Data Representativeness:
Are you observing what you want to observe or what you *need* to observe?

John Horel
Department of Meteorology
University of Utah
john.horel@utah.edu

- Acknowledgements

  – Dan Tyndall (Univ. of Utah)
  – Dave Myrick (WRH/SSD)

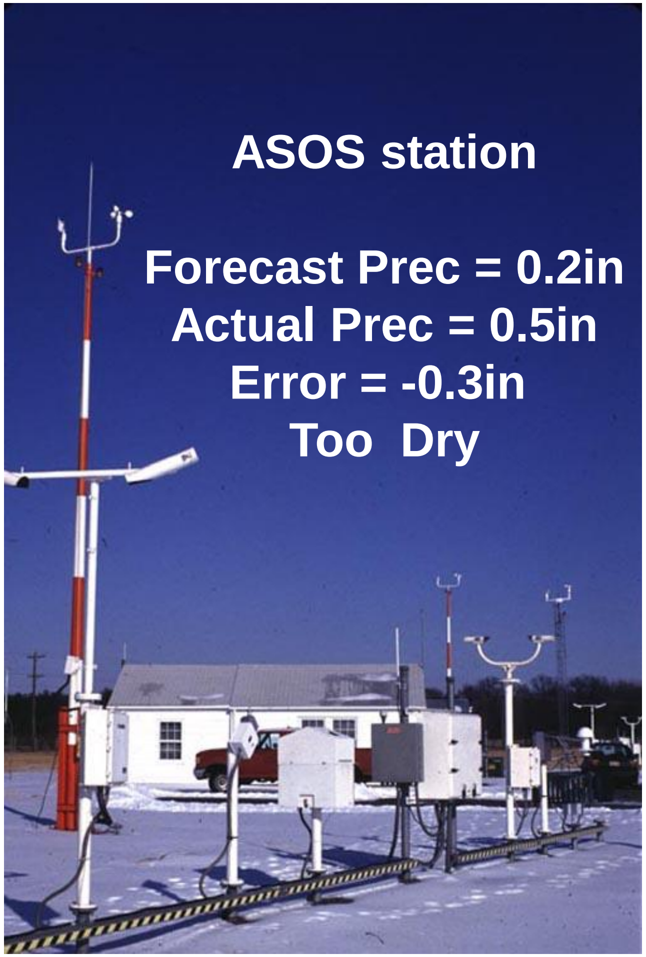
- References

  – Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge
  – Lockhart, T., 2003: Challenges of Measurements. *Handbook of Weather, Climate and Water*. Wiley & Sons. 695-710.
  – Myrick, D., and J. Horel, 2006: Verification over the Western United States of surface temperature forecasts from the National Digital Forecast Database. *Wea. Forecasting*, 21, 869-892.
  – Vasiloff, S. and coauthors, 2007: Improving QPE and Very Short Term QPF: An Initiative for a Community-Wide Integrated *Bull. Amer. Meteor. Soc.*, 88, 1899–1911

How Well Can We Observe, Analyze, and Forecast Conditions Near the Surface?

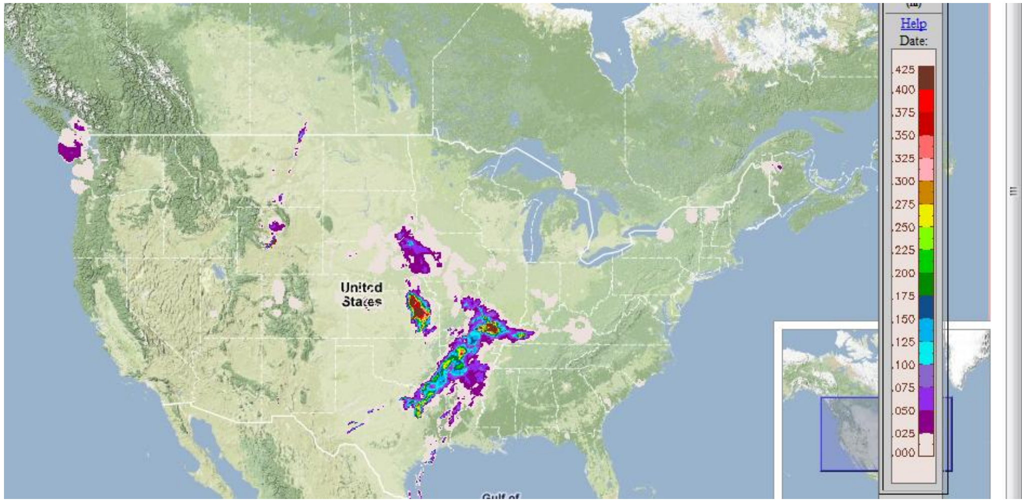
- Forecasters clearly recognize large variations in surface temperature, wind, moisture, precipitation exist over short distances:
  - in regions of complex terrain
  - when little lateral/vertical mixing
  - due to convective precipitation
- To what extent can you rely on surface observations to define conditions within 2.5 x 2.5 or 5 x 5 km^2 grid box?
- Do we have enough observations to do so?
- What is it going to take to get a national effort to collect, manage, and distribute mesonet observations necessary for verification as well as a myriad other applications?
- Need to support efforts to collect and manage metadata
- Need to recognize errors inherent in observations and use that error information for analyses, forecast preparation, & verification

Viewing the atmosphere in terms of grids vs. points



ASOS station

Forecast Prec = 0.2in
Actual Prec = 0.5in
Error = -0.3in
Too Dry



What about away from ASOS stations?

Need an integrated analysis of observations

Developing Mesoscale Meteorological Observational Capabilities to Meet Multiple National Needs

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

- Committee charged to:
  - develop an overarching vision for an integrated, flexible, adaptive, and multi-purpose mesoscale meteorological observation network
  - seek to identify specific steps to help develop a network that meets multiple national needs in a cost-effective manner.
- Starting from existing information:
  1. characterize the current state of mesoscale atmospheric observations and purposes;
  2. compare the U.S. mesoscale atmospheric observing system to other observing system benchmarks;
  3. describe desirable attributes of an integrated national mesoscale observing system;
  4. identify steps to enhance and extend mesoscale meteorological observing capabilities so they meet multiple national needs; and
  5. recommend practical steps to transform and modernize current, limited mesoscale meteorological observing capabilities to better meet the needs of a broad range of users and improve cost effectiveness.

Report due soon…

Observations are not perfect…

- Metadata errors
- Gross errors
- Local siting errors (e.g., artificial heat source, overhanging vegetation, observation at variable height above ground due to snowpack)
- Instrument errors (e.g., exposure, maintenance, sampling)
- Representativeness errors: correct observations that are capturing phenomena that are not representative of surroundings on broader scale (e.g., observations in vegetation-free valleys and basins surrounded by forested mountains)

All that is labeled data Is NOT gold!

Lockhart (2003)

Are All Observations Equally Good?

- Why was the sensor installed?
  - Observing needs and sampling strategies vary (air quality, fire weather, road weather)
- Station siting results from pragmatic tradeoffs: power, communication, obstacles, access
- Use common sense and experience
  - Wind sensor in the base of a mountain pass will likely blow from only two directions
  - Errors depend upon conditions (e.g., temperature spikes common with calm winds)
  - Pay attention to metadata
- Monitor quality control information
  - Basic consistency checks
  - Comparison to other stations

Real-Time Precipitation Data

- Hardest to manage due to differences in
  - Equipment and measurement technique
  - Measurement type (interval, sum)
  - Reporting interval (5 min-24 hour)
- Hardest to quality control
  - Unheated tipping buckets
  - Representativeness issues
- Difficult to integrate QC procedures developed for hydrologic applications (e.g., 24-h total QC'd data from NRCS) into real-time data stream

Observing Precipitation: Remote Sensors
Vasiloff et al. (2007)

Sensor	Time-Space Scales	Strengths	Weaknesses
Radar	5-10 min, 1km	• High spatial and temporal resolution • Good areal coverage*	• Range effects •Coverage in complex terrain •Z-R and Z-S uncertainties •Target contamination
Geostationary satellite	15 min, 4 km	Continuous spatial coverage	• Indirect measurement • Sorting out nonprecipitating clouds
Polar-orbiting satellite (passive microwave)	3-6h+, 15 km	Continuous spatial coverage	• Poor spatial/temporal resolution • Indirect measurement • Difficulty with non-ice clouds

Observing Precipitation: Gauges
Vasiloff et al. (2007) +

Sensor	Time-Space Scales	Strengths	Weaknesses
Unheated/heated tipping bucket, Belfort, ETI, Geonor weighing gauges, snow pillows	10 min-1 day, network dependent	•Direct measurement	• Nonuniform spatial distribution • Latency in real-time data transfer • Quality control • Frozen hydrometeors • Wind effects • Calibration issues as function of rain/snow rate

- Integrating gauge observations is a challenge…
- Integrating gauges AND remote sensing information is even more of a challenge…
- Integrating all observations AND prior model forecast/analysis is the greatest challenge

GOES Platforms

- RAWS
  - More agencies are using RAWS as a means to collect observations beyond simply fire weather applications (e.g., air quality)
- HADS: Accessing GOES DCPs
  - 2500+ mostly precipitation reporting stations received via HADS
  - We depend on WFOs (HADS focal points/service hydrologists) to manage station metadata updates via the NWSLI system

Some of the National & Regional Mesonet Data Collection Efforts

MADIS

IEM

NJ Weather & Climate NETWORK

CoCoRaHS
Community Collaborative Rain, Hail & Snow Network

Weather Underground

RAWS

NATIONAL WEATHER SERVICE

HOMELAND SECURITY WEATHER NETWORK

NOAA Regional Climate Centers

Northwest Atmospheric Modeling System

FAWN Florida Automated Weather Network

Citizen Weather Observer Program
APRSWXNET

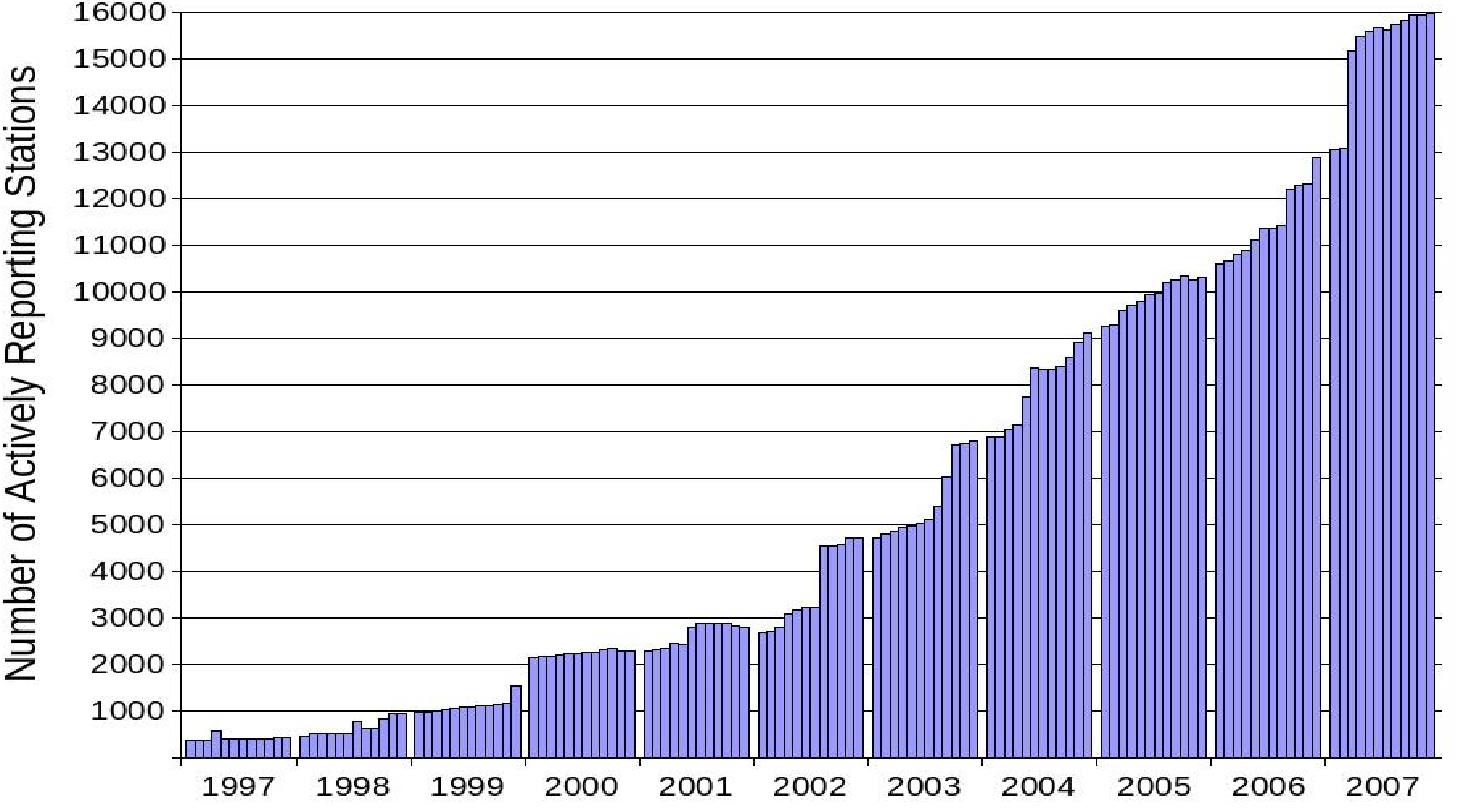
MESO WEST

Oklahoma Mesonet

Arizona Cooperative Extension
AZMET

WEST TEXAS MESONET

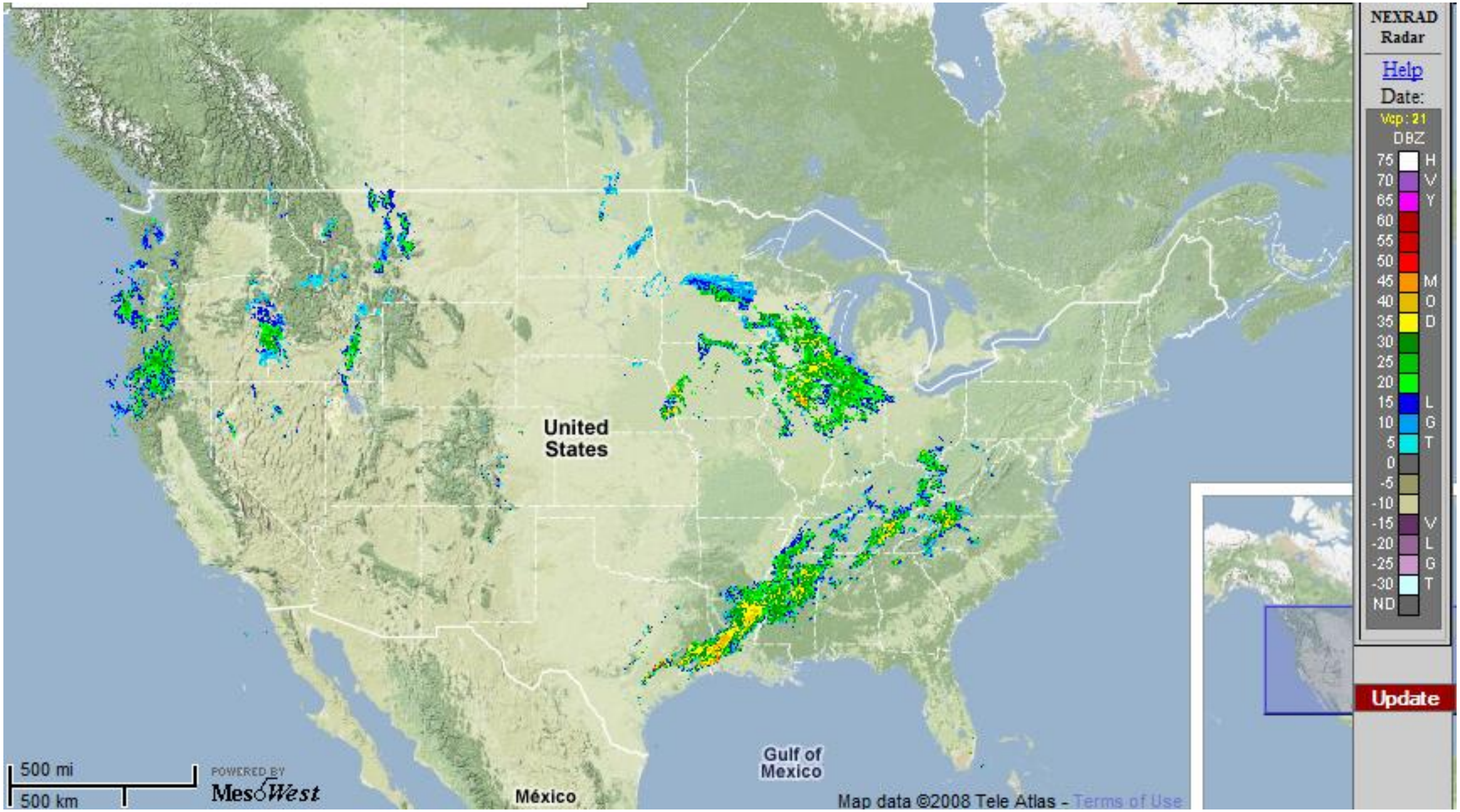
Number of Active Mesonet Stations in MesoWest



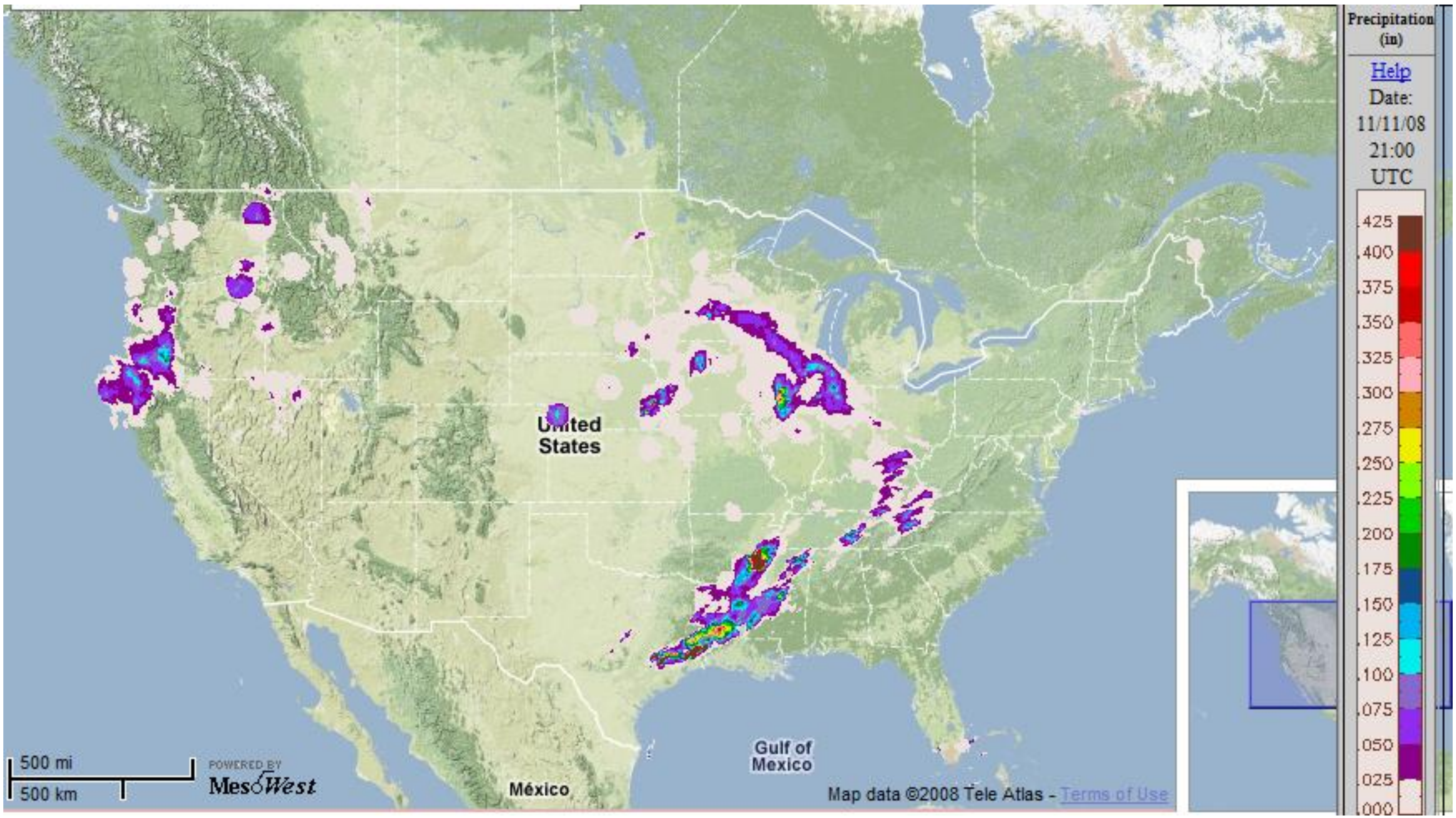
RTMA Precipitation Analysis

- NCEP Stage 2 Multisensor Precipitation Analysis on 4 km grid remapped to 5 km NDFD grid
- Gauge and Radar data only

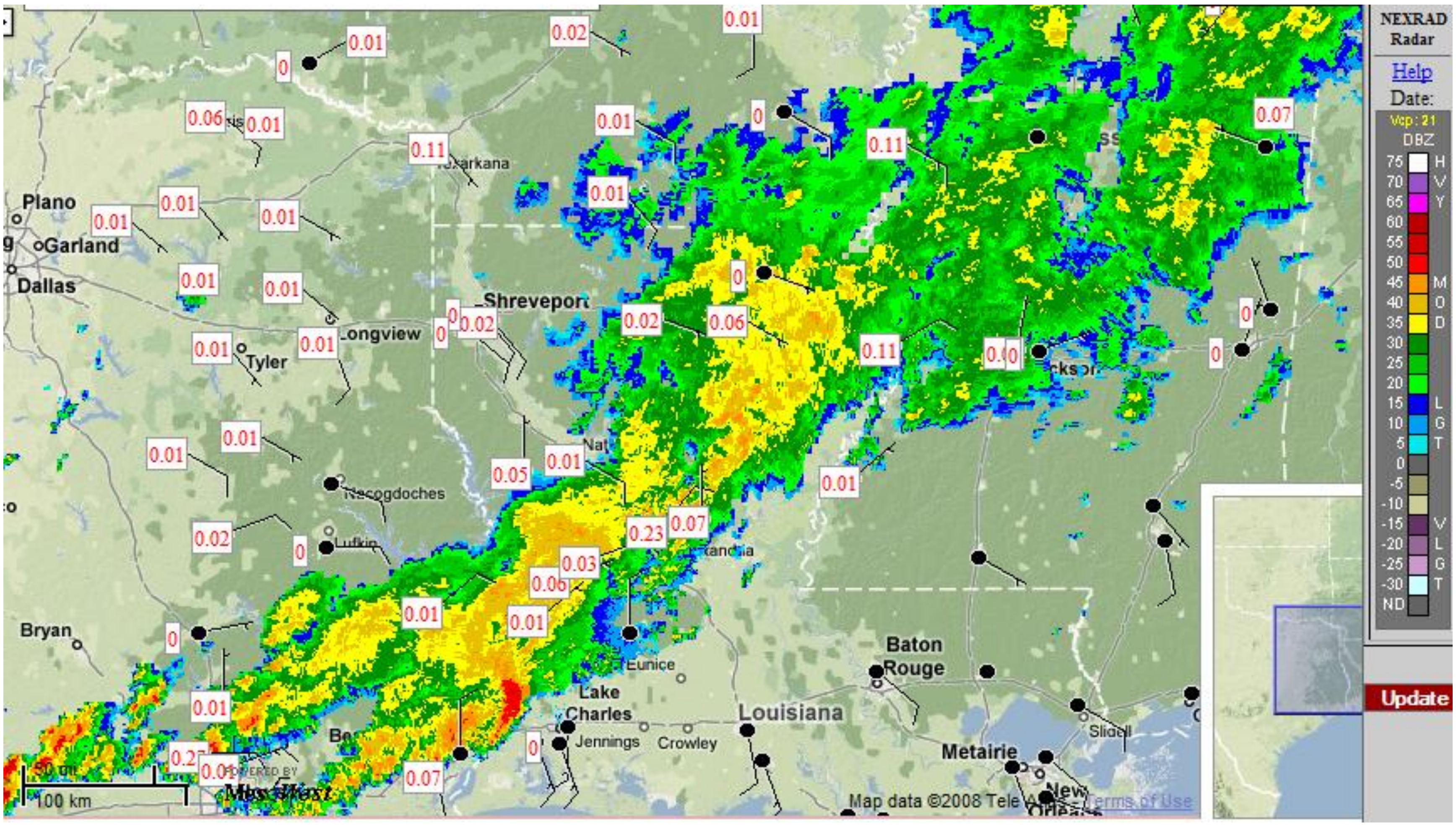
Reflectivity



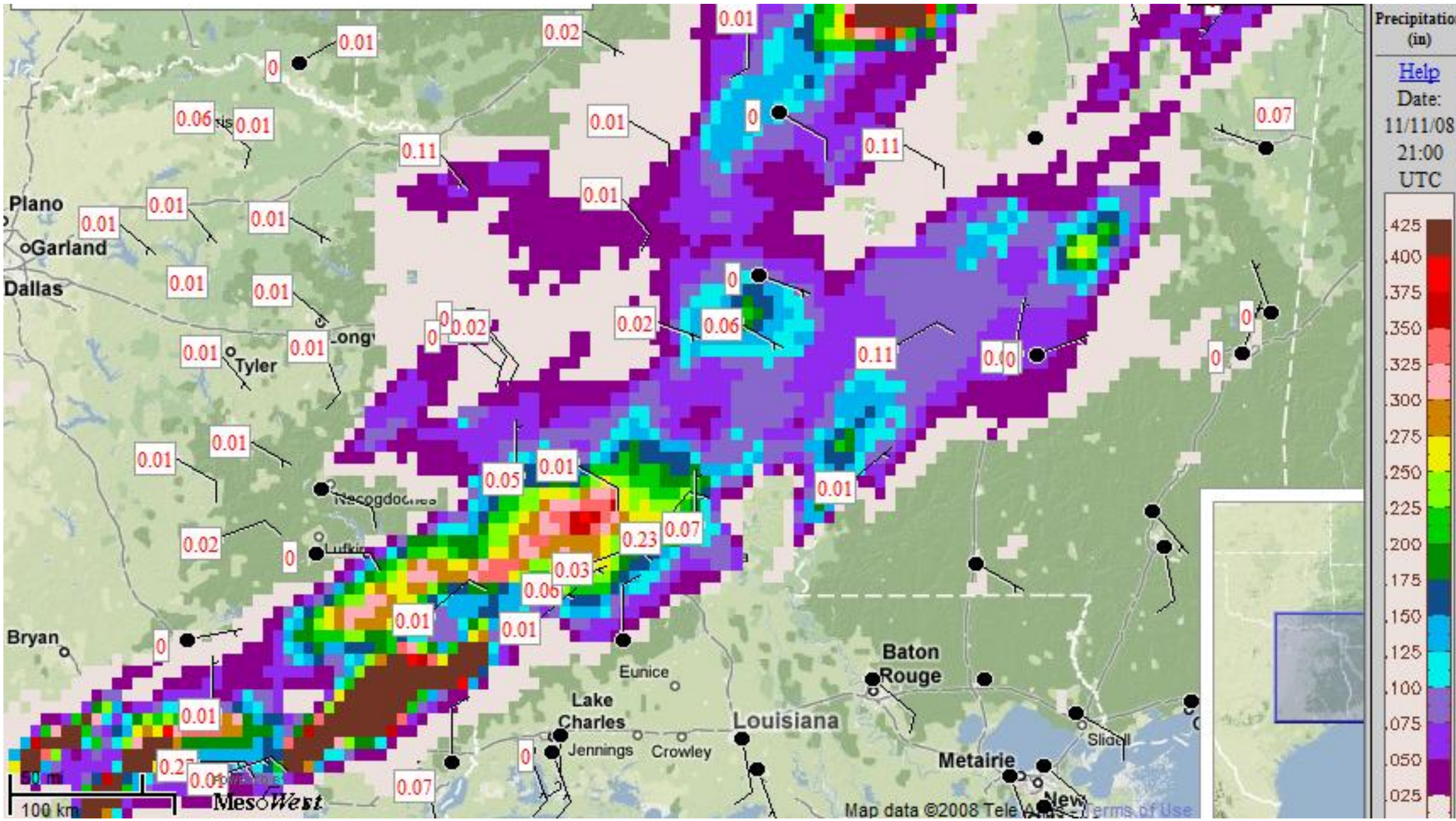
RTMA Precipitation



Gauge & Radar


NEXRAD Radar
Help
Date:
DBZ
Update
Plano
Garland
Dallas
Tyler
Longview
Shreveport
Nacogdoches
Lufkin
Bryan
Eunice
Lake Charles
Jennings
Crowley
Louisiana
Baton Rouge
Metairie
Slidell
100 km
Map data ©2008 Tele Atlas
Terms of Use

RTMA Precipitation



Observations vs. Truth?

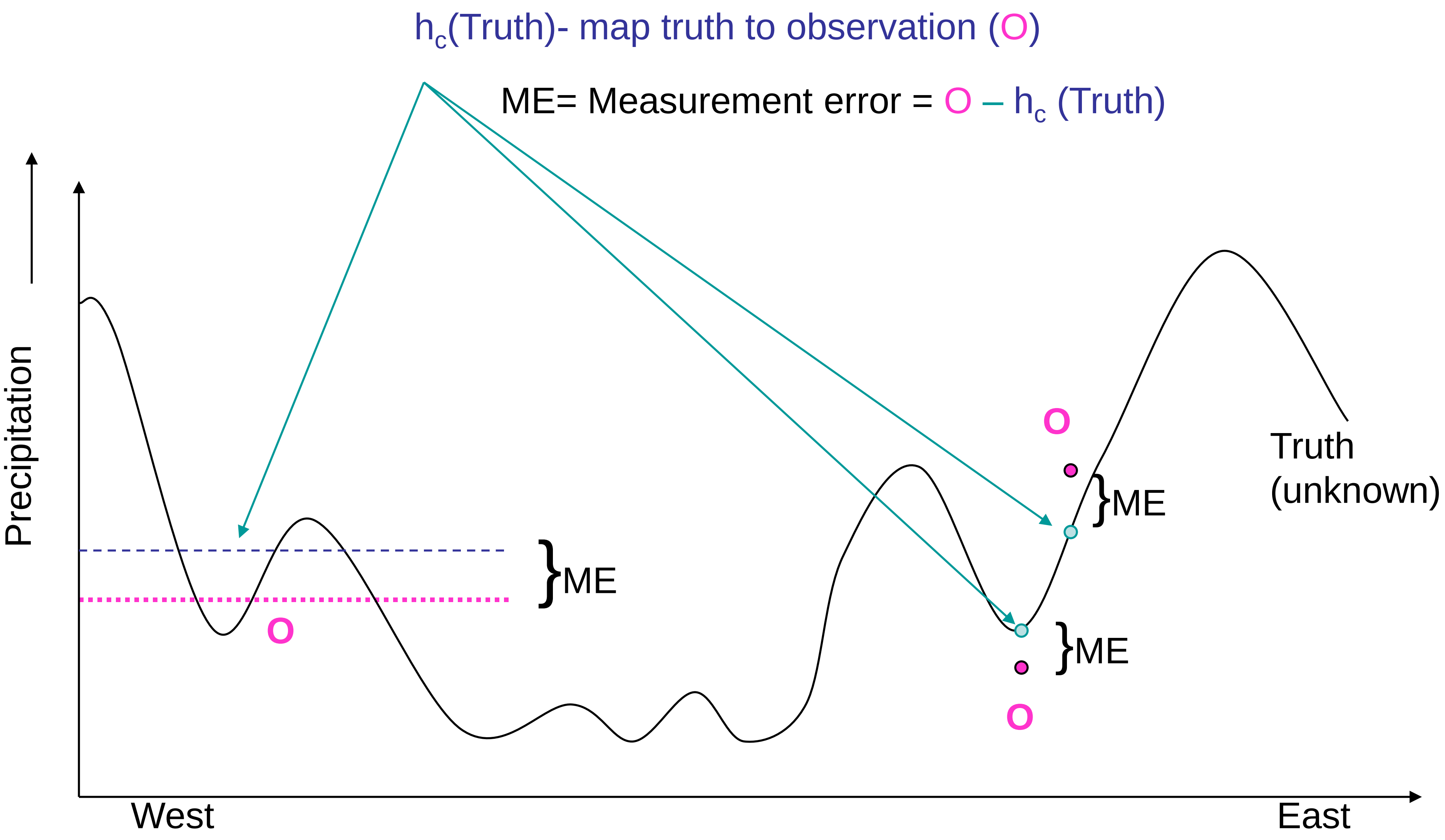
- A Few Good Men
- Truth is unknown and depends on application: “expected value for 5 x 5 km^2 area”
- Assumption: average of many unbiased observations should be same as expected value of truth
- However, accurate observations may be biased or *unrepresentative* due to any number of factors

Representativeness Errors

- Observations may be accurate…
- But the phenomena they are measuring may not be resolvable on the scale of the analysis
  - This is interpreted as an error of the *observation* not the analysis
- Common problem over complex terrain
- Also common when strong inversions
- Can happen anywhere

Sub-5km terrain variability (m)
(Myrick and Horel, WAF 2006)

Observation Errors



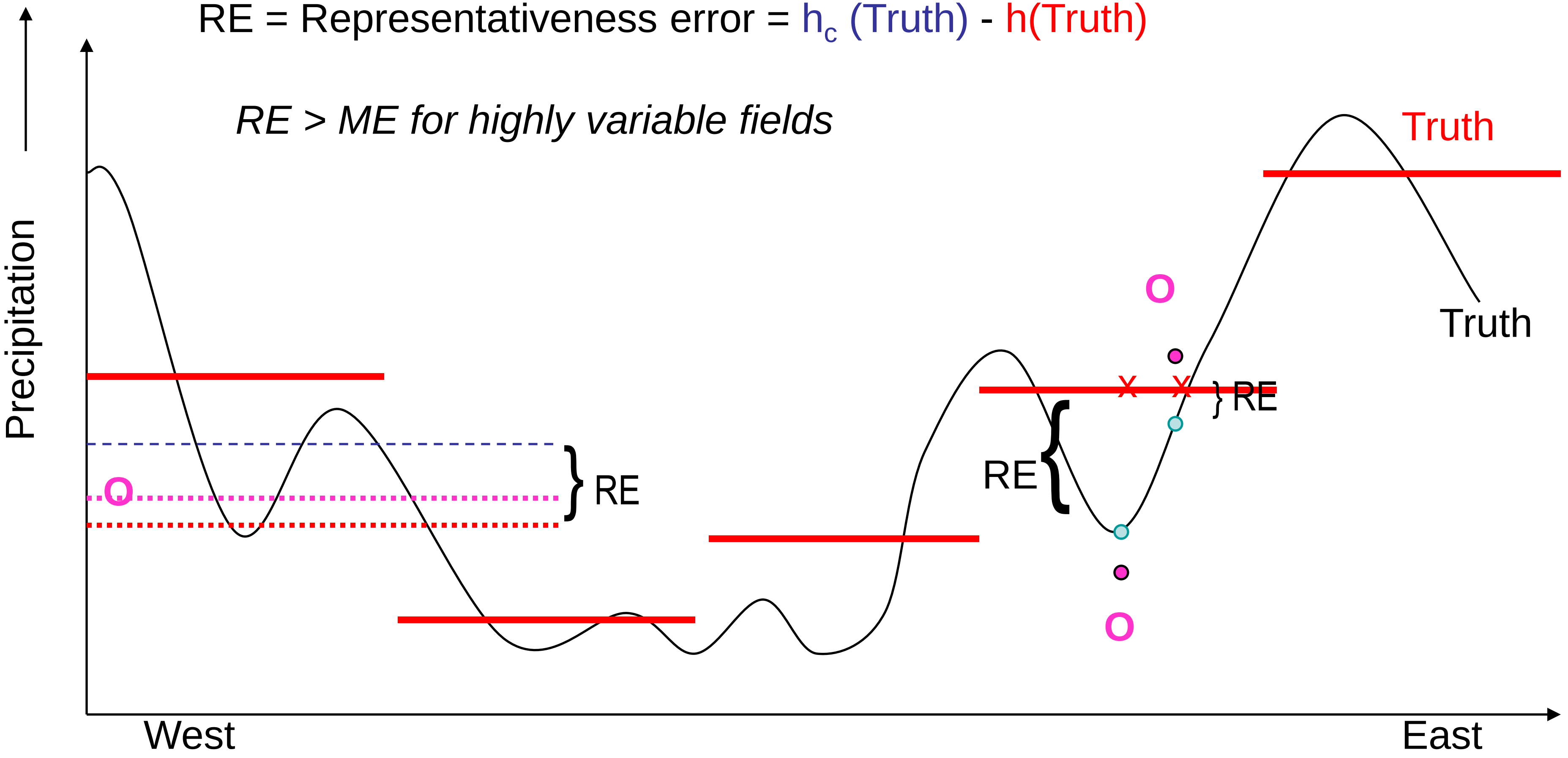
Observation Errors

Truth= H(Truth)- maps truth to scale of analysis grid

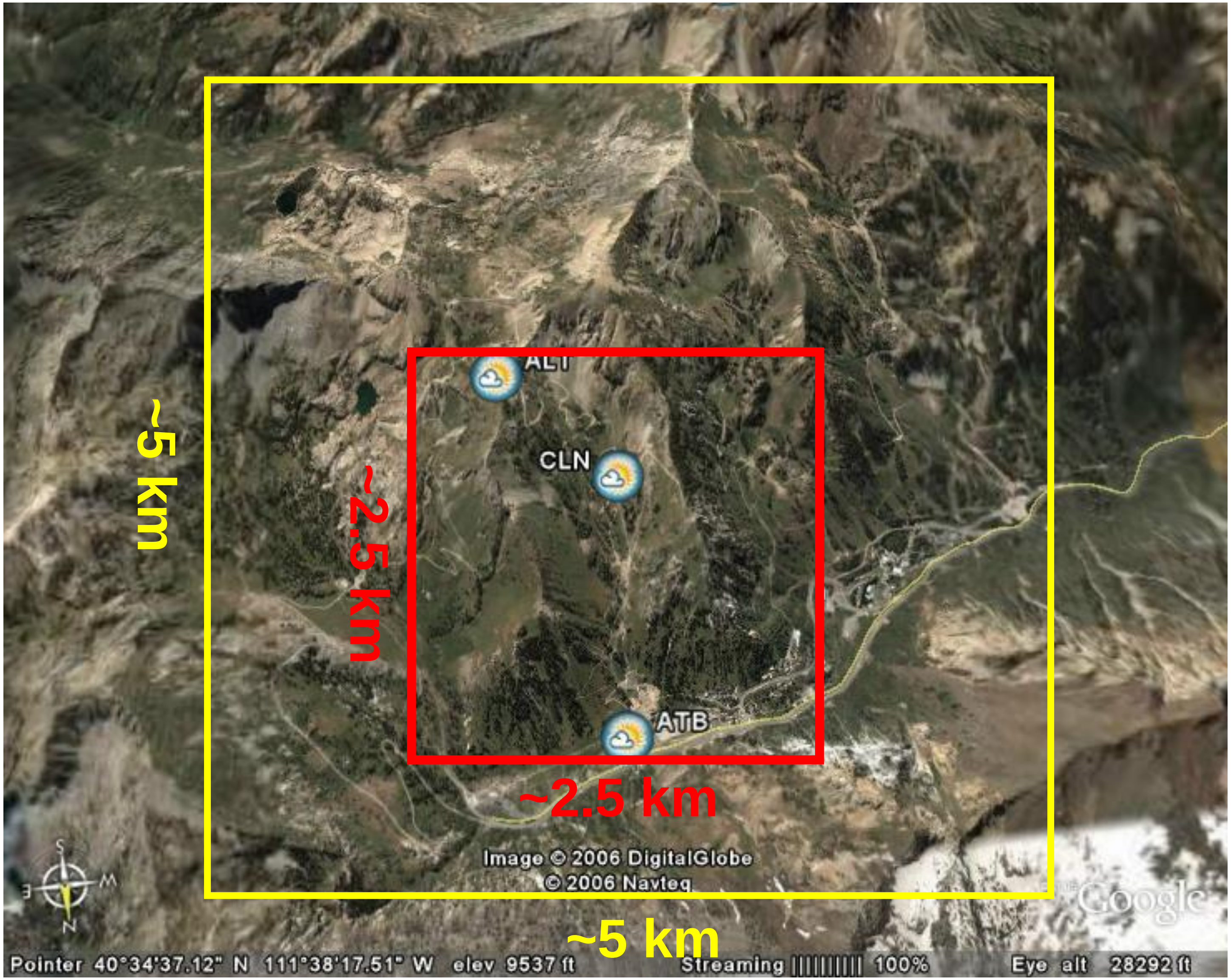
h(Truth)- maps Truth to observation

RE = Representativeness error = h_c (Truth) - h(Truth)

RE > ME for highly variable fields



Representative errors to be expected in mountains
Alta Ski Area

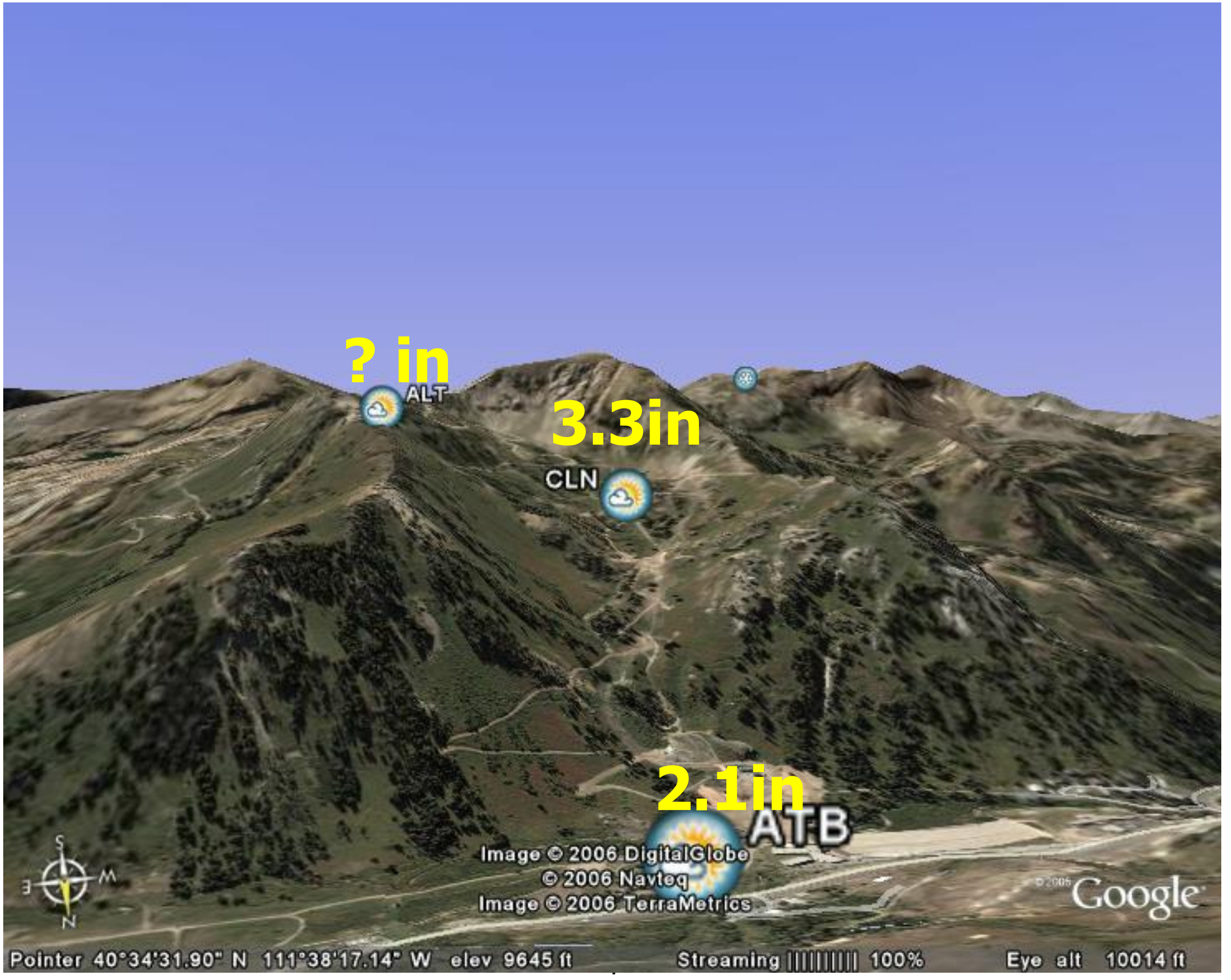


Alta Ski Area

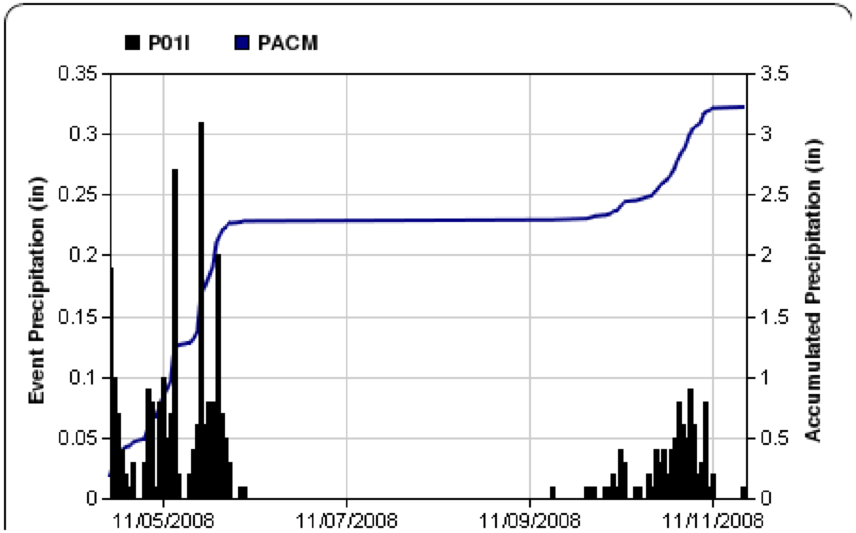
Looking up the mountain

Looking up Little Cottonwood Canyon

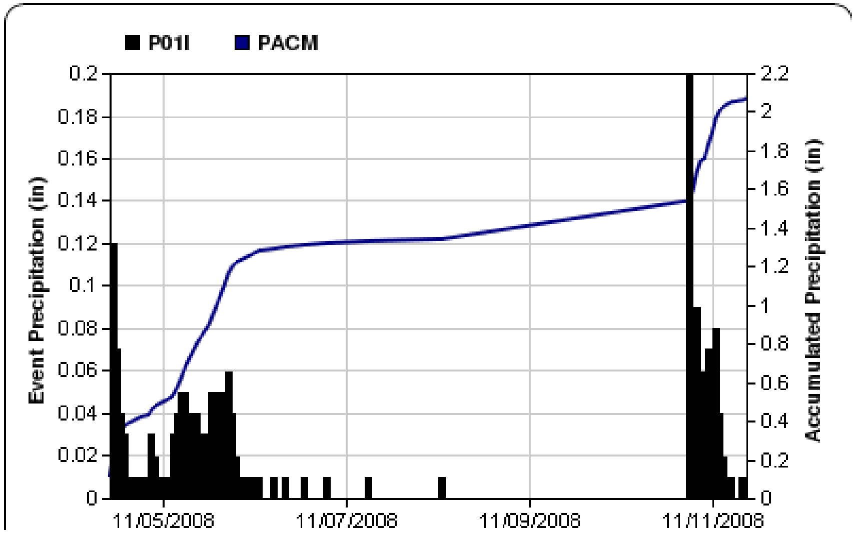
Alta Ski Area



Precipitation within grid box



CLN



ATB

Key Points

- Assuming an observation is “truth” may seem simpler if you have only that one observation
- Magnitudes of observational errors are only a piece of the puzzle
  - Analyses assume observational errors at one location are independent of errors at another
  - Observational biases (equipment, siting, etc.) especially during specific synoptic conditions (light winds, cold pools) can contribute to correlations between observational errors
- Verification procedures need to incorporate uncertainty information about the observational assets
  - Don’t sweat the small stuff
  - ASOS observations are far from perfect
  - Monitor error characteristics of observations over space and time