Integrating Objective Evaluation into Operations The Development Testbed Center's Approach

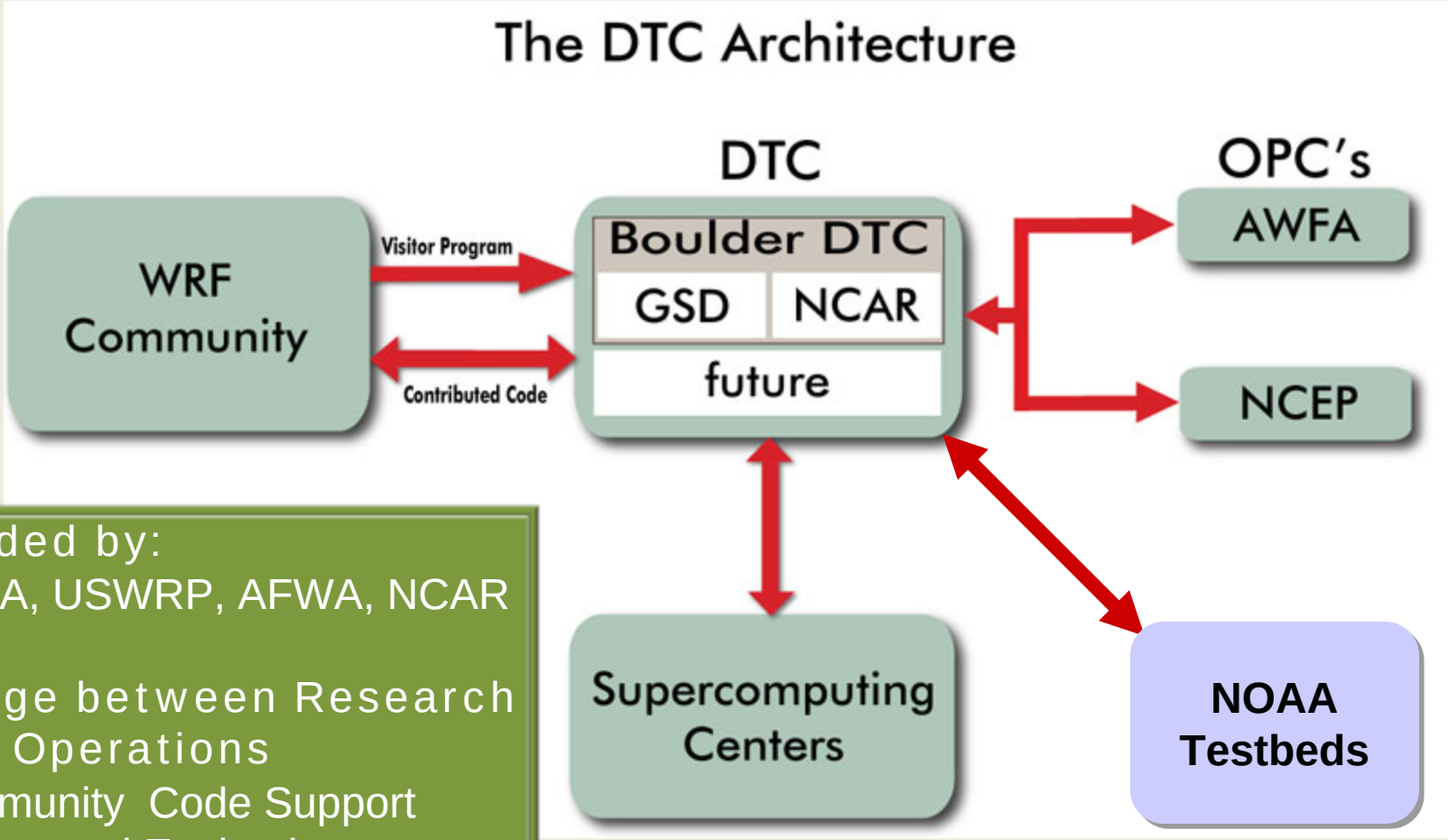
Tara Jensen
Development Testbed Center
and NCAR/Research Applications Lab
Boulder, Colorado

National Severe Weather Workshop
Norman, OK
4-6 March 2010

THE DEVELOPMENTAL TESTBED CENTER (DTC)

The Development Testbed Center (DTC)

Who we are and what we do…



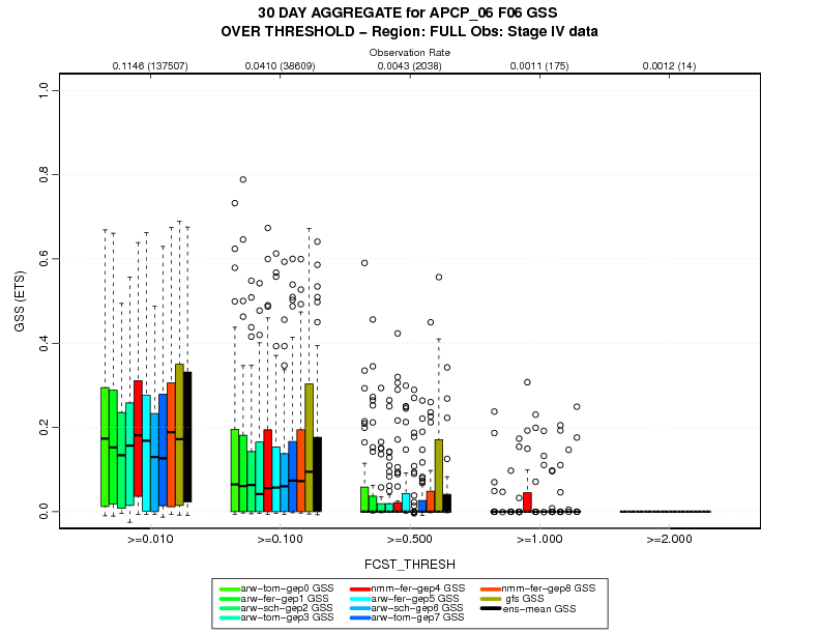
Funded by:
NOAA, USWRP, AFWA, NCAR

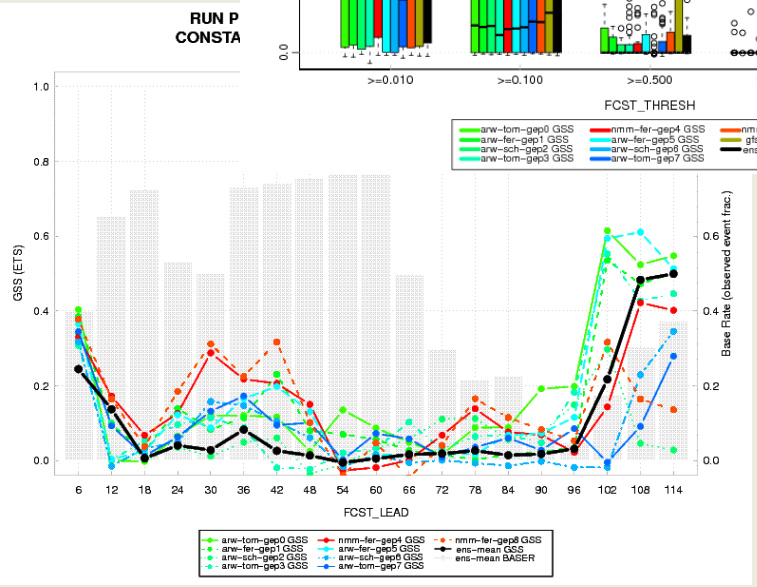
Bridge between Research And Operations
Community Code Support
Testing and Evaluation
Verification Research

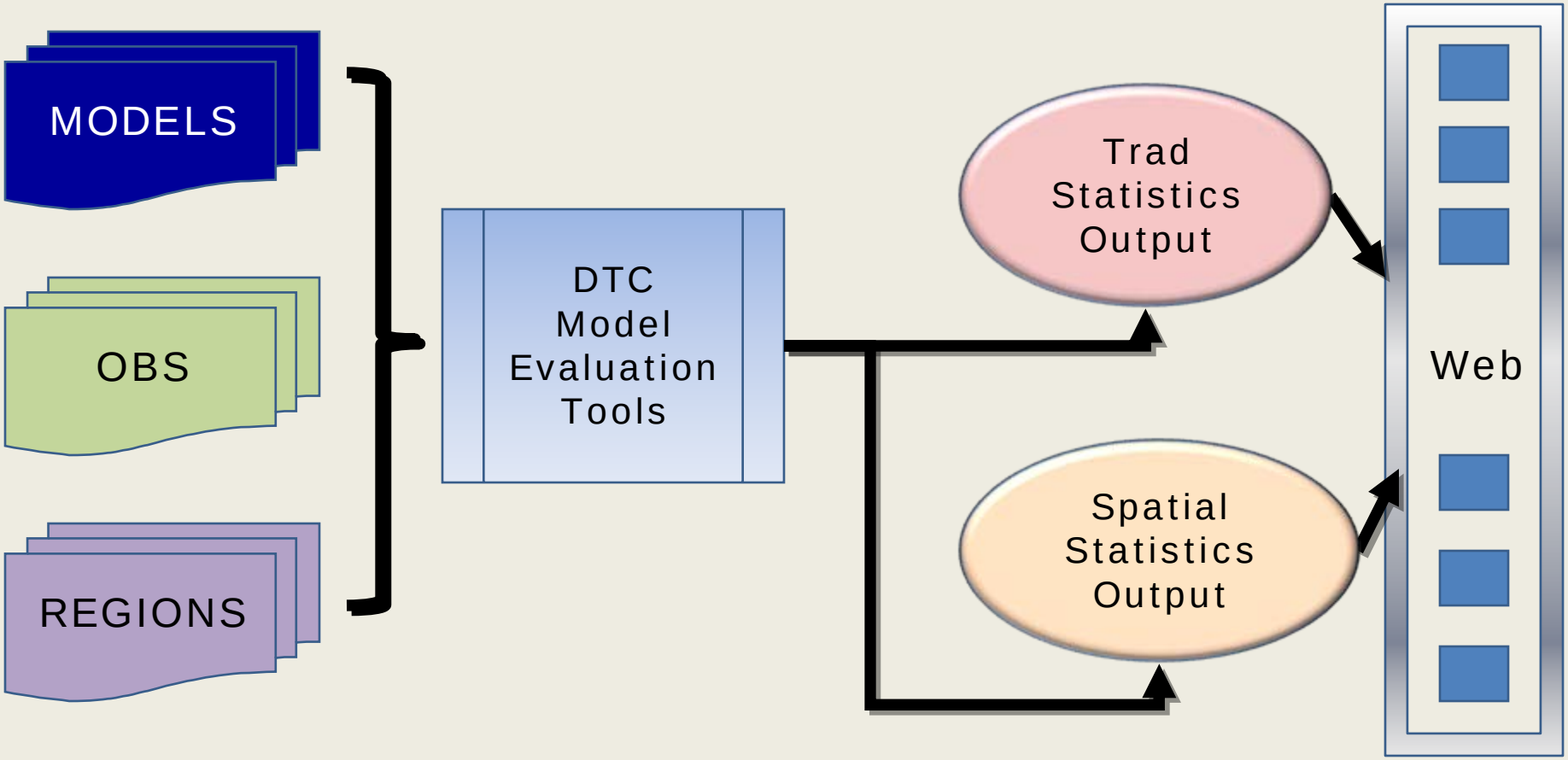
DTC: Objective Evaluation
Direct Infusion into Operations

- Hazardous Weather Testbed (HWT)
  - SPC and NSSL
  - Focus – Severe Weather over US
  - Spring Experiments: 2008, 2009, 2010
- Hydrometeorology Testbed (HMT)
  - WFOs and RFCs, GSD, HPC, OHD, NSSL
  - Focus – Extreme Precipitation Events in Pacific NW (especially American River Basin)
  - Winter Forecast Exercise: 2010

Objective versus Subjective







General Approach for

Objective Evaluation of Contributed Research Models

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

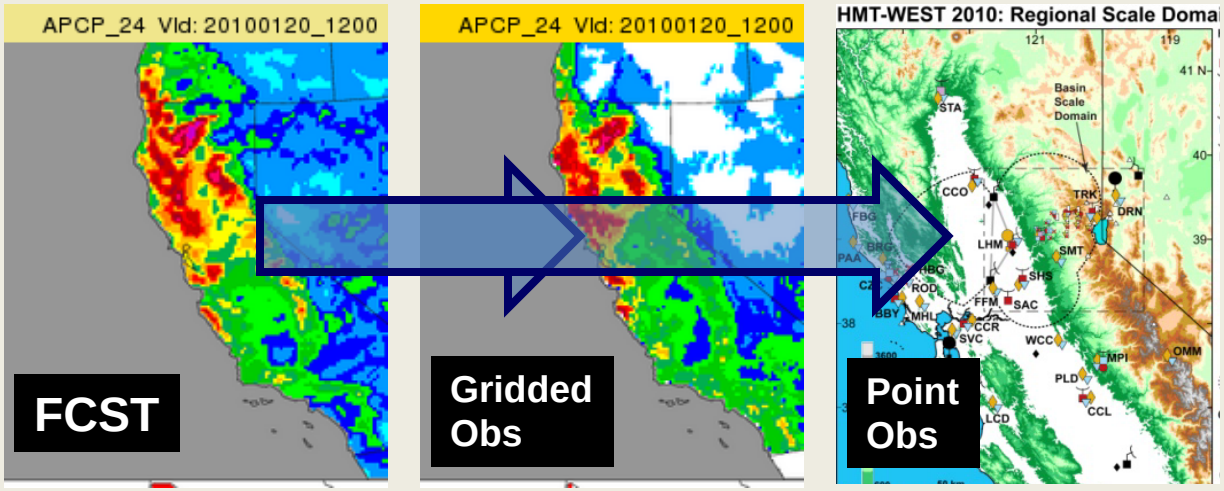
- Traditional methods
- Spatial methods

Post-processing

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

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

- 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

What’s Can Users See on the Web

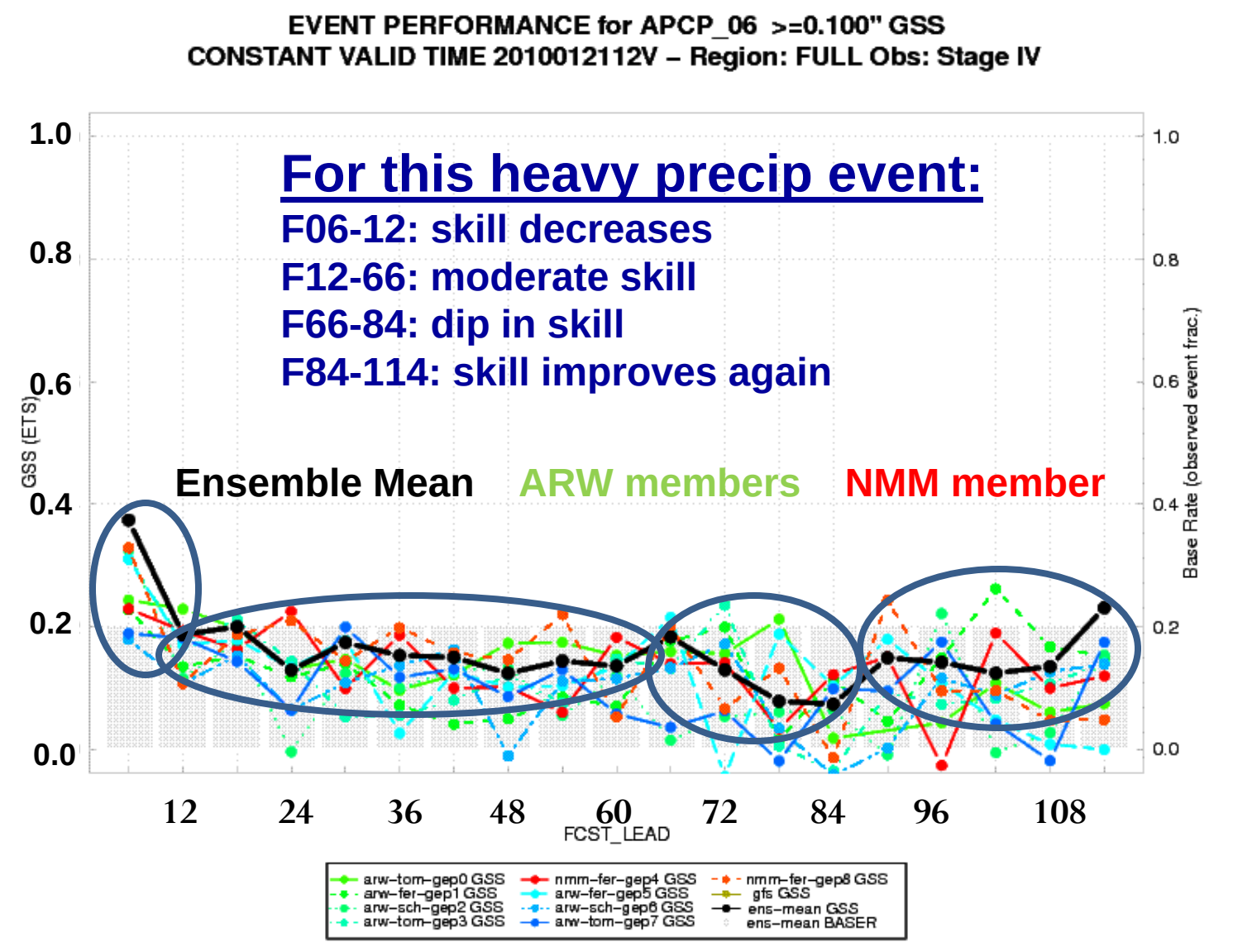
For HMT

Looking at 6hr and 24hr accumulated precip

Thresholds: Trace, 0.1, 0.5, 1.0, 2.0, 3.0 inch
Primary interest in events with >1.0 in

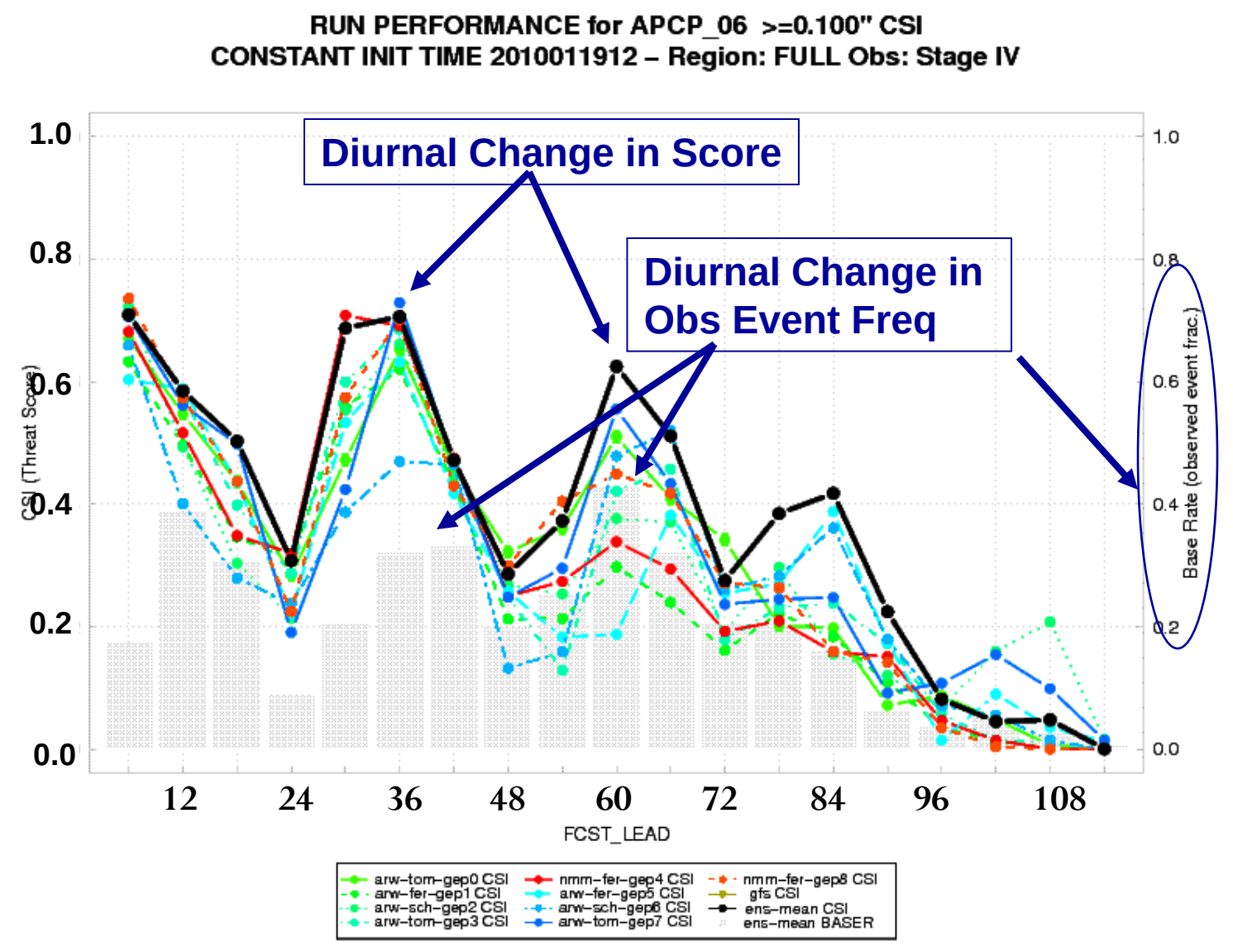
NOTE: ALL RESULTS IN THIS PRESENTATION ARE PRELIMINARY

Event Performance – From HMT



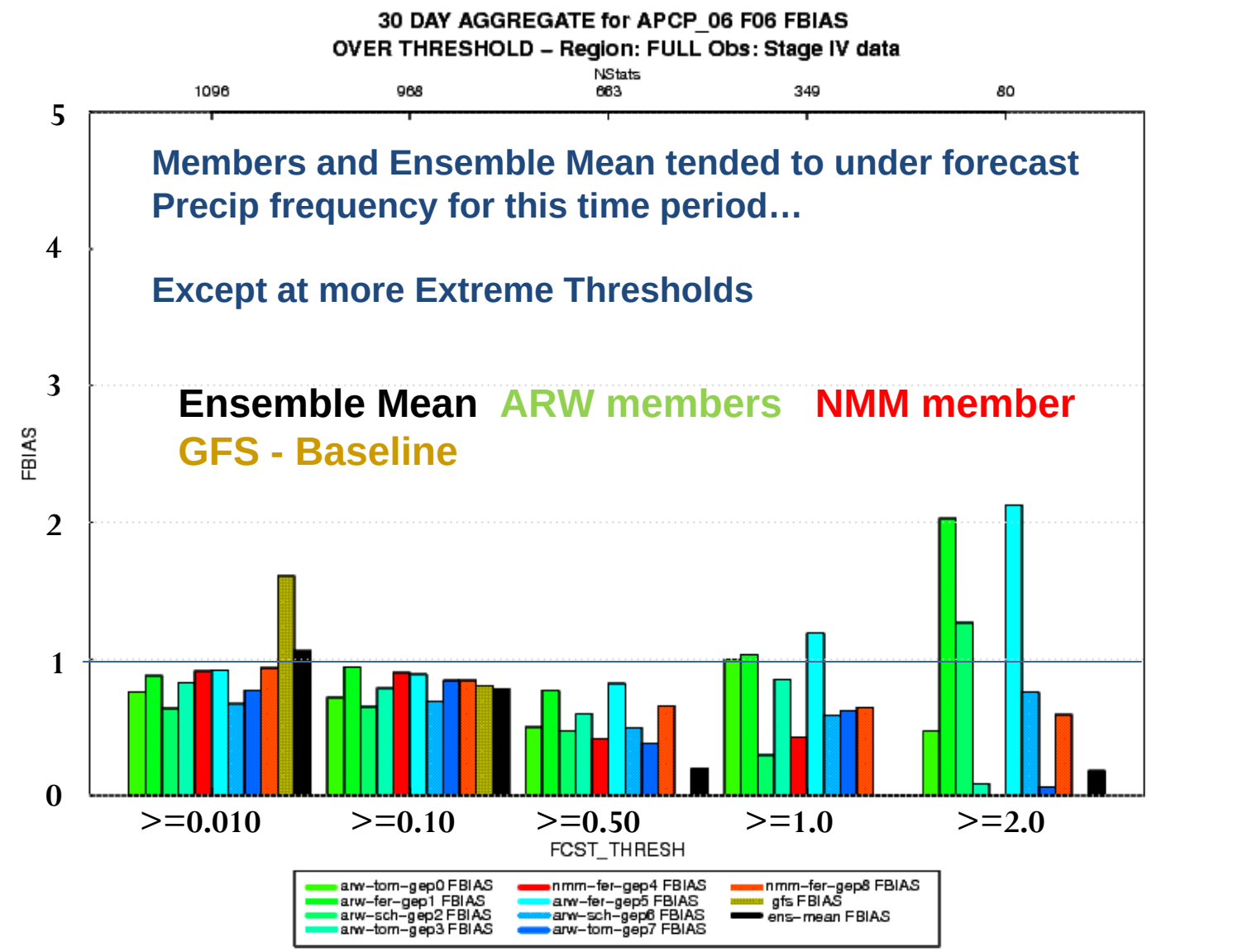
Look at how forecasts change over time for a given event

Run Performance – From HMT



Look at how forecast skill changes for a given run

Aggregated Performance – From HMT



Look at Statistics Aggregated over 30 days and stratified by threshold

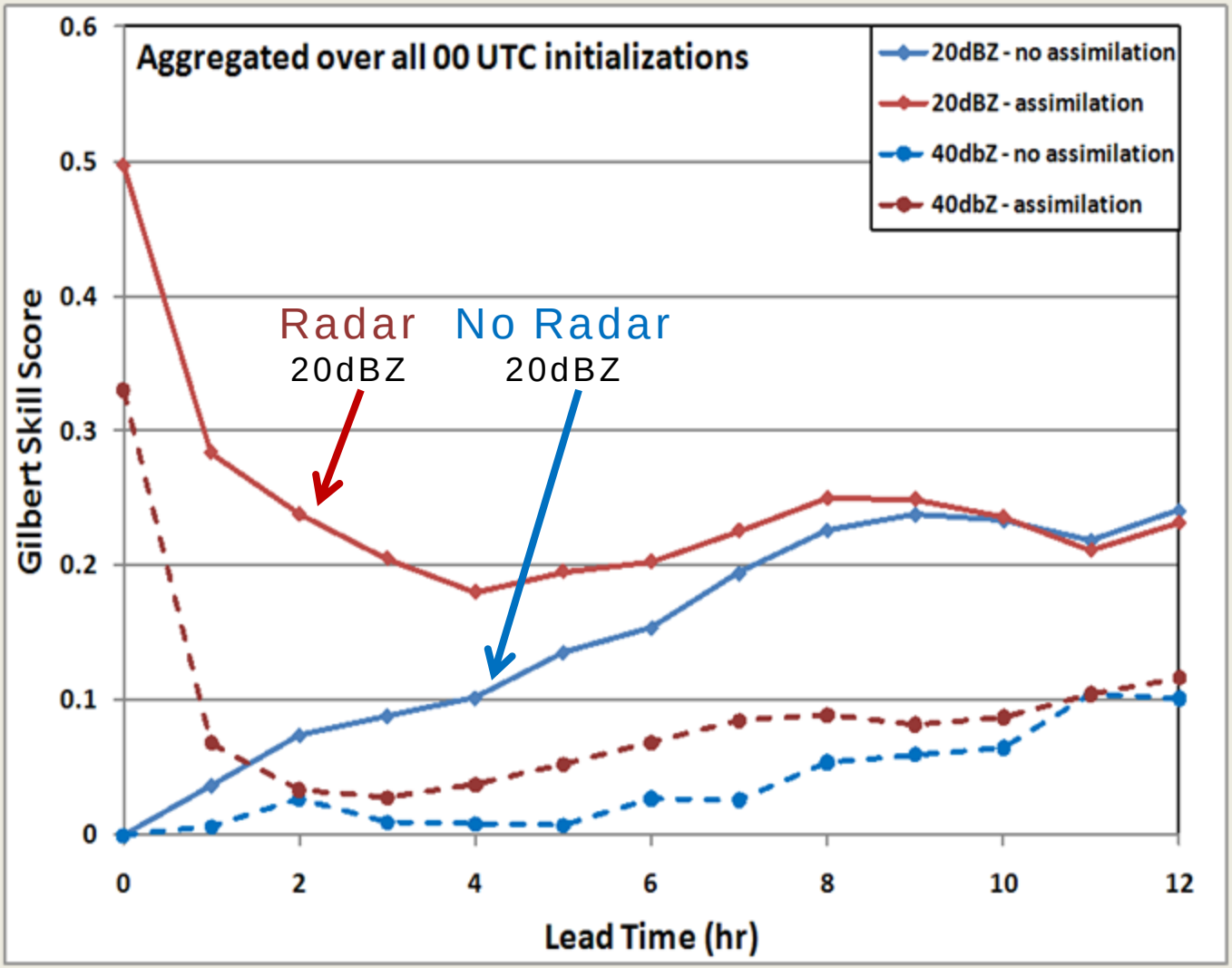
It’s Not All About Accumulated Precip

For HWT

Looked at Simulated Radar Reflectivity Field also

Impact of radar assimilation on two members of the CAPS Storm Scale Ensemble Forecast System (4 km grid spacing)

Aggregated Statistics – From HWT



Results were aggregated over Spring Experiment time period and the median values are plotted

RESULTS:

Radar assimilation appears to improve 0-6hr skill scores

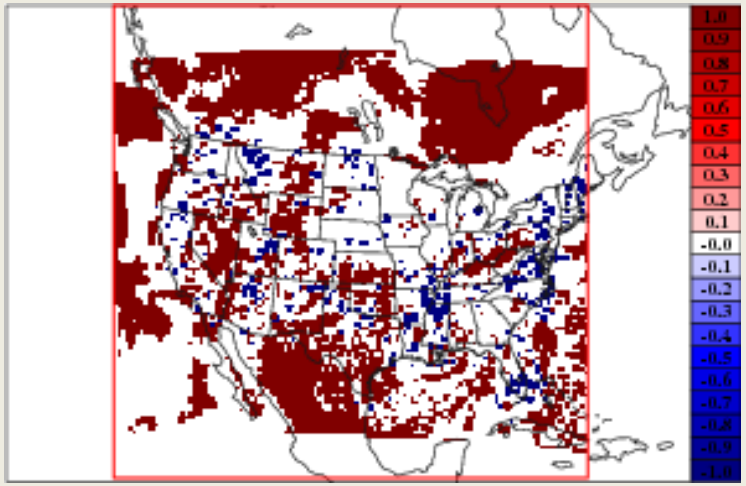
Lack of clear difference in skill scores during 6-12 hr lead times suggests model physics taking over

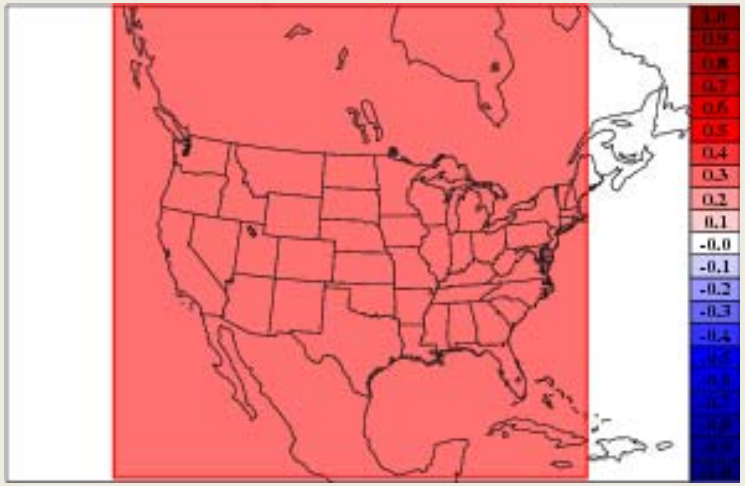
MET Spatial Verification (using MODE)

Higher Resolution forecasts of spatially-coherent fields (e.g., precipitation) are typically penalized using traditional statistics

Questions Spatial Methods might answer:

- Is there typically a position, intensity, or orientation error?
- How does the forecast perform on attributes that are of interest to users?

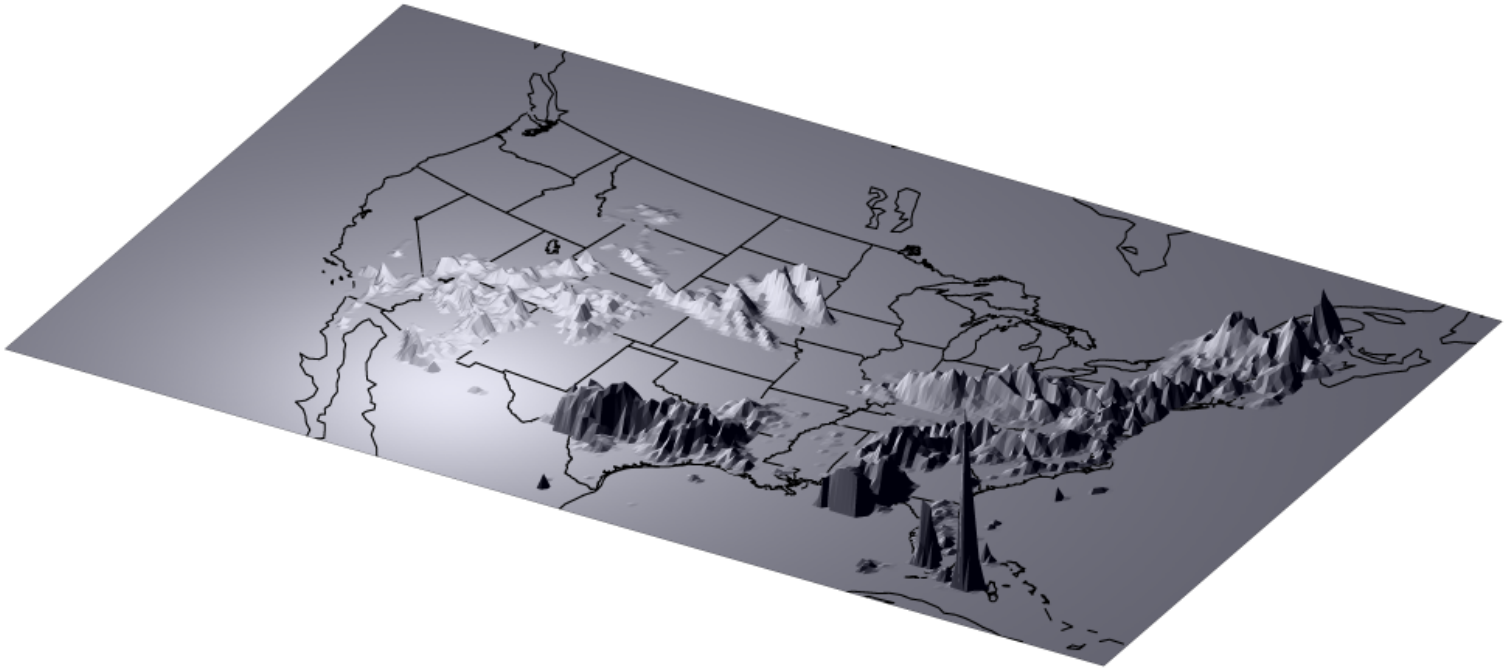




It's much easier to have a "hit" at coarser resolution

MODE – Spatial Verification

Method For Object-Based Diagnostic Evaluation



Step #1

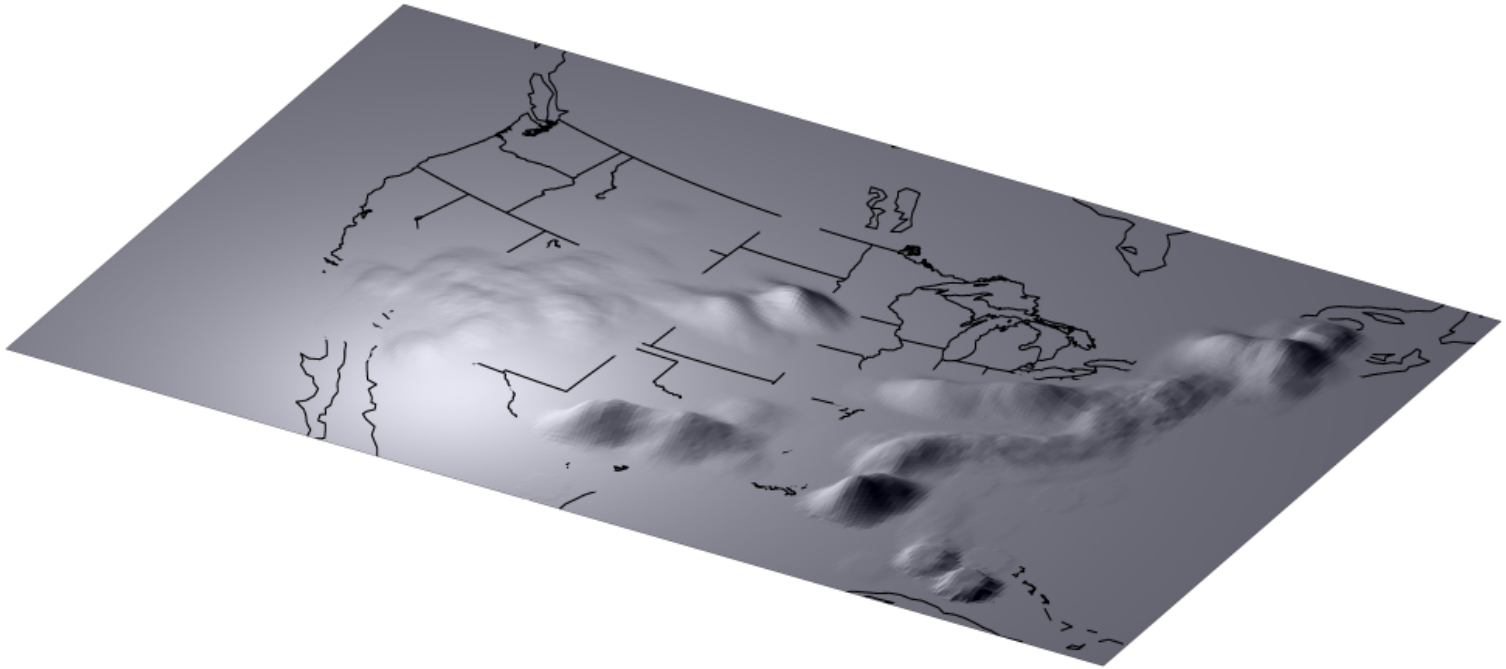
Start with the raw data field.

In this case, a precipitation field.

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

MODE – Spatial Verification

Method For Object-Based Diagnostic Evaluation



Step #2

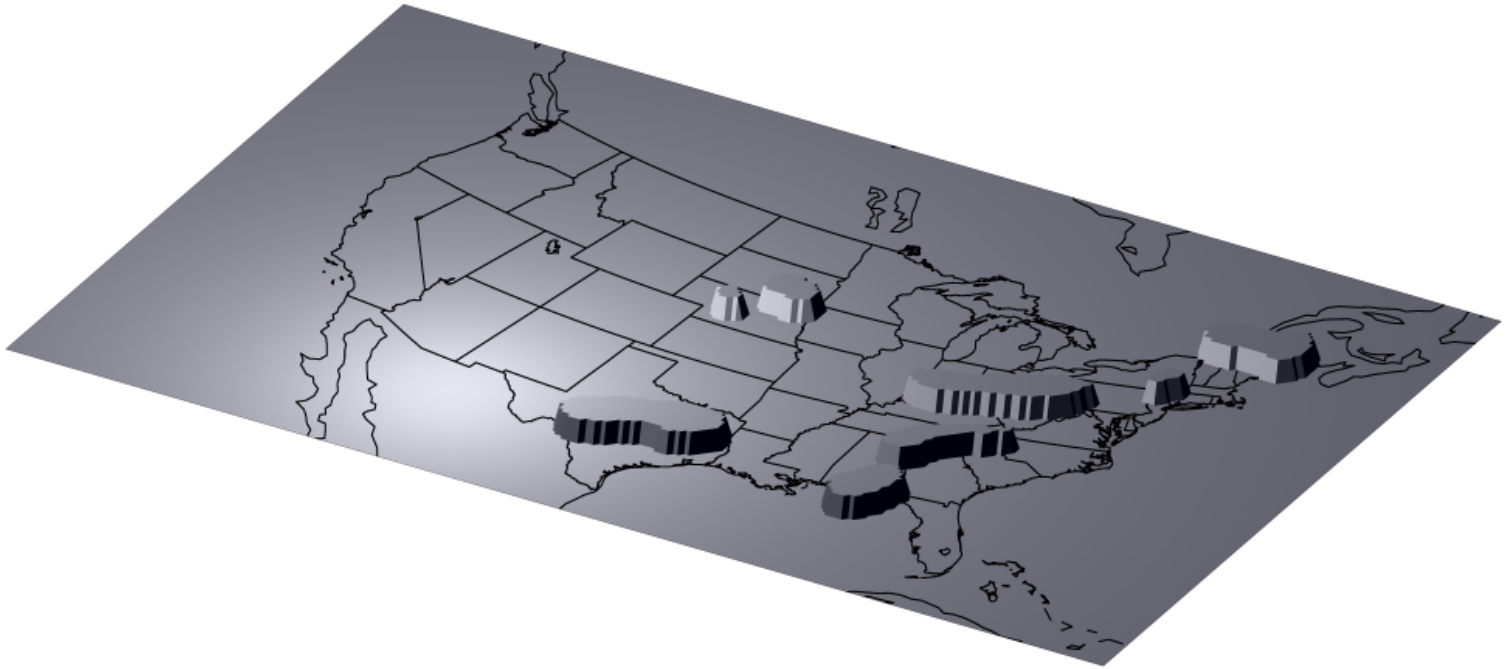
Apply convolution operator.

This is basically a smoothing operation.

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

MODE – Spatial Verification

Method For Object-Based Diagnostic Evaluation



Step #3

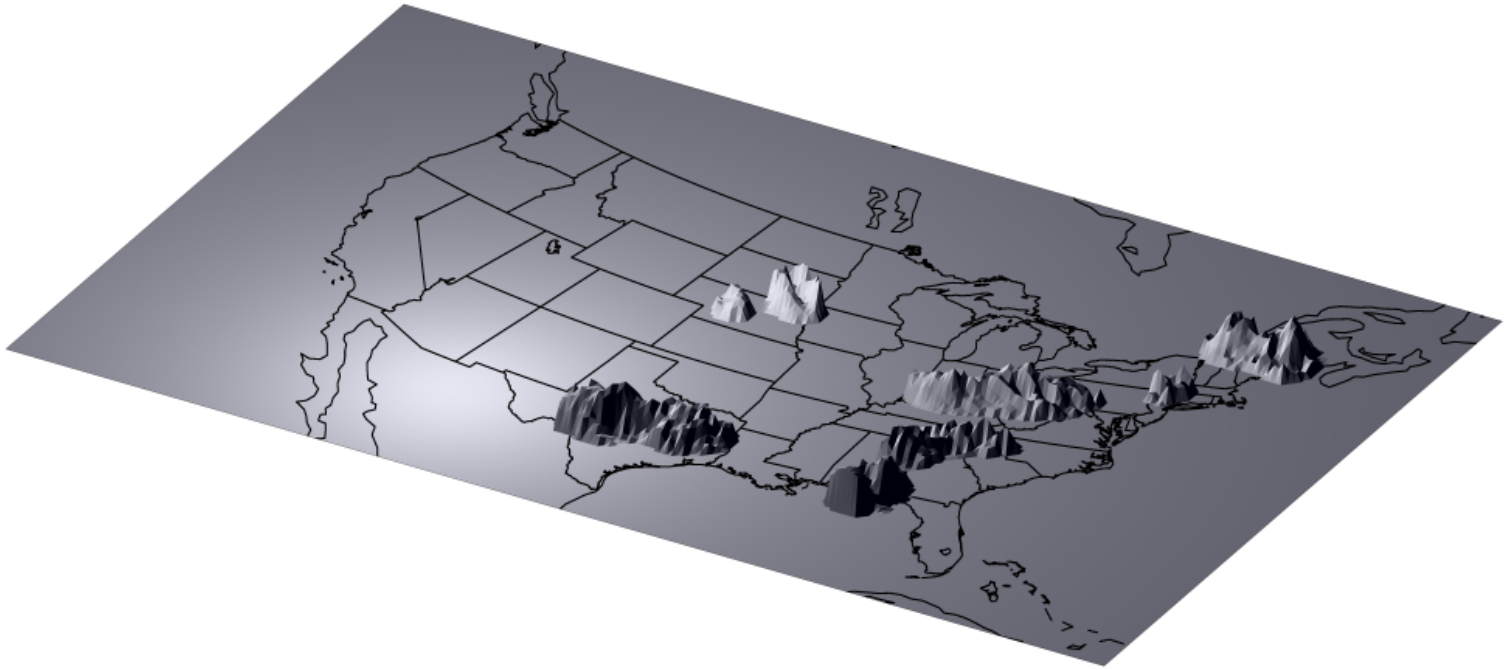
Threshold the smoothed field.

This produces an on/off mask field.

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

MODE – Spatial Verification

Method For Object-Based Diagnostic Evaluation



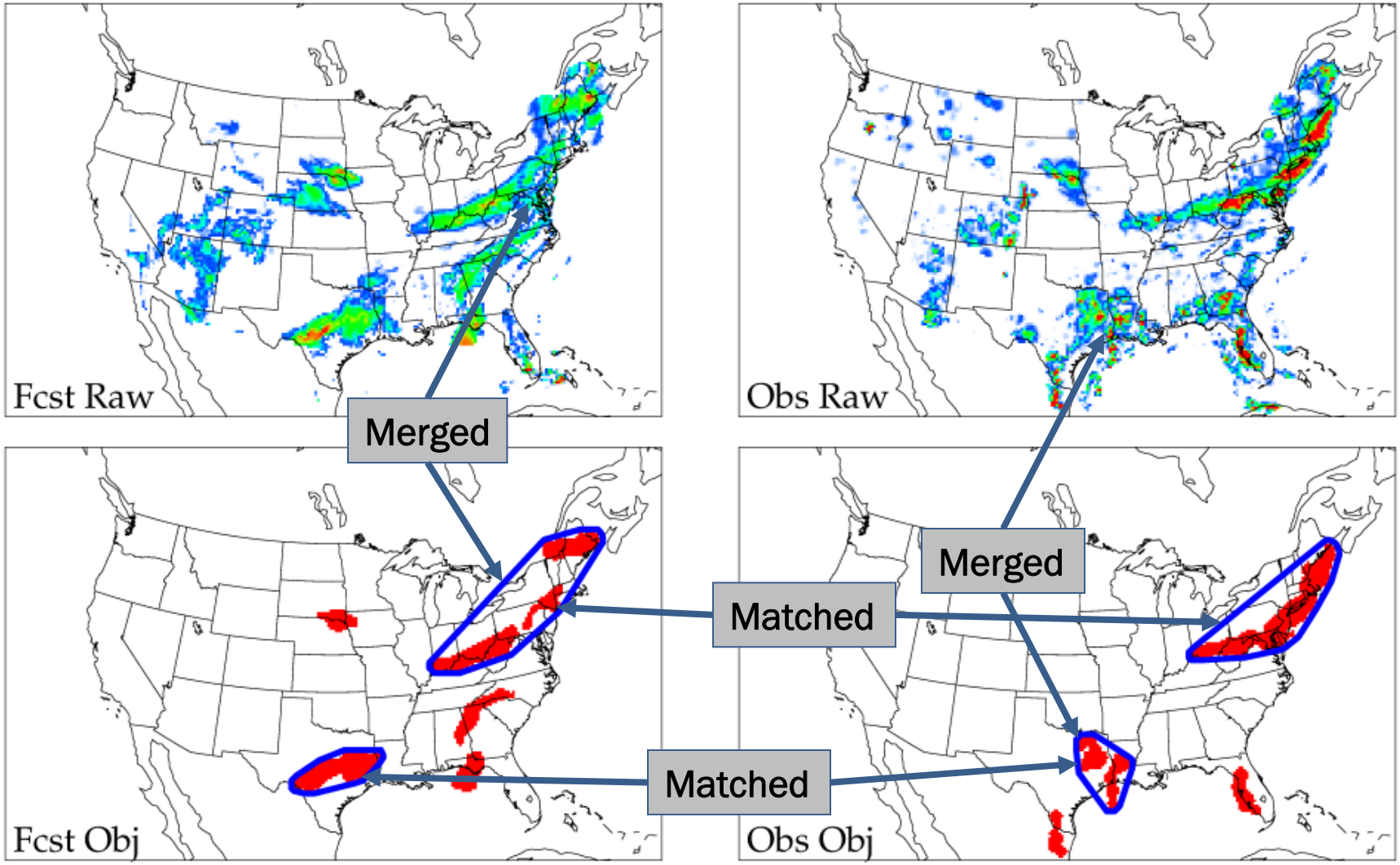
Step #4

Restore original data to object interiors.

This gives us our objects.

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

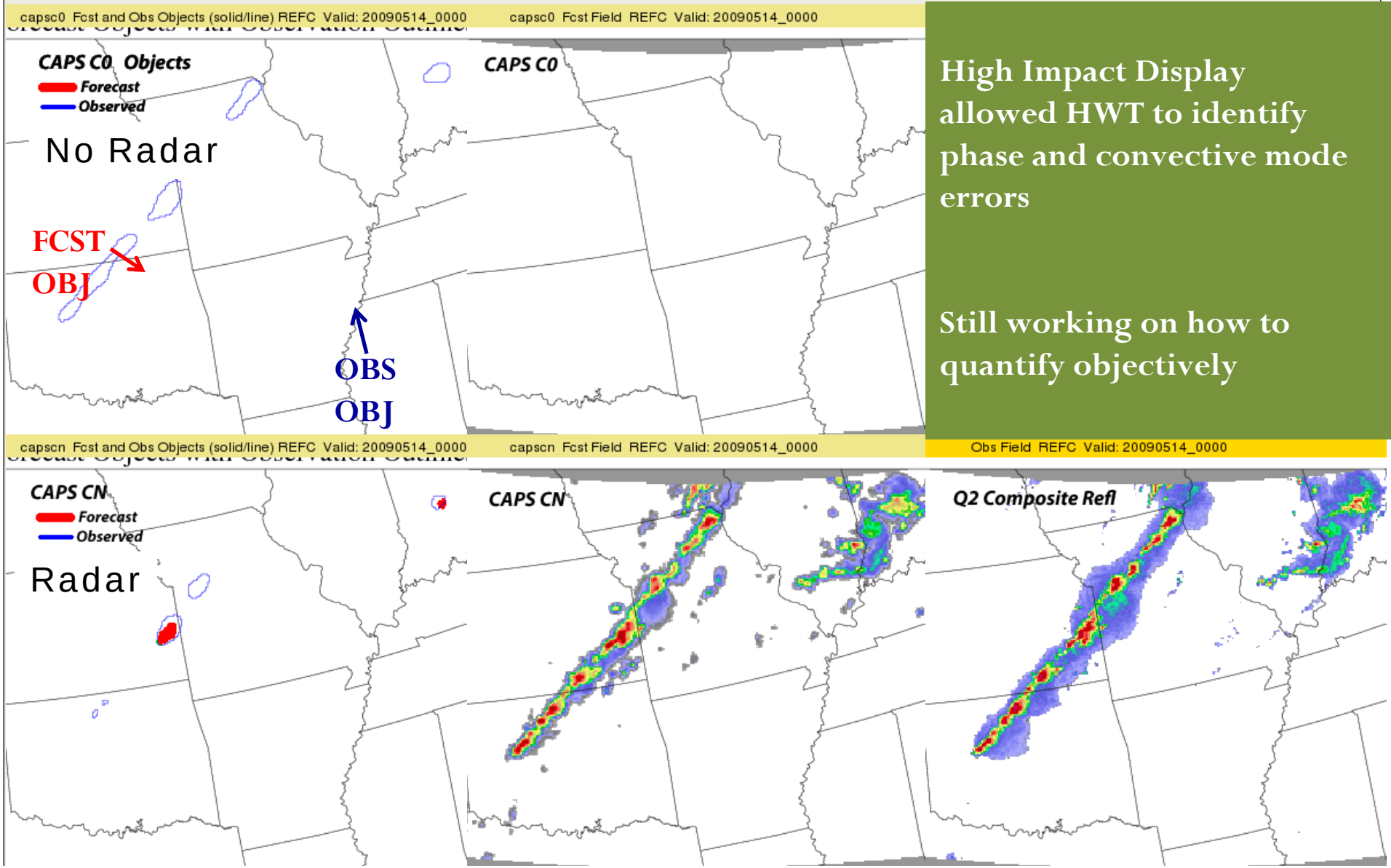
Example of Matching & Merging



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

How Objects Helped HWT Identify a Timing/Location Error

14 May 2009 Init: 00 UTC Spatial Thresh: 30dBZ



High Impact Display allowed HWT to identify phase and convective mode errors

Still working on how to quantify objectively

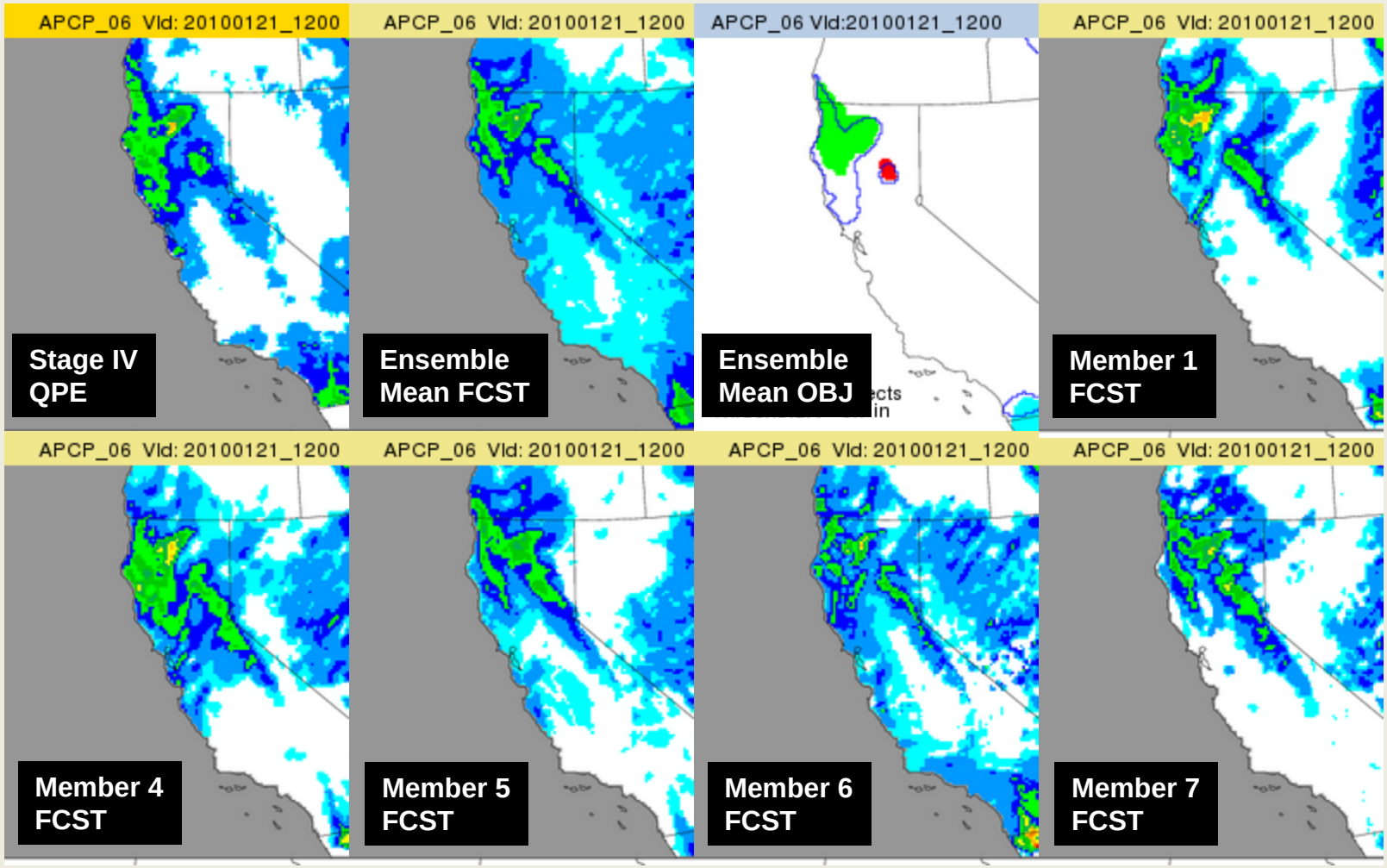
Objects

Forecast Field

Observed Field

How Objects Help with Understanding Ensemble Output

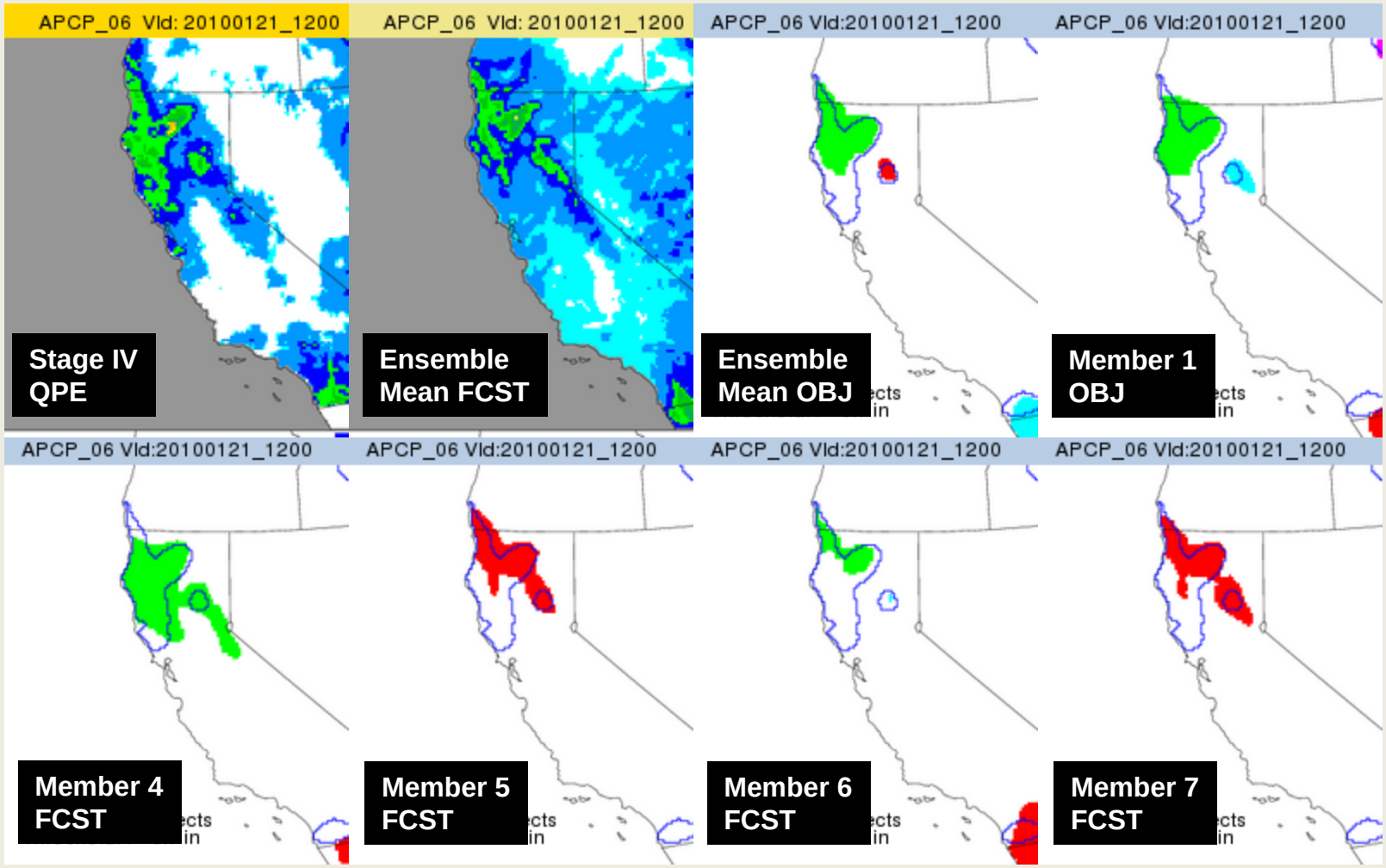
HMT Ensemble Forecast Fields



4

More Plots

HMT Ensemble Object Fields



4

More Plots

Where do we go from here?

- Continued Education at Testbeds on current web displays
- Integration of MET and METviewer (database and display system) into computing at Forecast Offices and Prediction Centers
- Pre-generated information in NOAA/ESRL ALPS system to demonstrate utility for AWIPS II
- ULTIMATELY: Coupling between AWIPS-type systems and MET system

Questions?

Useful Links:

http://www.dtcenter.org/plots/hwt/index.php
http://verif.rap.ucar.edu/eval/hmt/2010/index.php

Thanks to the following agencies for their support of this work!!

USWRP The U.S. Weather Research Program

NCAR

NOAA

DTC The Developmental Testbed Center (DTC)

Air Force Weather Agency