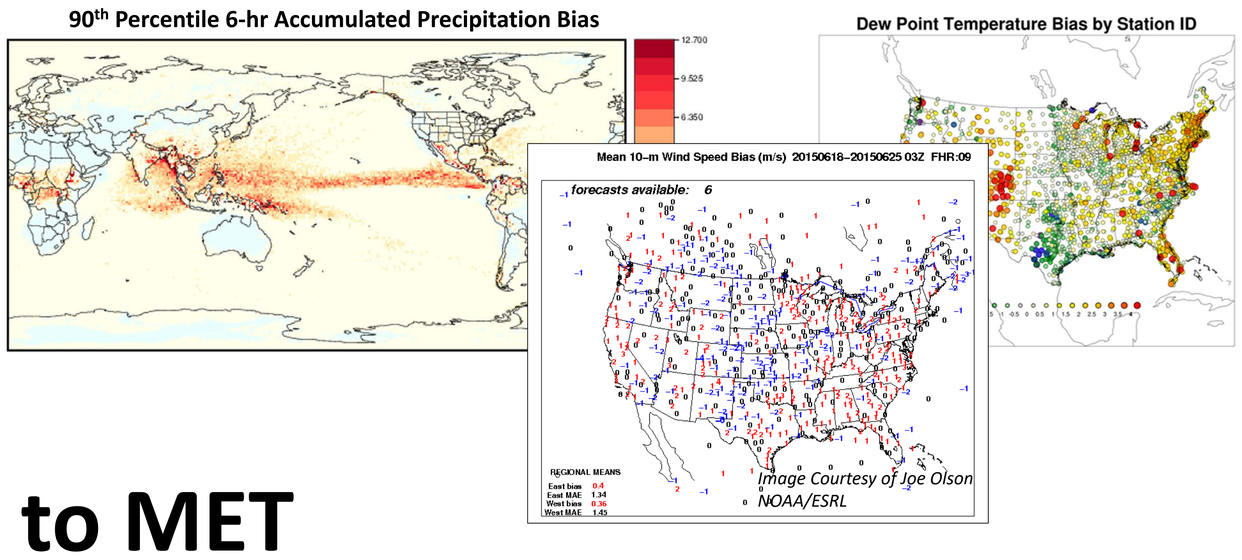


Image Courtesy of Joe Olson NOAA/ESRL

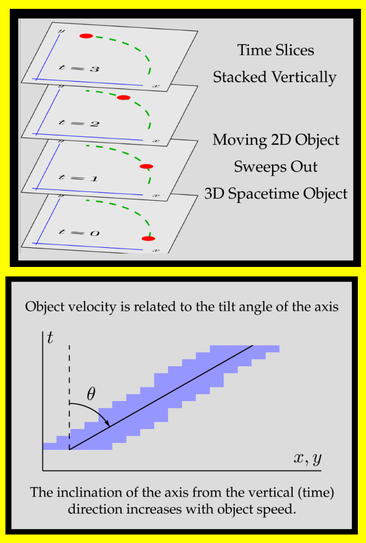
Most Recent Enhancements to MET

MODE-TD

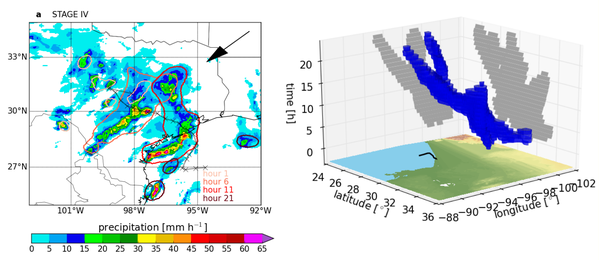
Method for Object Diagnostic Evaluation Time Domain

- ❑ Convolution (smoothing) and thresholding performed analogous to traditional 2D MODE then the raw values are re-established within the object..
- ❑ 3D objects are then formed by connecting to objects immediately adjacent in space +/-1 time step.
- ❑ Forecasts must have high enough temporal resolution so 2D objects overlap
- ❑ This method has been applied to many spatial and temporal scales: Convection Allowing Models (CAMs) *(Clark et. al, 2014)* as well as Climate Models

Reference:
Prein AF, C Liu, K Ikeda, R Bullock, R Rasmussen, G Holland, M Clark (2016) Simulating Convective Storms: A New Benchmark for Climate Modeling, Bull. Am. Meteorol. Soc., submitted



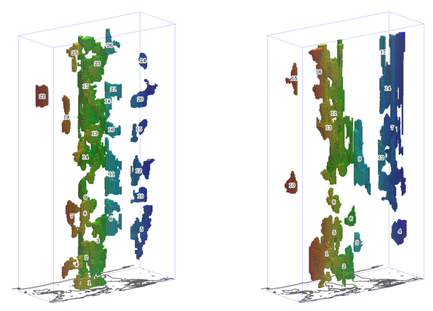
Precipitation from Regional Climate



Graphics Courtesy of Andreas Prein, NCAR

Precip Field with Time — MTD Object

Drought Index from Climate Model



Simulation — Analysis

Errors to be detected: Timing, Velocity, Duration, Build-up & Decay

Weighting

Cosine Latitude and Grid-Box Area Weighting

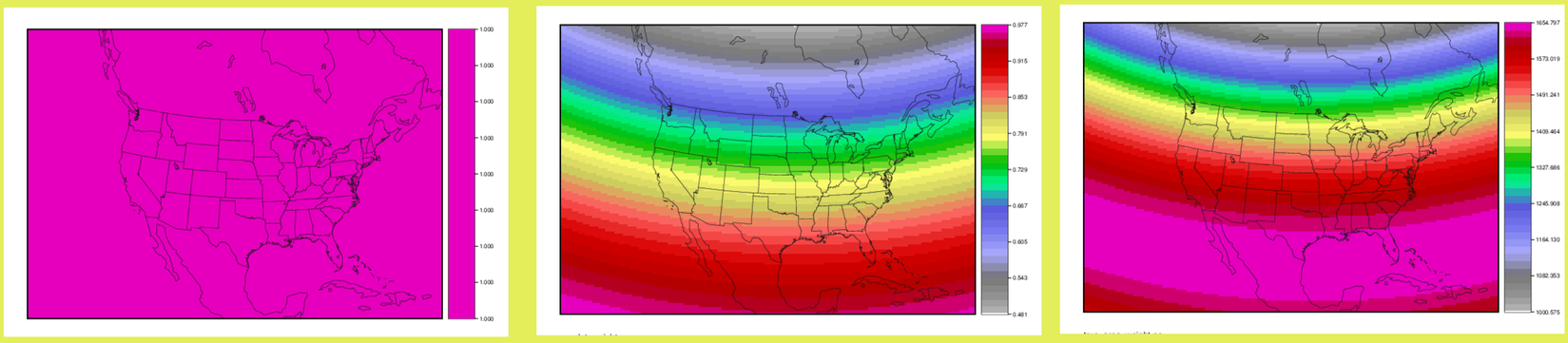
Cosine Latitude weighting
- ❑ Applied to Latitude-Longitude grids to account for meridional convergence at higher latitudes
- ❑ Assumes that the data to be weighted have a distribution that is uniform per degree of latitude.

Grid-box area weighting
- ❑ Another option that may be beneficial for randomly, or quasi-randomly distributed data

Reference:
Gleisner, H, (2011): Latitudinal Binning and Area-Weighted Averaging of Irregularly Distributed Radio Occultation Data. GRAS SAF Report 10.

WMO, 2010: Manual on the Global Data-processing and Forecast System. WMO-No. 485, 1.

No Weighting — Cosine Latitude Weighting — Grid-box Area Weighting



WMO Standard

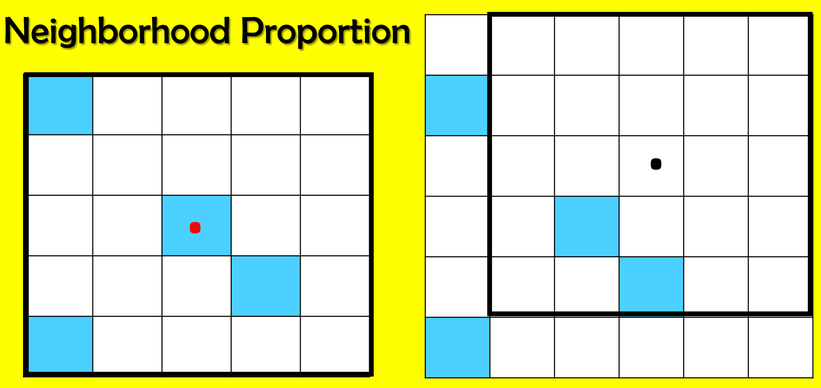
HiRA Framework

High Resolution Analysis Framework

- ❑ Applied to deterministic forecasts matched to point observations.
- ❑ Uses neighborhood fraction of events in place of probability forecast.
- ❑ Point observation location defines neighborhood.
- ❑ As with all neighborhood methods, allows for some spatial / temporal uncertainty in either model or observation 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.

Reference:
Mittermaier, M., 2014: A strategy for verifying near-convection-resolving model forecasts at observing sites. *Wea. Forecasting,* **29**, 185-204.

Neighborhood Proportion



Nbrhd size 1: 1/1
Nbrhd size 9: 2/9
Nbrhd size 25: 4/9

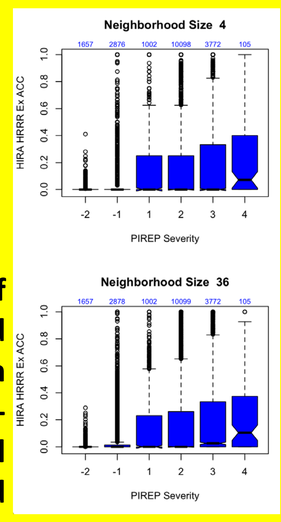
Nbrhd size 1: 0/1
Nbrhd size 9: 1/9
Nbrhd size 25: 2/9

Compute Brier Score using Neighborhood Proportion

$$\mathrm{BS} = \frac{1}{T}\sum_i (p_i - o_i)^2$$

- (> Thresh)
- (<=Thresh)
- O_i=1 (> Thresh)
- O_i= 0 (<= Thresh)

Example of Neighborhood Proportion For Multi-Categorical Field



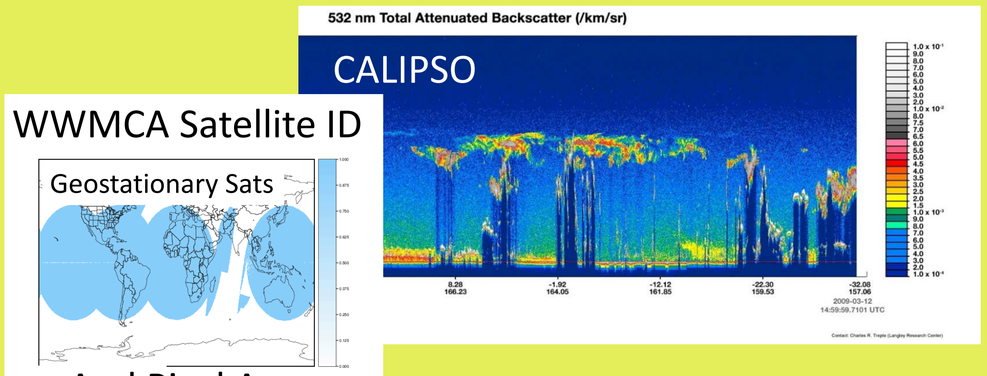
Other Additions

Other Notable Enhancements

- ❑ READ NETCDFv4
- ❑ READ CALIPSO Layers and compute Cloud Base, Height
- ❑ Addt'l Specialty Masking
- ❑ Support for ATCF E-Decks (probability forecasts)

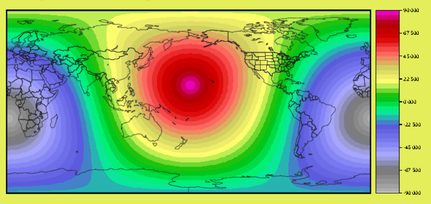
The Developmental Testbed Center is supported by National Oceanic and Atmospheric Organization (NOAA), National Center for Atmospheric Research (NCAR), U.S. Air Force (USAF) and National Science Foundation (NSF). NCAR is sponsored by NSF.

Support for New Observations



Specialty Masking

Day/Night



Lat/Lon File
- Polyline (poly)
- Box
- Circle
- Track

Gridded Data File
- Grid
- Data
- *Lat or Lon bands*

Or Timestamp
- Solar Altitude
- Solar Azimuth

Also Available: Union, Intersection, Symm. Diff