

DTC/HMT Verification Overview

Tara Jensen and Ed Tollerud

for

DTC/HMT Team

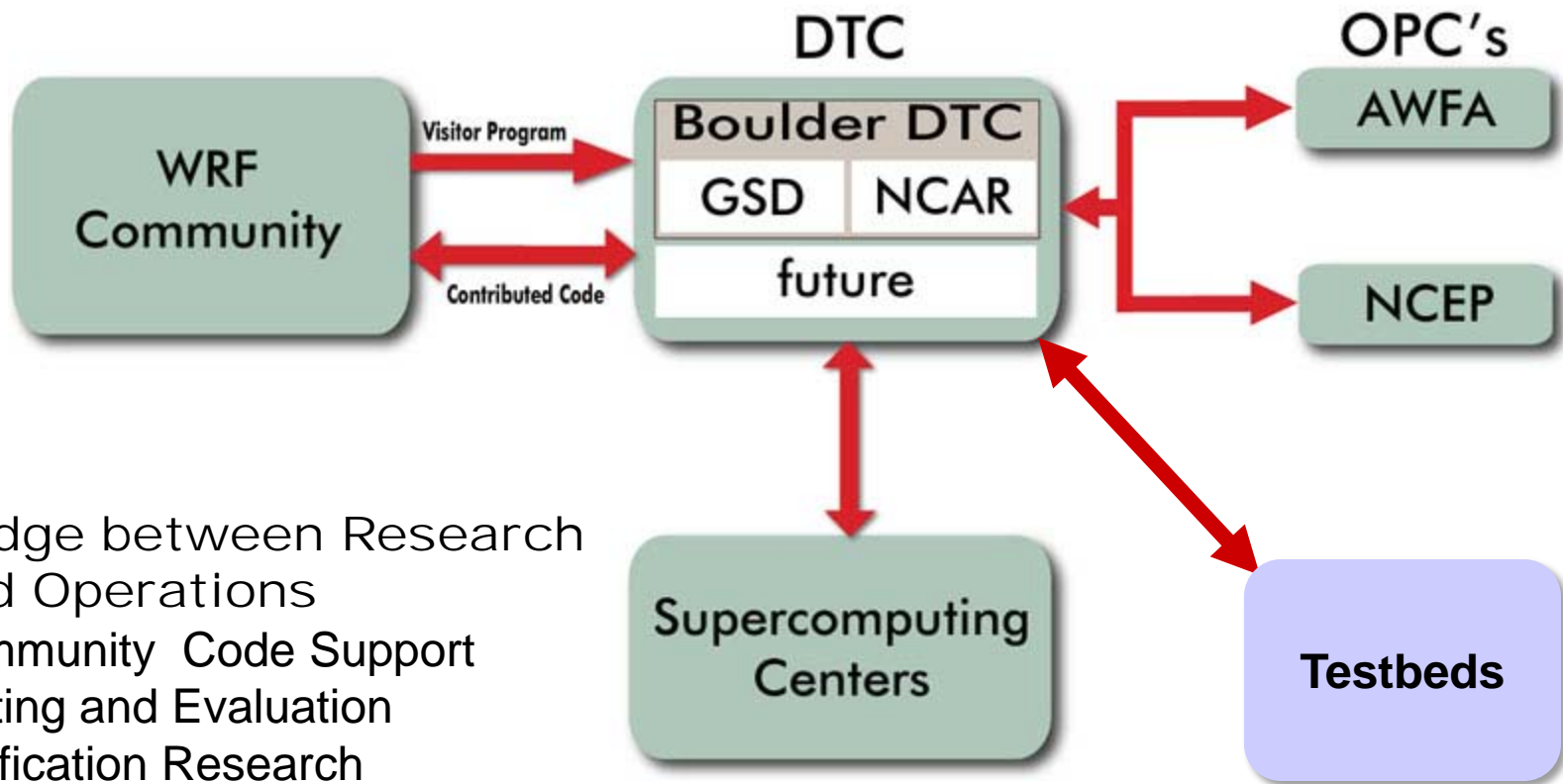
Telecon – 5 Feb 2010



The Development Testbed Center (DTC)

What do we do?

The DTC Architecture



Evaluation system built on DTC Model Evaluation Tool (MET)

MET is a set of tools for evaluating model forecasts.

- **Preprocessing**

- Point Obs
- Precip Accumulation
- Sub-domain Masking

- **Statistics**

- Traditional methods
- Spatial methods

- **Post-processing**

- Aggregation over time and regions
- METviewer database and display system

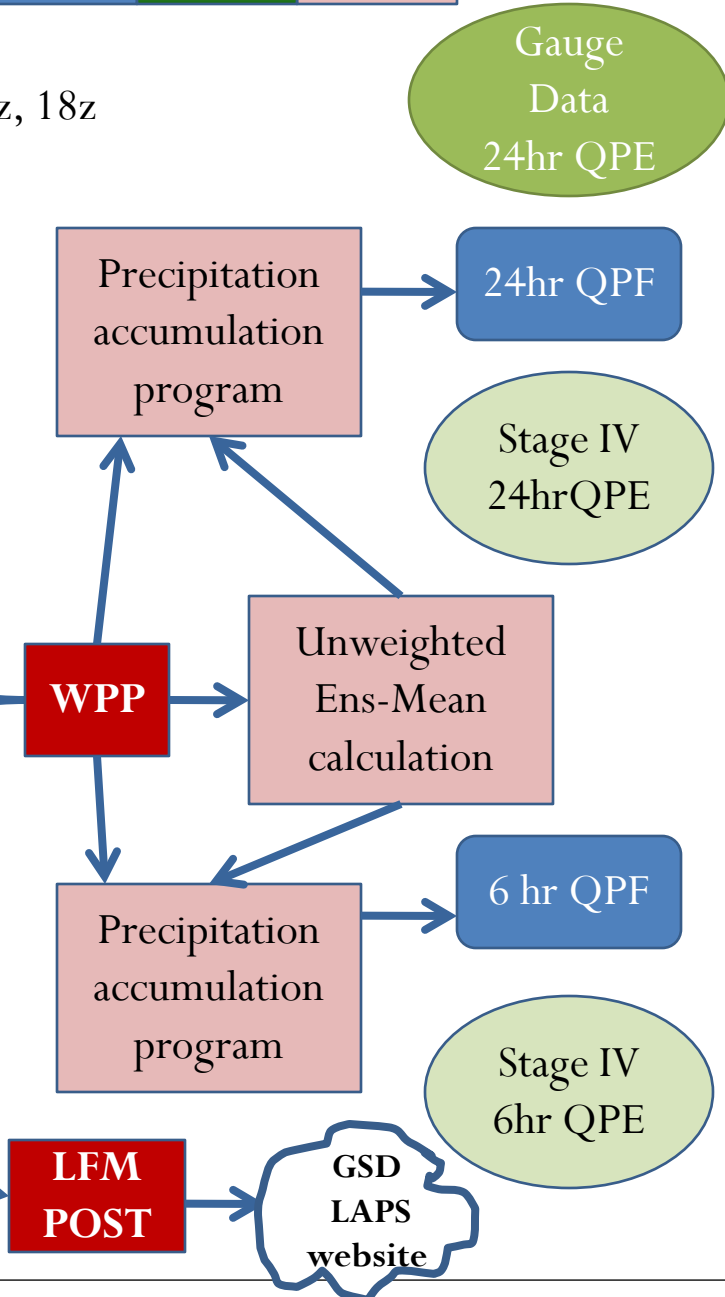


d01 – 9km
6 hr and 24hr
APCP

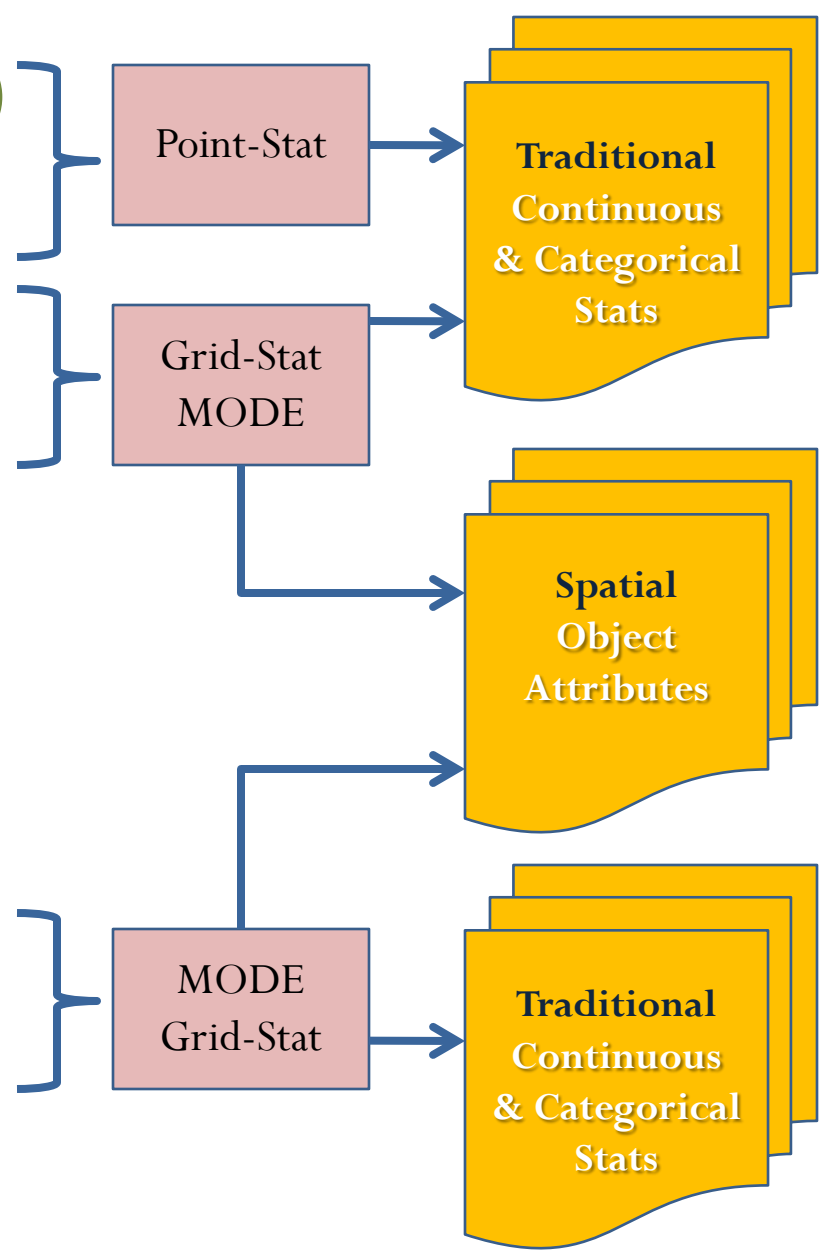
Init:00z, 06z, 12z, 18z

- arw-tom-gep1
- arw-fer-gepf
- arw-sch-gep2
- arw-tom-gep3
- nmm-fer-gep4
- arw-fer-gep5
- arw-sch-gep6
- arw-tom-gep7
- nmm-fer-gep8

**E
n
s
e
m
b
l
e**

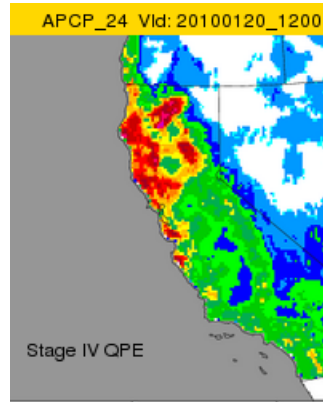
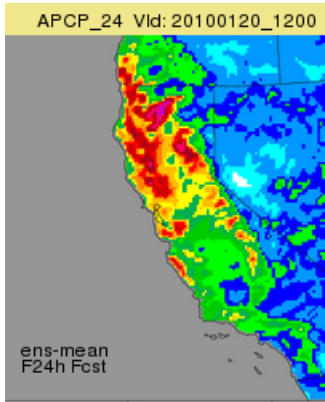


SYSTEM DIAGRAM



MET Traditional Measures (from Point-Stat and Grid-Stat)

- **Gridded and point verification**
 - Multiple interpolation and matching options



Matching approaches:

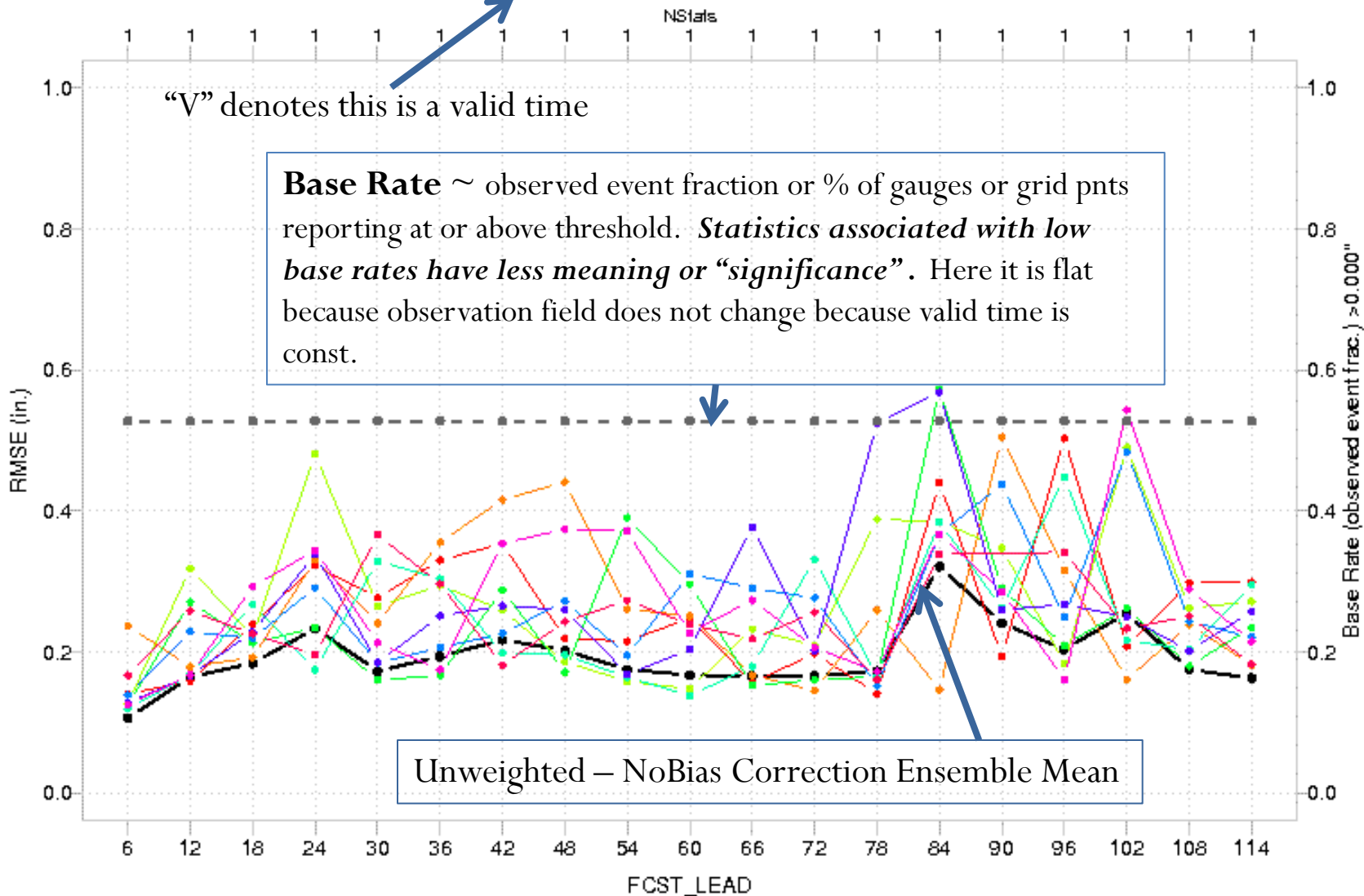
MET allows users to select the number of forecast grid points to match to a point observations and the statistic to use to summarize the forecasts.

- **Statistics**
 - **Continuous** – MAE, RMSE, ME, Correlation, BCRMSE, etc.
 - **Categorical** - POD, FAR, CSI (Threat), GSS (ETS), Freq Bias, etc.
 - **Probabilistic** - Brier Score, Reliability, ROC, Rank Histogram*, CRPS*
*in spring release

How does model performance
change over time??

VALID TIME: CONSTANT INIT TIME: CHANGES

Valid: 2010012112V APCP_06 RMSE vs. FCST_LEAD – Obs: Stage IV



“V” denotes this is a valid time

Base Rate ~ observed event fraction or % of gauges or grid pnts reporting at or above threshold. *Statistics associated with low base rates have less meaning or “significance”*. Here it is flat because observation field does not change because valid time is const.

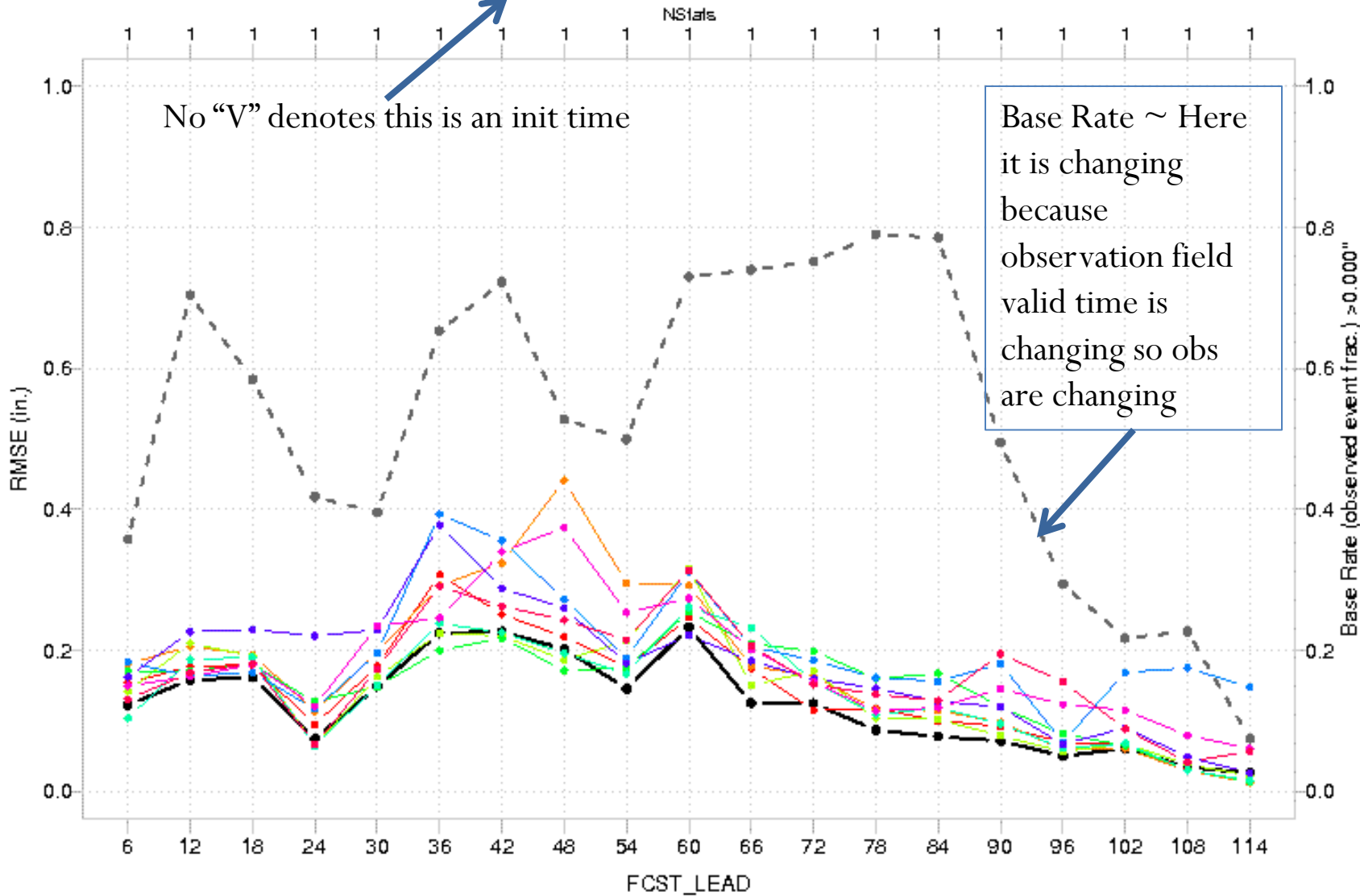
Unweighted – NoBias Correction Ensemble Mean

- ens-mean RMSE
- arw-10m-gcp0 RMSE
- arw-10m-gcp1 RMSE
- arw-sch-gcp2 RMSE
- arw-10m-gcp3 RMSE
- nmm-10m-gcp4 RMSE
- arw-10m-gcp5 RMSE
- arw-sch-gcp6 RMSE
- arw-10m-gcp7 RMSE
- nmm-10m-gcp8 RMSE
- ens-mean >0.000 BASER

How does each model perform on a run-by-run basis??

VALID TIME: CHANGES INIT TIME: CONSTANT

Init: 2010011912 ACP0_06 RMSE vs. FCST_LEAD - Obs: Stage IV



- ens-mean RMSE
- arw-lom-gcp0 RMSE
- arw-lor-gcp1 RMSE
- arw-sch-gcp2 RMSE
- arw-lom-gcp3 RMSE
- nmm-lor-gcp4 RMSE
- arw-lor-gcp5 RMSE
- arw-sch-gcp6 RMSE
- arw-lom-gcp7 RMSE
- nmm-lor-gcp8 RMSE
- ens-mean >0.000 BASER

MET Spatial Verification approaches

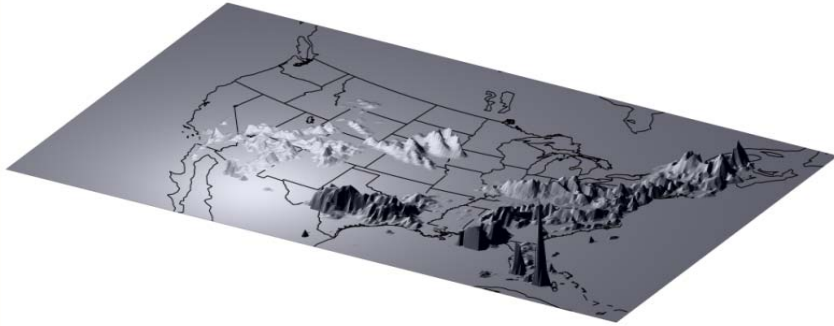
(using MODE)

- Higher Resolution forecasts of spatially-coherent fields (e.g., precipitation) are typically penalized using traditional statistics
- Spatial techniques provide a potentially more meaningful evaluation
- Examples
 - *What* is wrong with the forecast?
 - At what scales does the forecast perform well?
 - How does the forecast perform on attributes of interest to users?

MODE – Spatial Verification

*Method For Object-
Based Diagnostic
Evaluation*

Steps #1-4: How objects are identified in MODE...

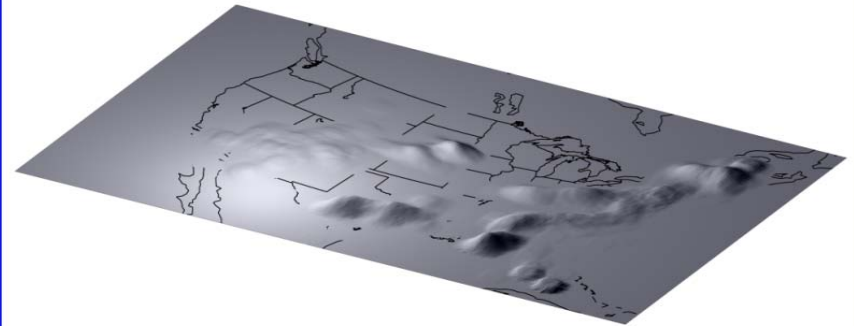


Step #1

Start with the raw data field.

In this case, a precipitation field.

© 2009 University Corporation for Atmospheric Research. All Rights Reserved.

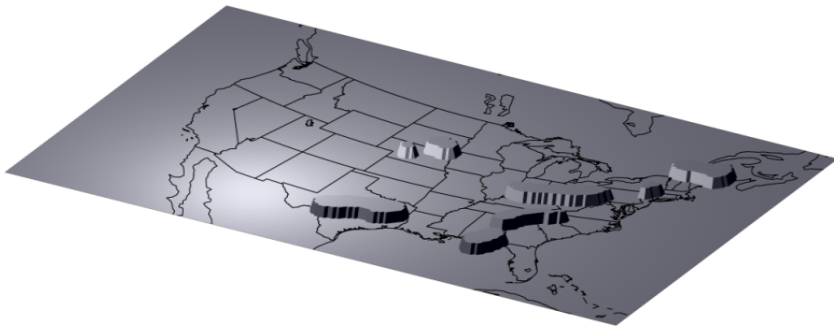


Step #2

Apply convolution operator.

This is basically a smoothing operation.

© 2009 University Corporation for Atmospheric Research. All Rights Reserved.



Step #3

Threshold the smoothed field.

This produces an on/off mask field.

© 2009 University Corporation for Atmospheric Research. All Rights Reserved.



Step #4

Restore original data to object interiors.

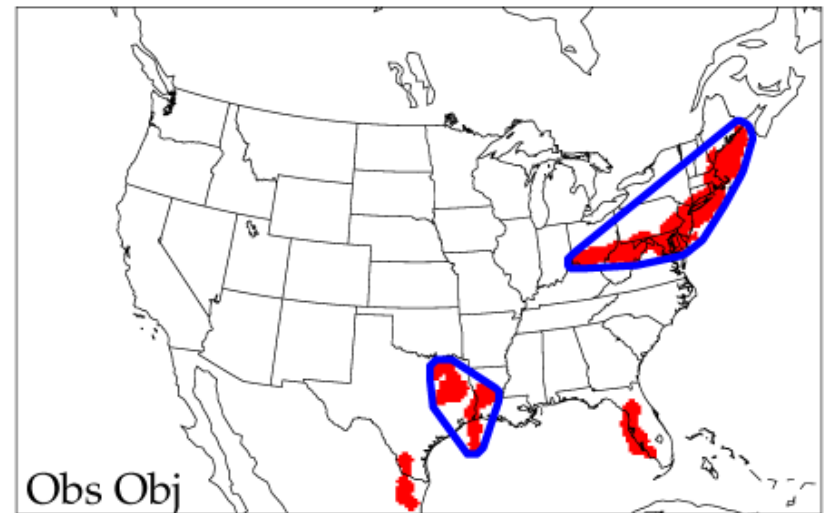
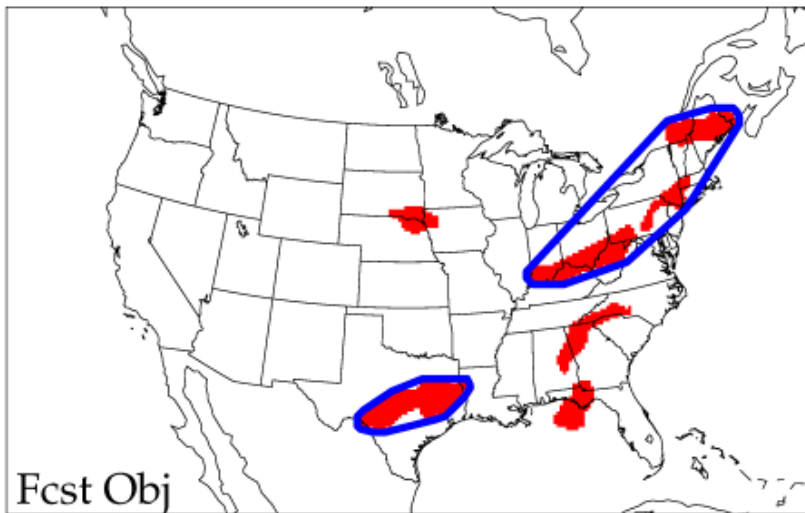
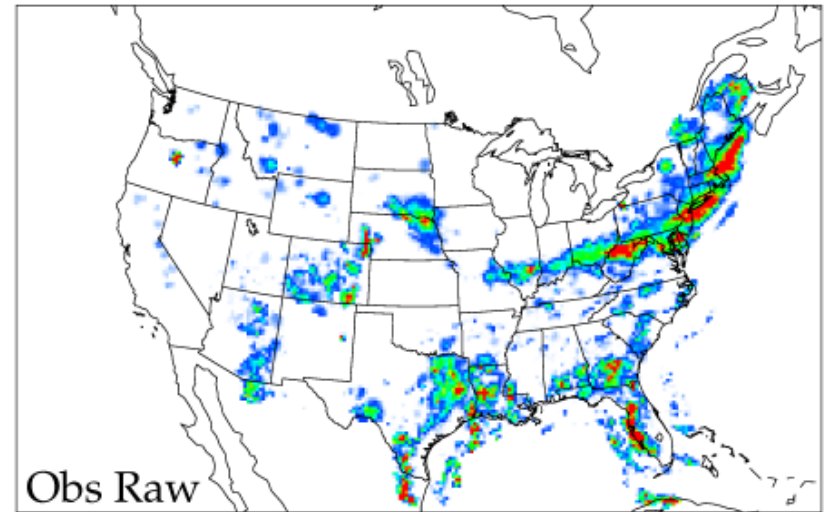
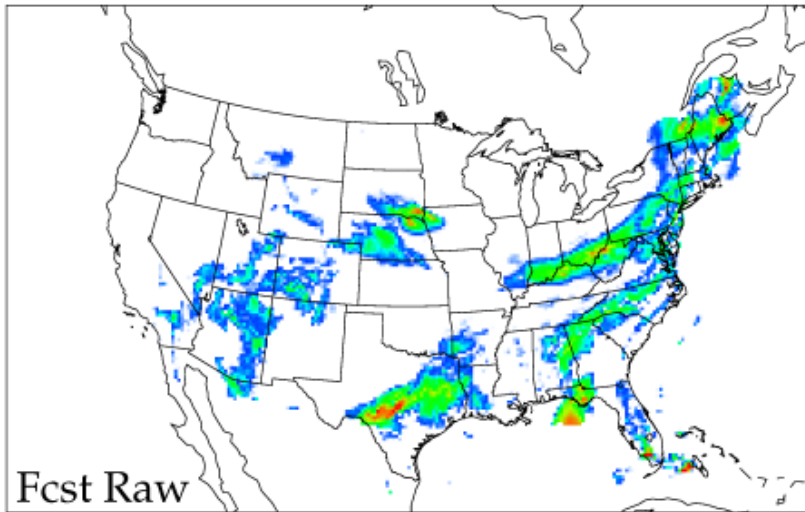
This gives us our objects.

© 2009 University Corporation for Atmospheric Research. All Rights Reserved.

Once you have objects

- **Merging** - associating objects in the same field to form “clusters” that may be more representative of a broader scale organization (i.e. within the forecast field and the observation field)
 - Uses a fuzzy logic algorithm - and/or -
 - A second – slightly lower threshold
- **Matching** – associating objects and clusters in different fields (i.e. between forecast and observation field)
 - Uses a fuzzy logic algorithm

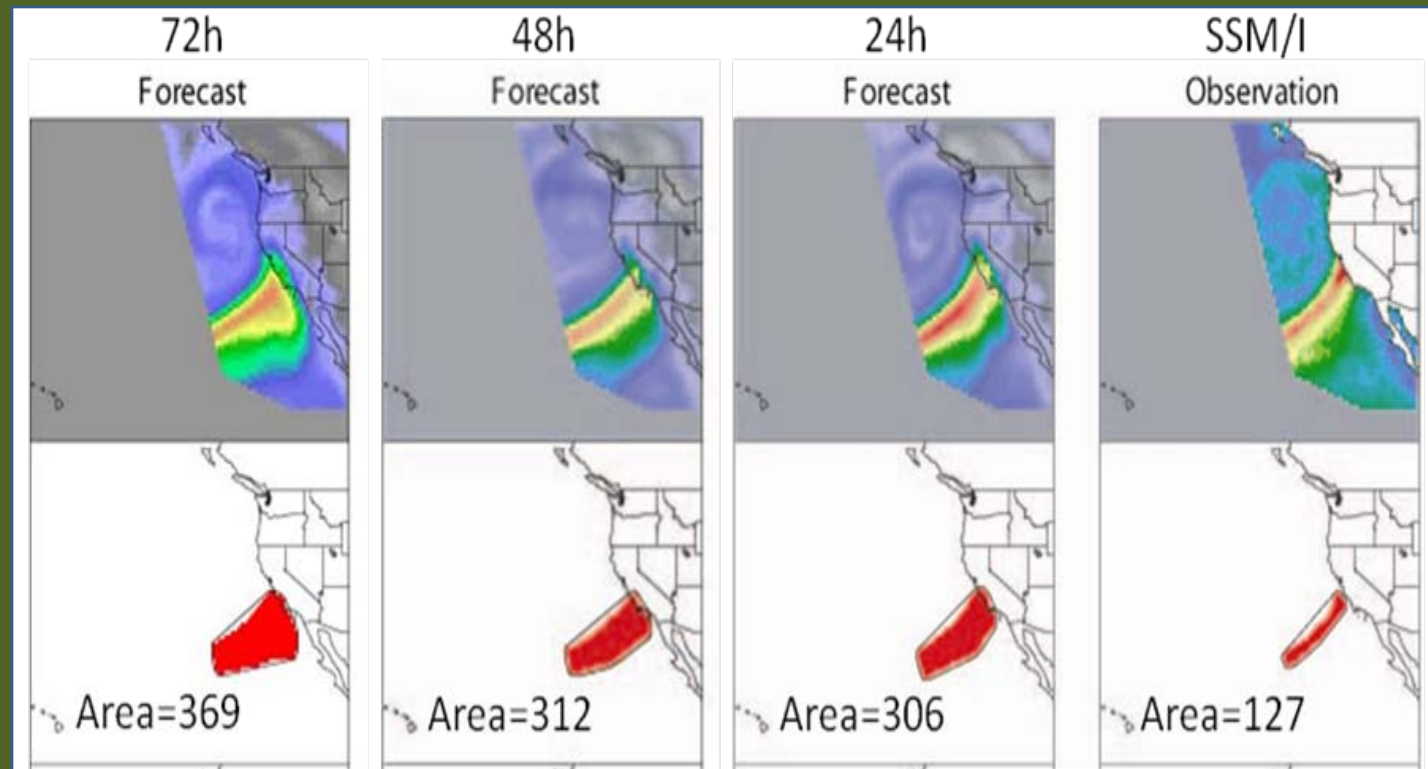
Example of Matching & Merging



Use of MODE to show changes in IWV Component of AR PWAT forecasts vs SSM/I Observations

MODE Object Comparison of
GFS Forecasts with SSM/I Observation
for 25 February, 2004

In this case the forecast width and location of AR landfall appears to be a function of forecast lead time.



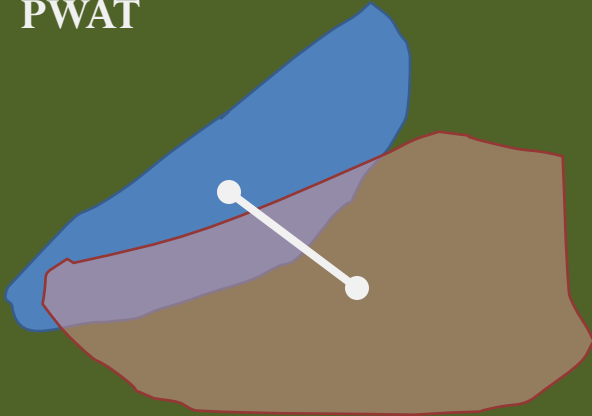
Use of Attributes of Objects defined by MODE



Forecasted
PWAT

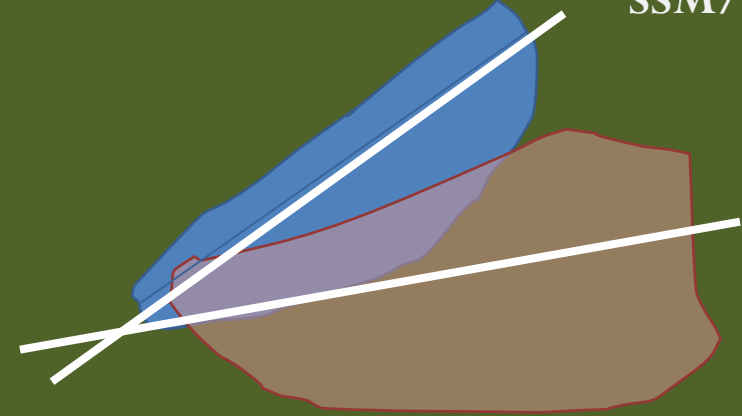


Observed
SSM/I IWV



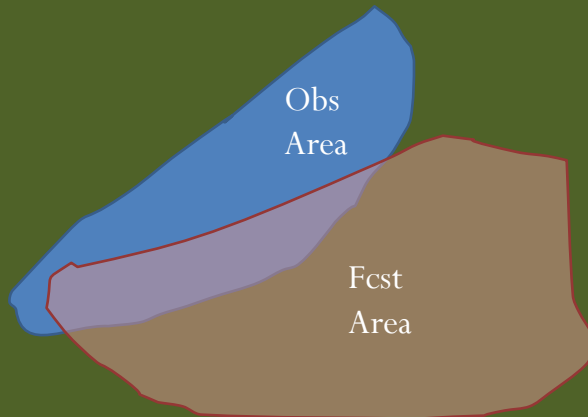
Centroid Distance: Provides a quantitative sense of spatial Displacement of AR core.

Small is good



Axis Angle: Provides an objective measure of how well the AR impact on terrain is captured. *Small is good*

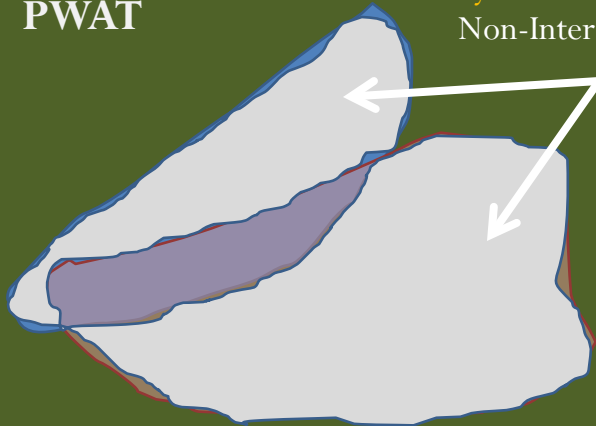
Area Ratio =
 $\frac{\text{Fcst Area}}{\text{Obs Area}}$



Area Ratio: Provides an objective measure of whether there is an over- or under-prediction of areal extent of AR. *Close to 1 is good*

Use of Attributes of Objects defined by MODE

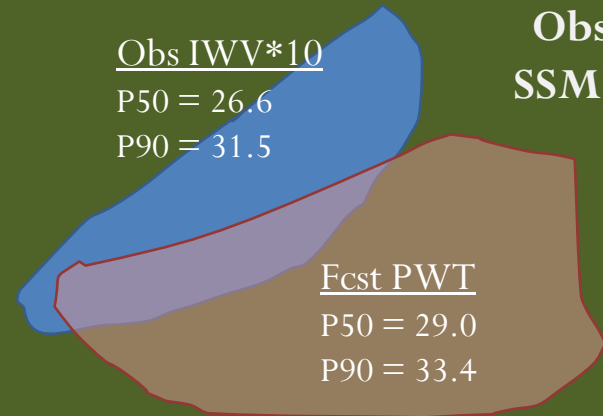
Forecasted
PWAT



Symmetric Difference:
Non-Intersecting Area

Symmetric Diff: May be a good summary statistic for how well Forecast and Observed objects match. *Small is good*

Observed
SSM/I IWV

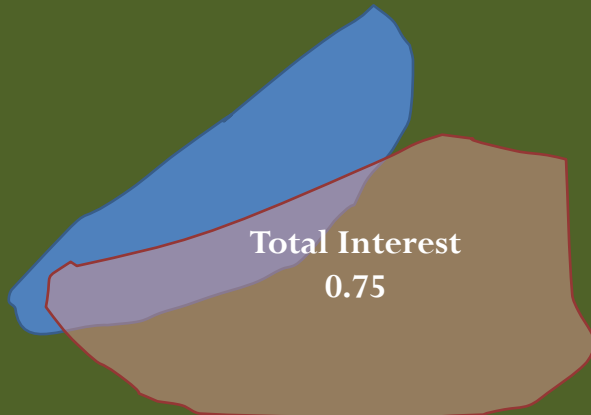


Obs IWV*10
P50 = 26.6
P90 = 31.5

Fcst PWT
P50 = 29.0
P90 = 33.4

P50/P90 Int: Provides objective measures of Median (50th percentile) and near-Peak (90th percentile) intensities found in objects.

Ratio close To 1 is good



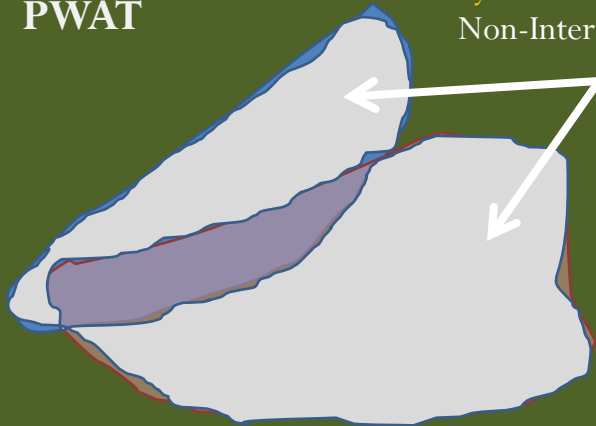
Total Interest
0.75

Total Interest: Summary statistic derived from fuzzy logic engine with user-defined Interest Maps for all these attributes plus some others.

Close to 1 is good

Use of Attributes of Objects defined by MODE

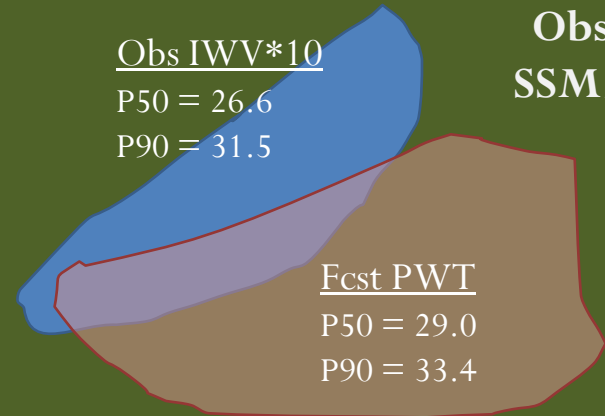
Forecasted
PWAT



Symmetric Difference:
Non-Intersecting Area

Symmetric Diff: May be a good summary statistic for how well Forecast and Observed objects match. *Small is good*

Observed
SSM/I IWV



Obs IWV*10

P50 = 26.6

P90 = 31.5

Fcst PWT

P50 = 29.0

P90 = 33.4

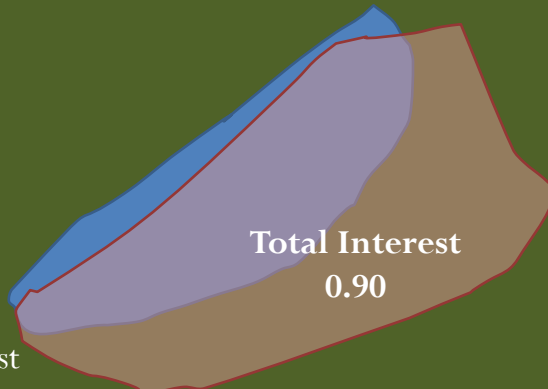
P50/P90 Int: Provides objective measures of Median (50th percentile) and near-Peak (90th percentile) intensities found in objects.

Ratio close To 1 is good

Total Interest: Summary statistic derived from fuzzy logic engine with user-defined Interest Maps for all these attributes plus some others.

Close to 1 is good

If forecast
was rotated
and moved
North –
Total Interest
may increase



Total Interest
0.90

Question?

- Next – run through the website
- After – Ed will discuss the 20 January case study to demonstrate how this site might be used ([DTC-HMT eval 20jan2010 case.pdf](#))