Using the Composite and Neighborhood Methods as Verification Tools

Jason E. Nachamkin*

Spatial Forecast Verification Methods Intercomparison Project, Fall 2009 Workshop

24-25 August 2009

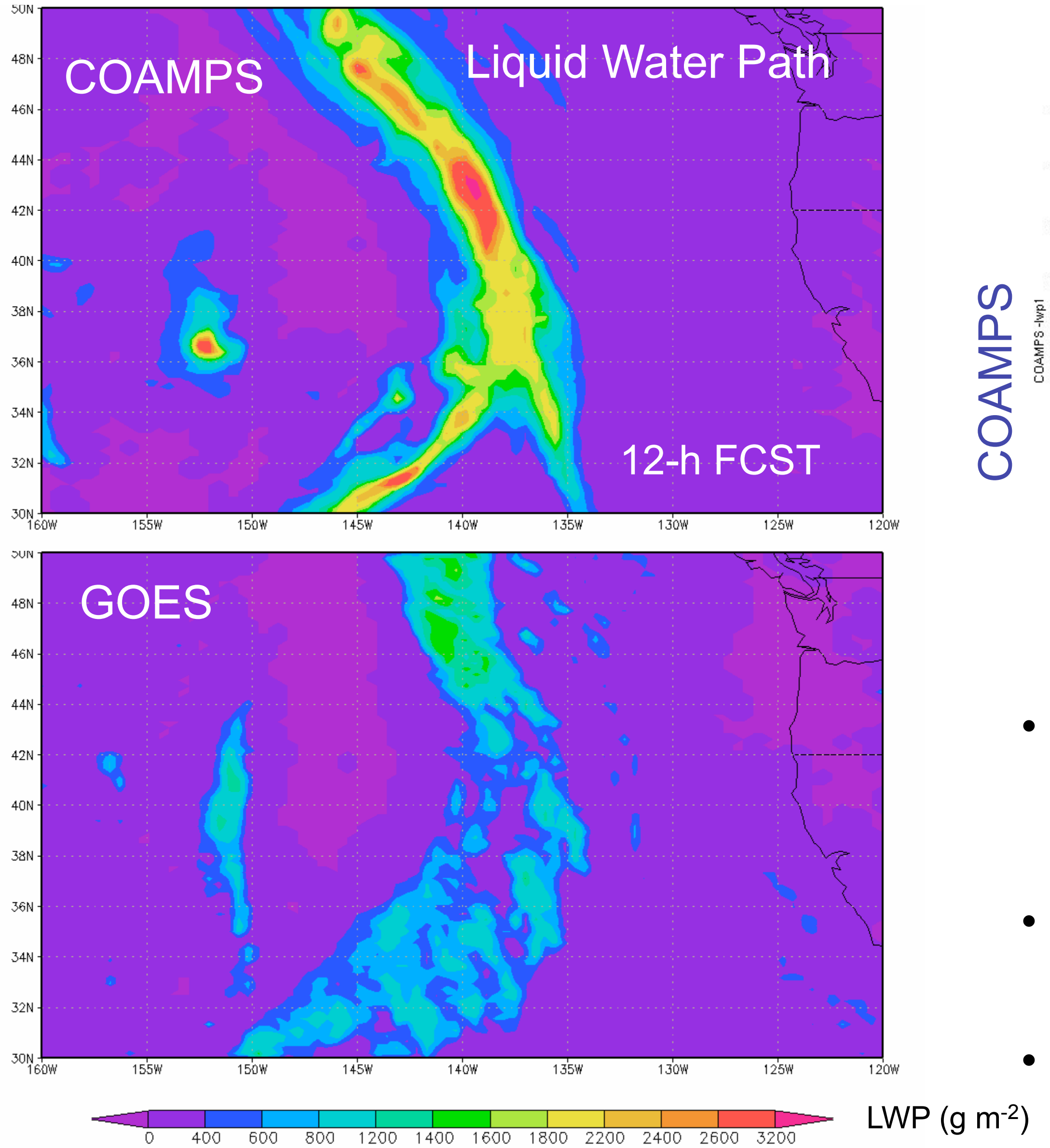
Boulder, Colorado

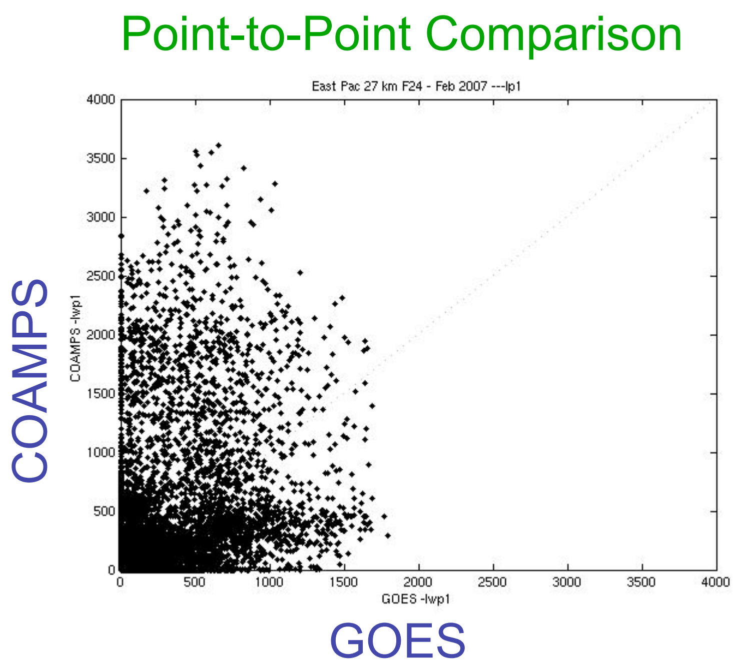
*Naval Research Laboratory, Monterey, California

Outline

- Quantifying forecast value
- Composite Method
  - Description/examples
  - Advantages/disadvantages
- Neighborhood Method
  - Description/modification/examples
  - Advantages/disadvantages
- Complementary Usage
- Summary

Is this a good Forecast?



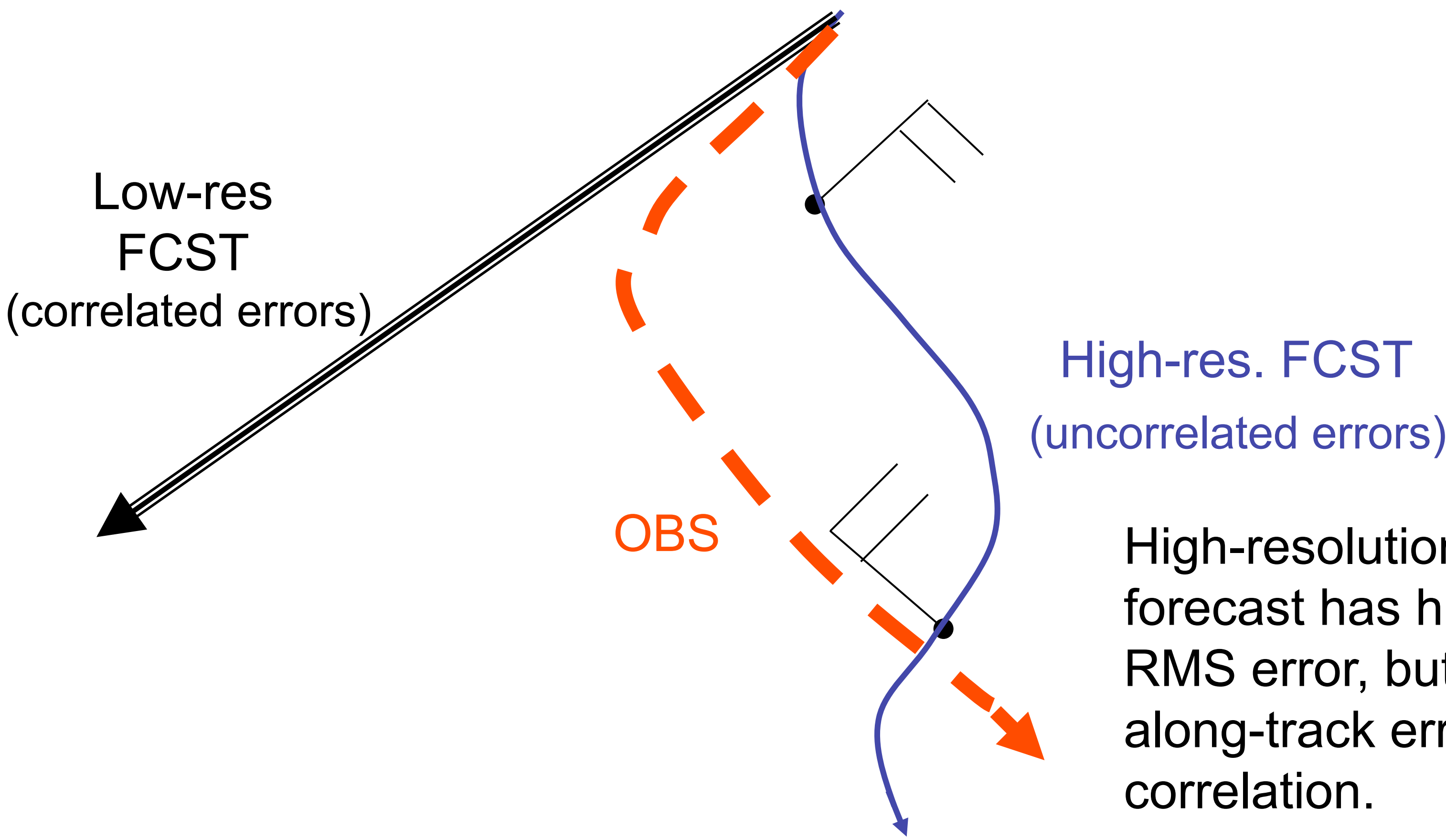


- Increased resolution increases variance and reduces spatial error correlation.
- Good forecasts may have large errors at small scales.
- How to quantify this?

Why Spatial Correlations Matter

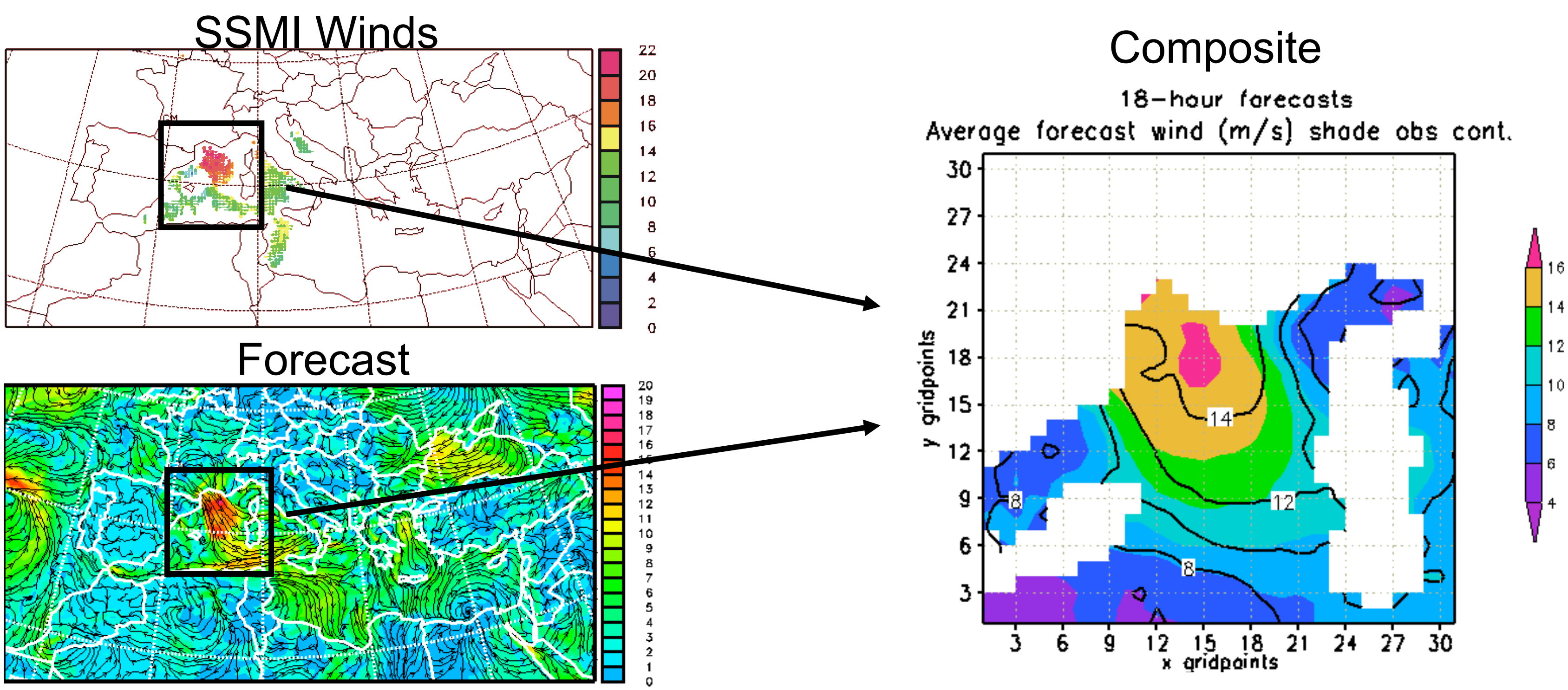
Contaminant dispersion forecasts

A Tale of Three Trajectories…



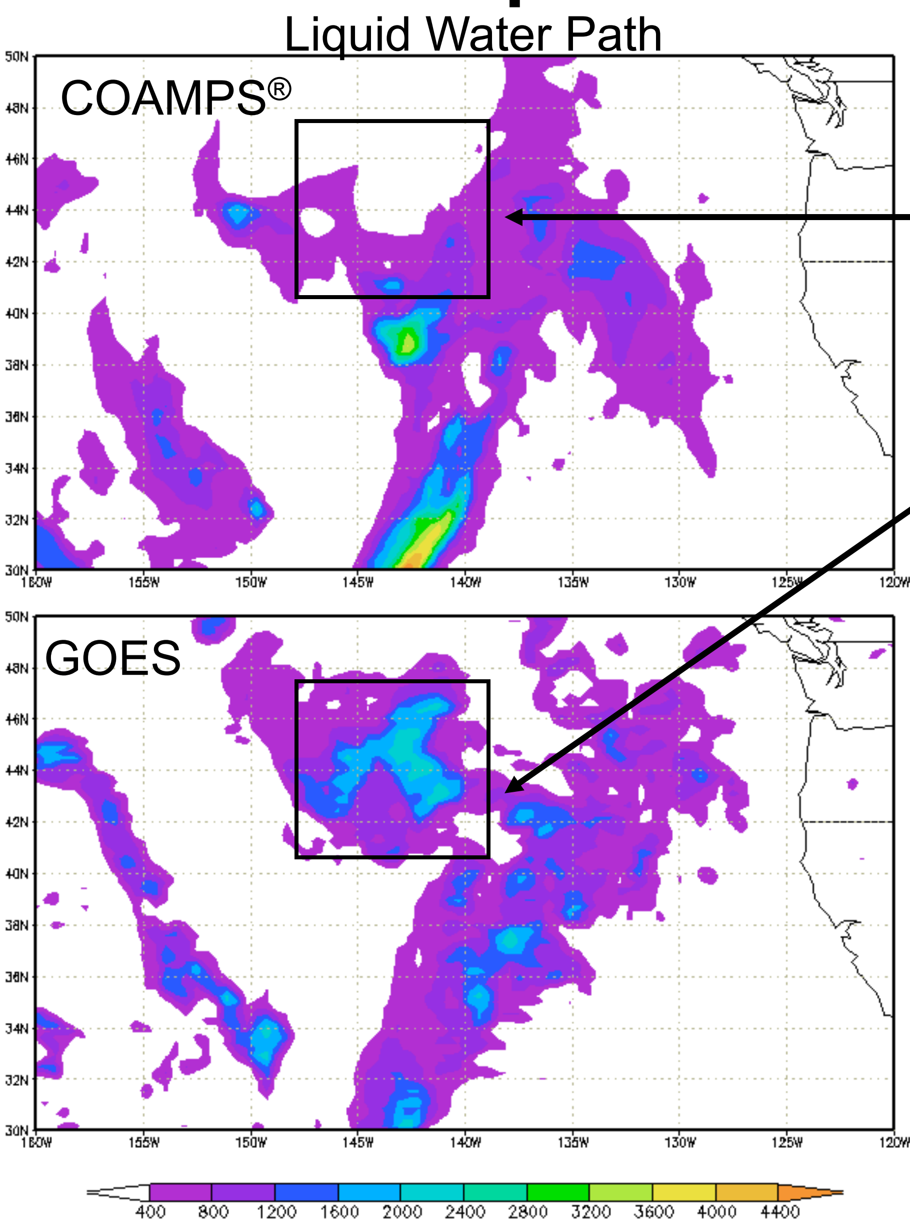
High-resolution forecast has higher RMS error, but less along-track error correlation.

Why Composite?



- Incomplete observations
  - Direct, unfiltered data
  - Avoid pitfalls of matching
- Probabilistic framework

Composite Verification



Event Statistics

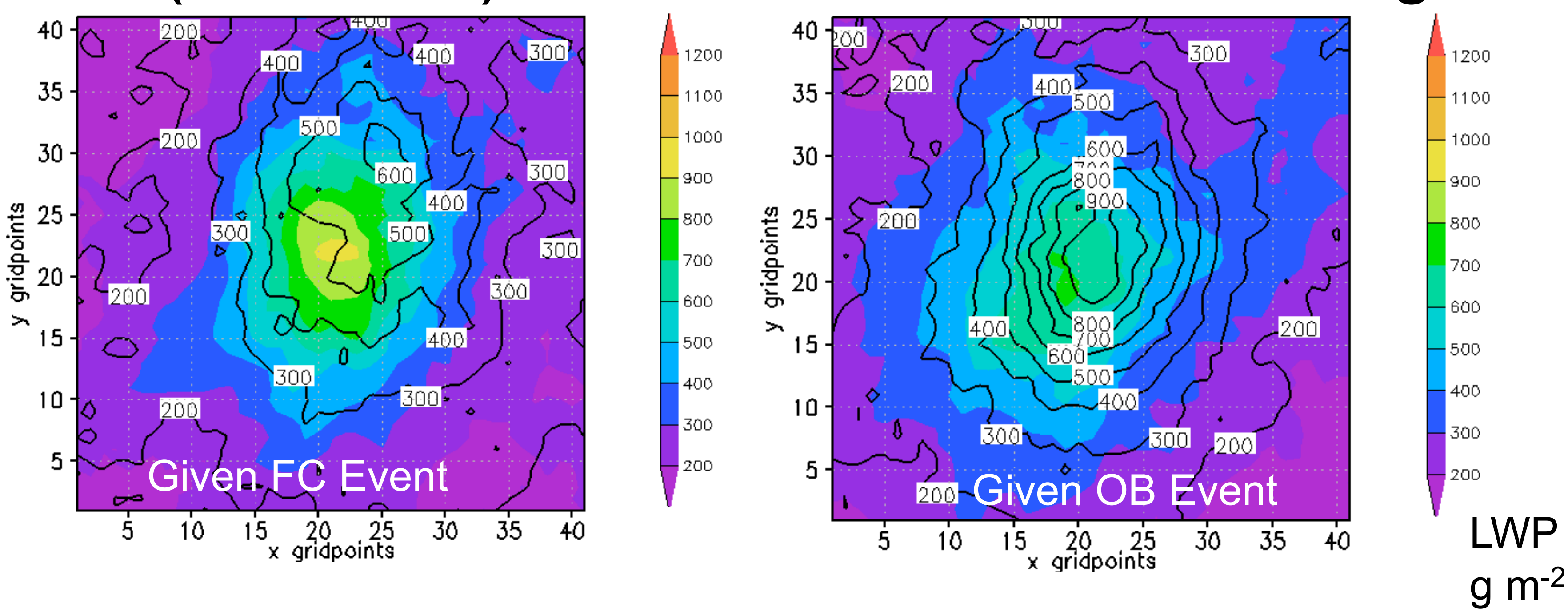
- Collect samples of multiple events of similar scale
- Investigate systematic forecast errors

Method:

- Identify events of interest
  - all events with LWP ≥ 500 g m^{-2}
  - 100-600; 600-3000 points
- Composite all predicted events
- Composite all observed events

Composite Verification

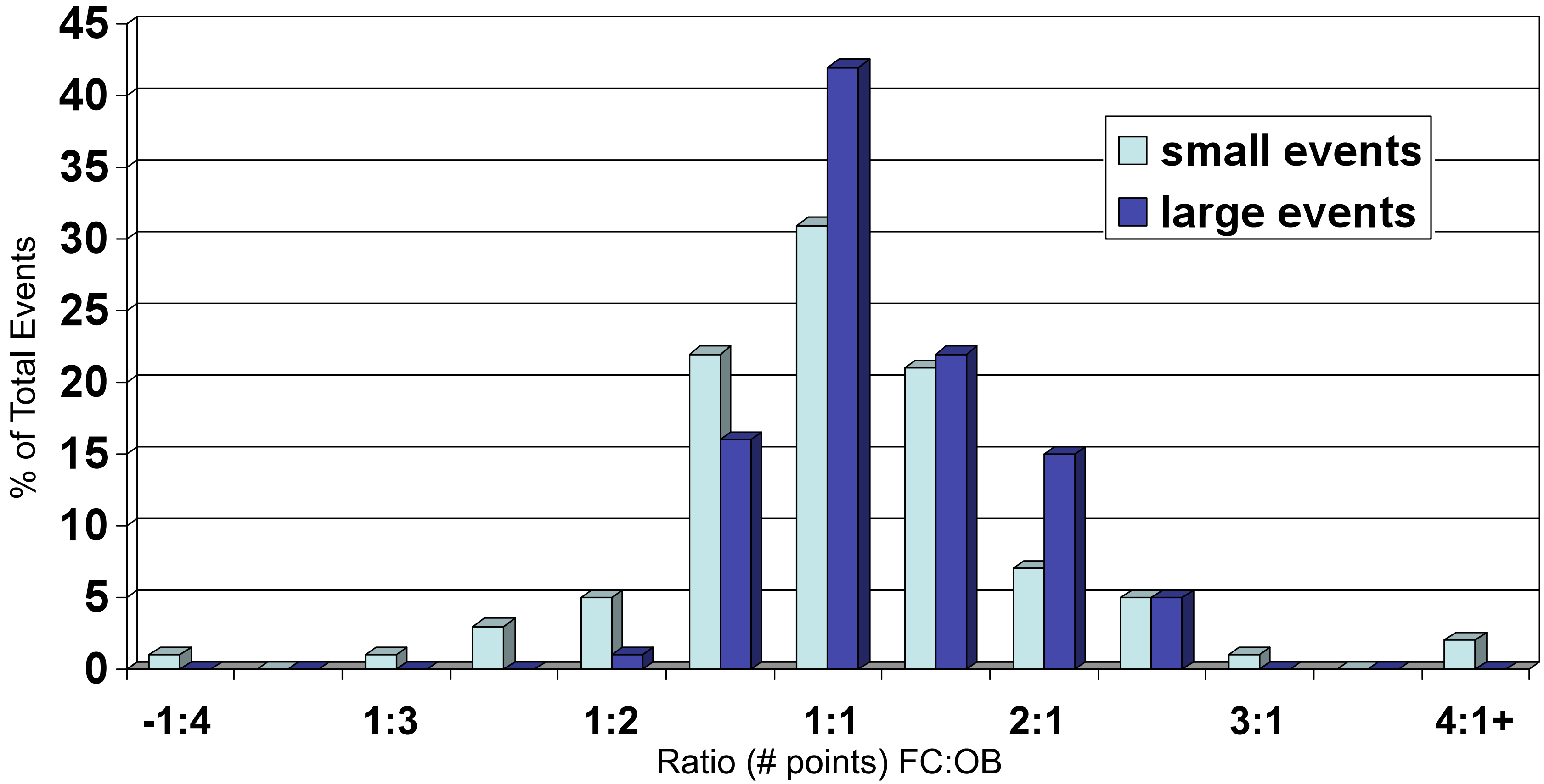
Small (~350 km) Cloud Events LWP ≥ 500 g m^{-2}



- Spatial phase errors revealed by overlaying distributions
- Must display predicted and observed events separately

Percentage of Events with Given Fcst:Obs Ratios

Feb-May 2007



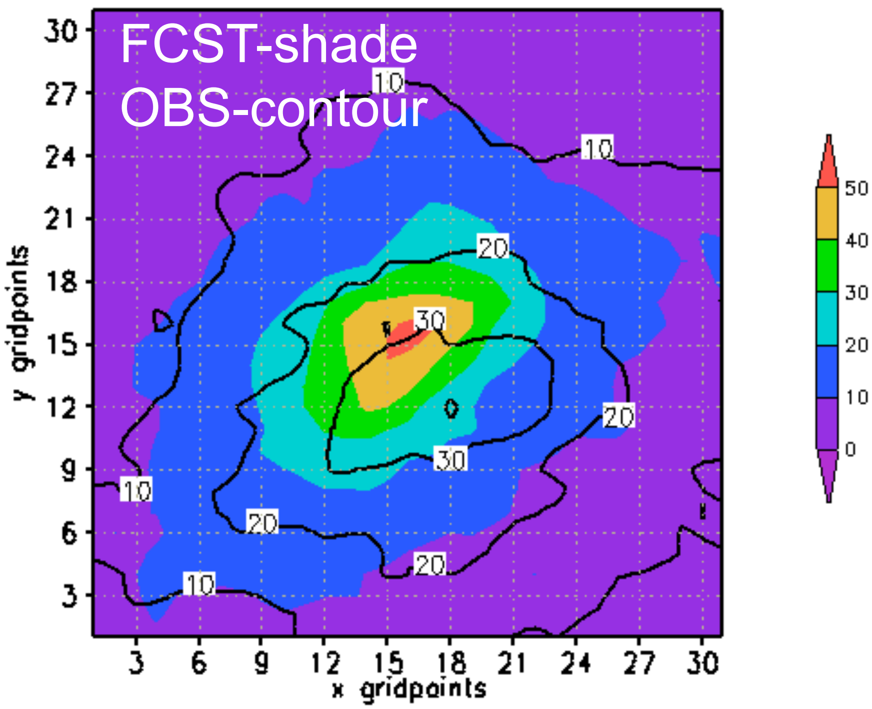
Under-forecasts (too few deep clouds)

Hits (Fcst similar to obs) 80% LG; 74% SM

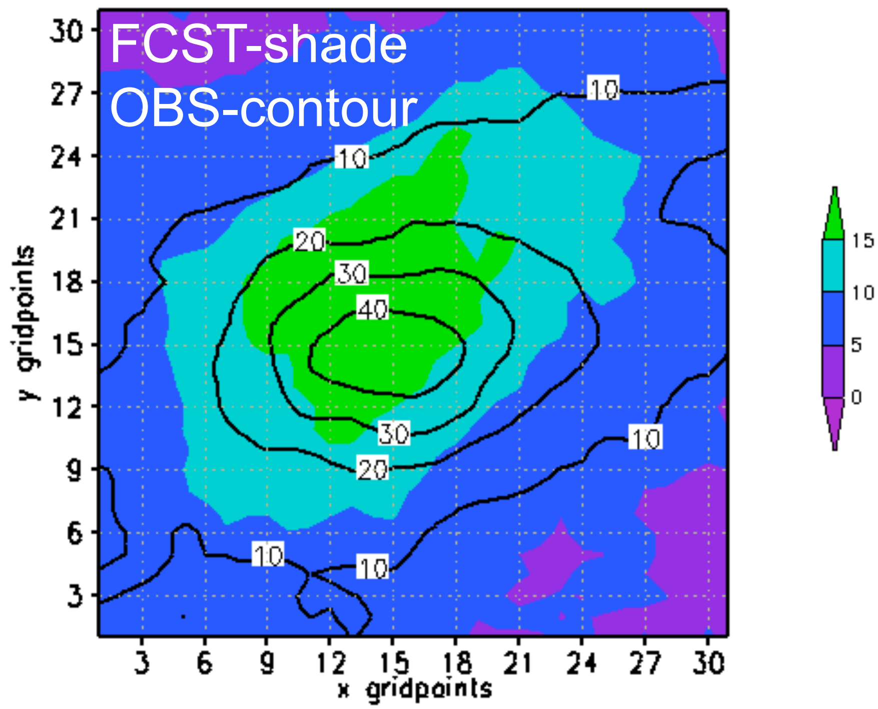
Over-forecasts (too many deep clouds)

Heavy Precipitation Composites

Average rain (mm) given an event was predicted



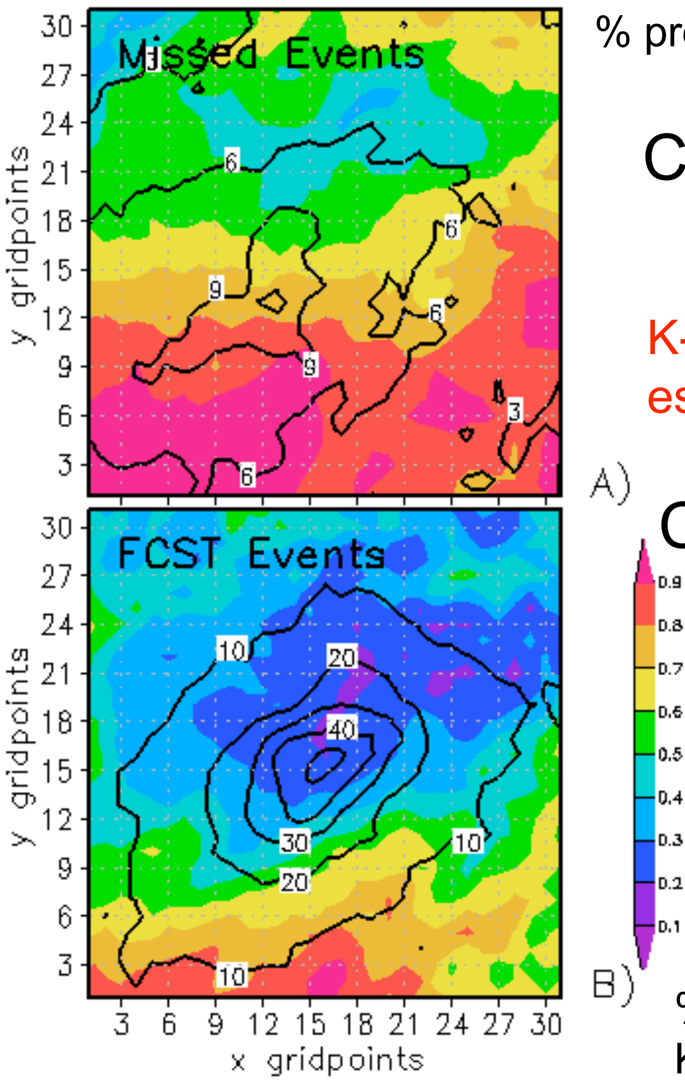
Average rain (mm) given an event was observed



- Composites behave differently for observed events vs. predicted events.
- False alarms and missed forecasts are associated with different errors.

Diagnostic Statistics

Ave. COAMPS precipitation contoured
% precip from Kain-Fritsch scheme shaded



Composite of all missed forecasts

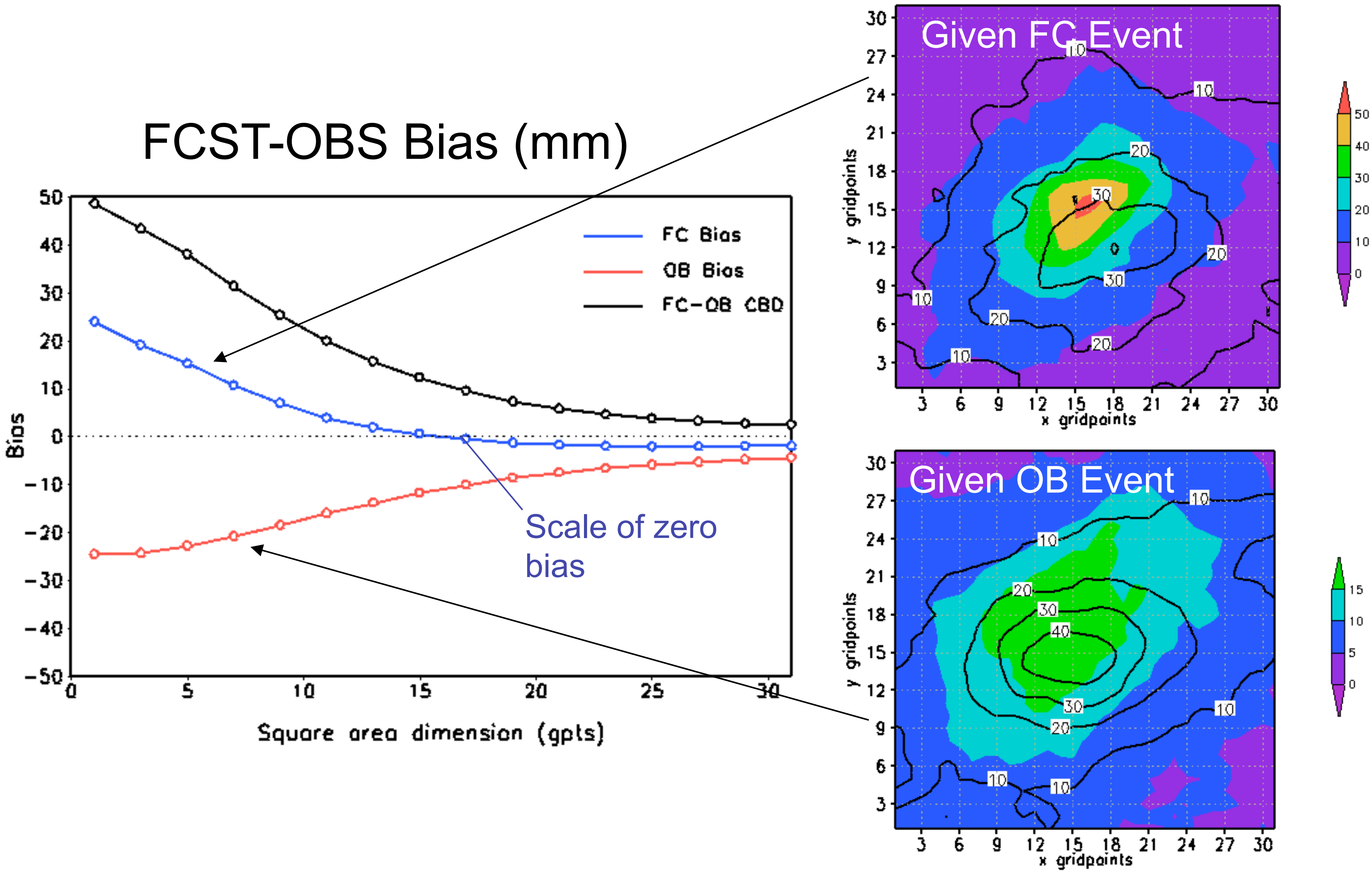
K-F does not produce enough precip, especially in the warm sector

Composite of all forecasts

Well predicted events contain more explicitly resolved precipitation in the northern portions.

% precip from Kain-Fritsch

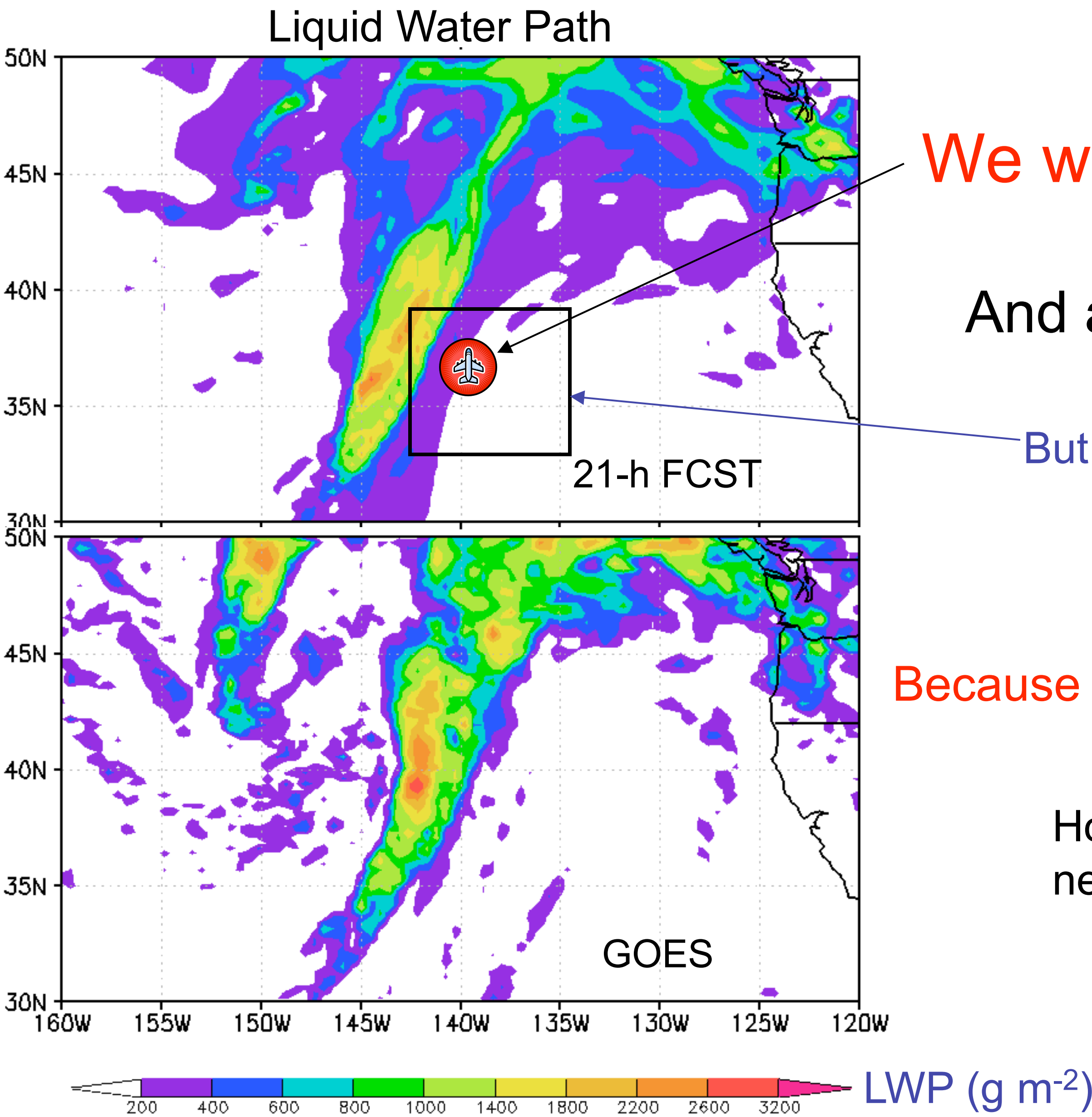
Conditional Bias Difference as a Summary Score



Composite Method Recap

- Strengths
  - Composite statistics are easily viewed
  - Works with limited data
  - Results can be databased
  - No dependence on matching
- Weaknesses
  - Not good for large, complex events (clouds)
  - No deterministic shape/rotation information
  - Observed and predicted events sampled separately
  - Scores have limited applicability

Quantifying Uncertainty



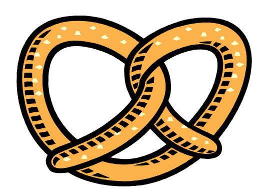
We want to go here

And avoid the front

But we're flexible…

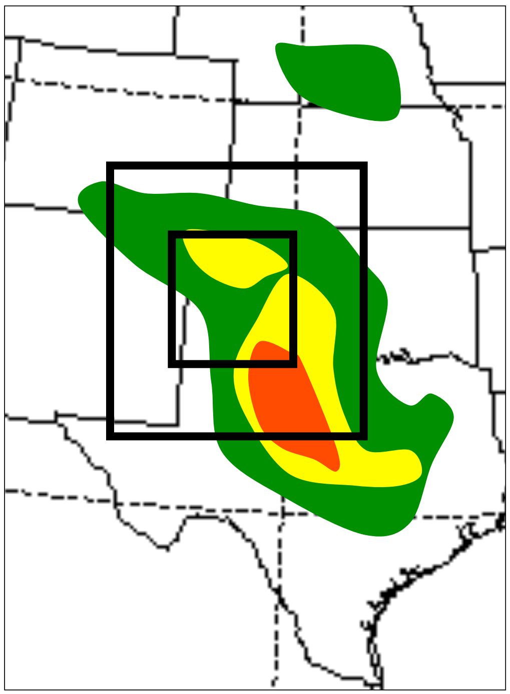
Because the forecast is not perfect

How flexible do we need to be?

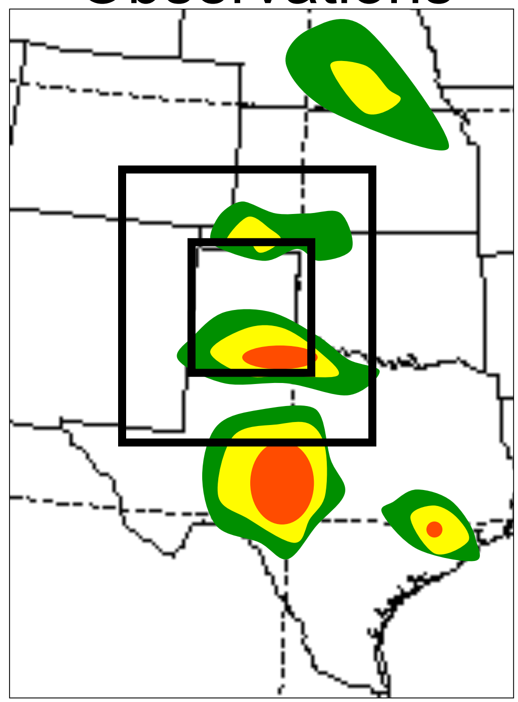


Fuzzy Neighborhood Method

Forecast



Observations

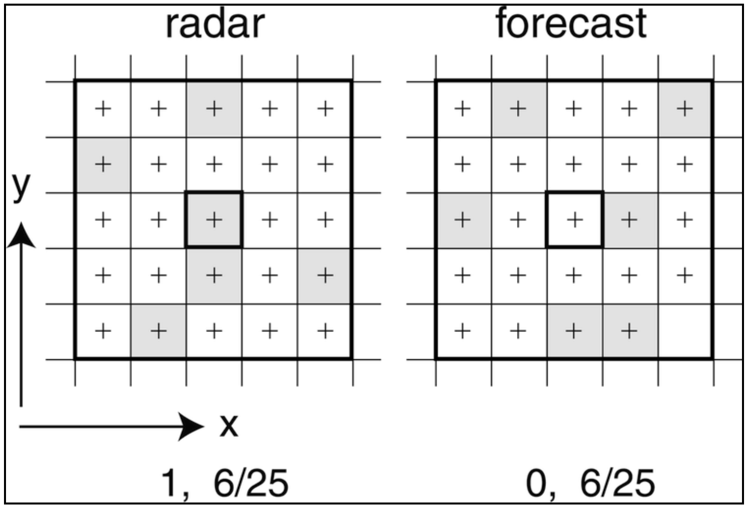


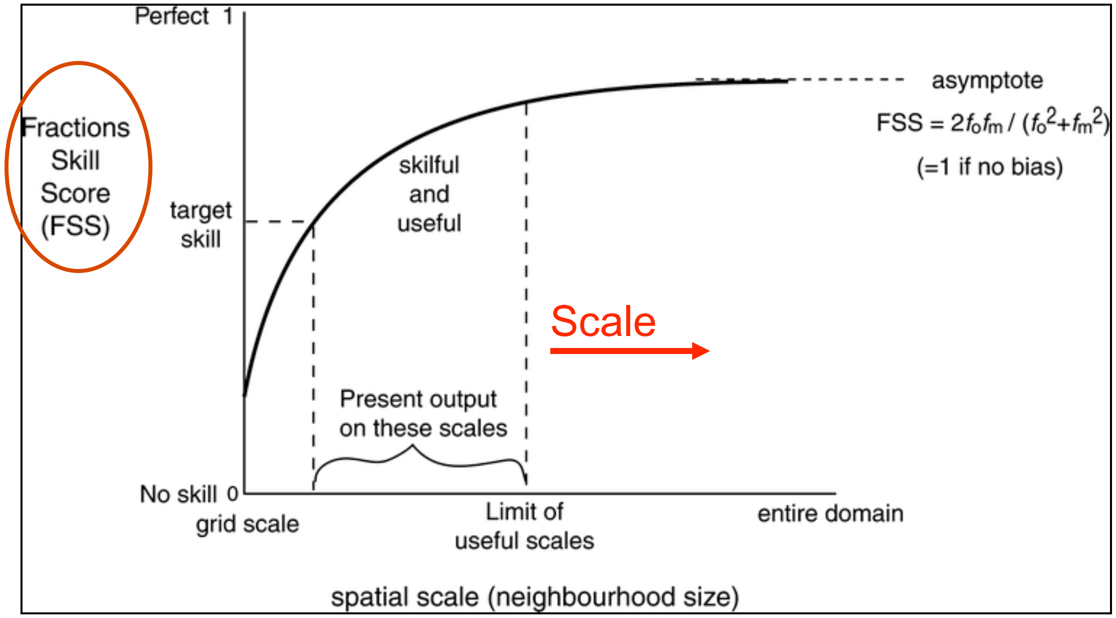
Roberts and Lean (2008)

- Compute a skill score over an increasing range of scales.
- Accuracy increases with scale, precision decreases with scale.
- Collect samples at every grid point, not just events.
- “Scale” is directly associated with the sampling area.
- For composites, “scale” is influenced by the events in the sample.

Fuzzy Neighborhood Method

From Roberts and Lean (2008)

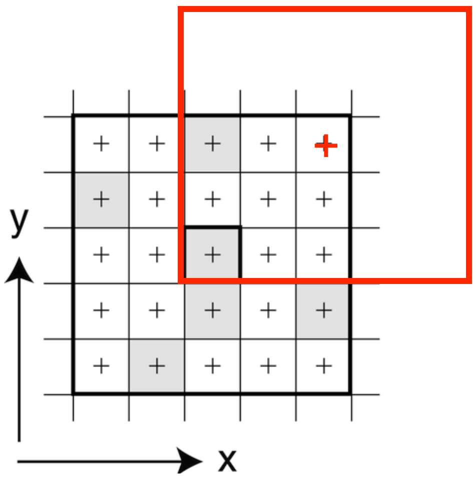




- Use threshold to create a binary field.
- Calculate Fractions Skill Score.
- FSS=1 (perfect) when forecast coverage=obs coverage.

Fuzzy Neighborhood Method

A Few Caveats



$$FSS_{(n)} = 1 - \frac{MSE_{(n)}}{MSE_{(n)ref}}$$

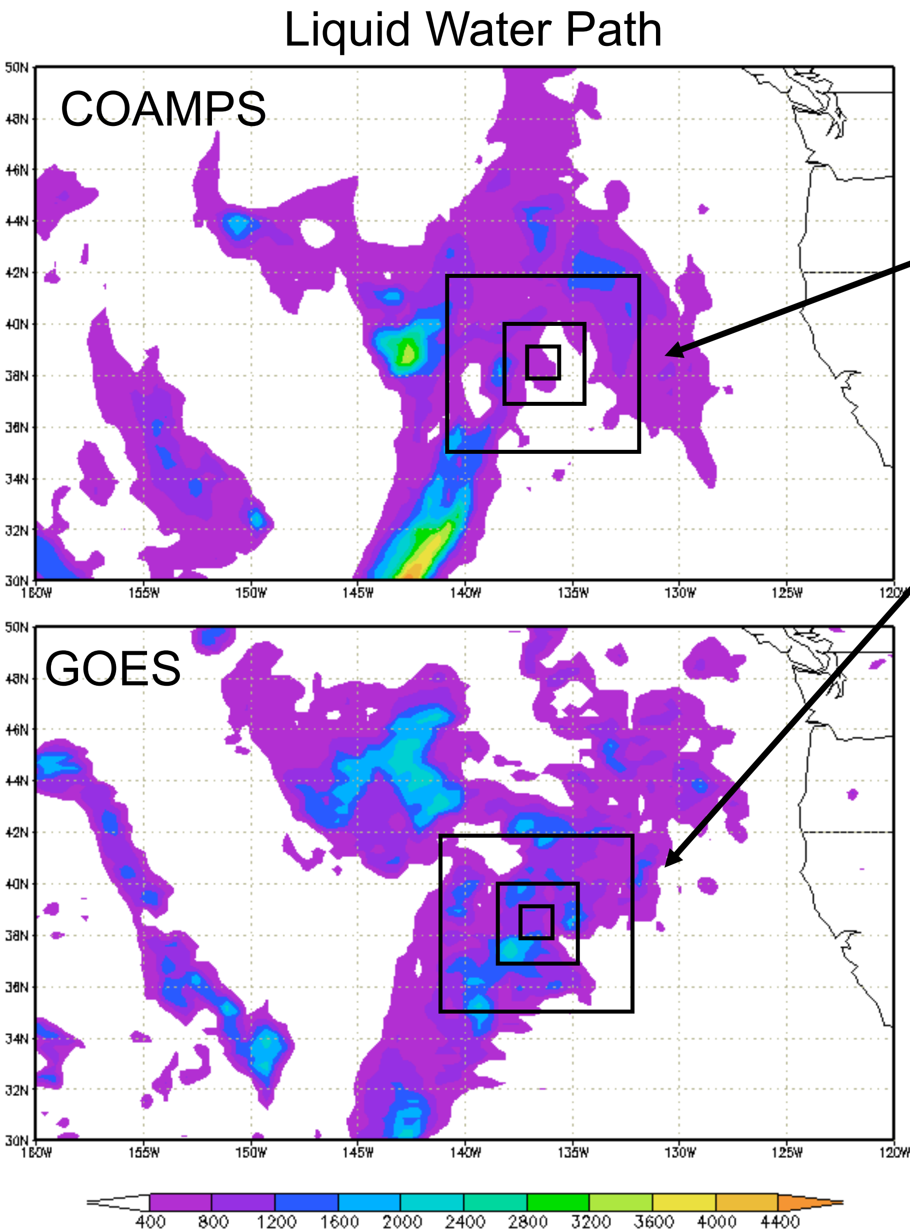
$$MSE_{(n)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (O_{(n)i,j} - M_{(n)i,j})^2$$

$$MSE_{(n)ref} = \frac{1}{N_x N_y} \left[\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} O_{(n)i,j}^2 + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} M_{(n)i,j}^2 \right]$$

- Large scale samples exceed grid bounds
- May cause aliasing

- Reference MSE not a true climatology
- Changes with each forecast
- Positive bias leads to large MSE_{ref} and improved FSS

Calculating Scale-Dependent Scores



Fuzzy Statistics

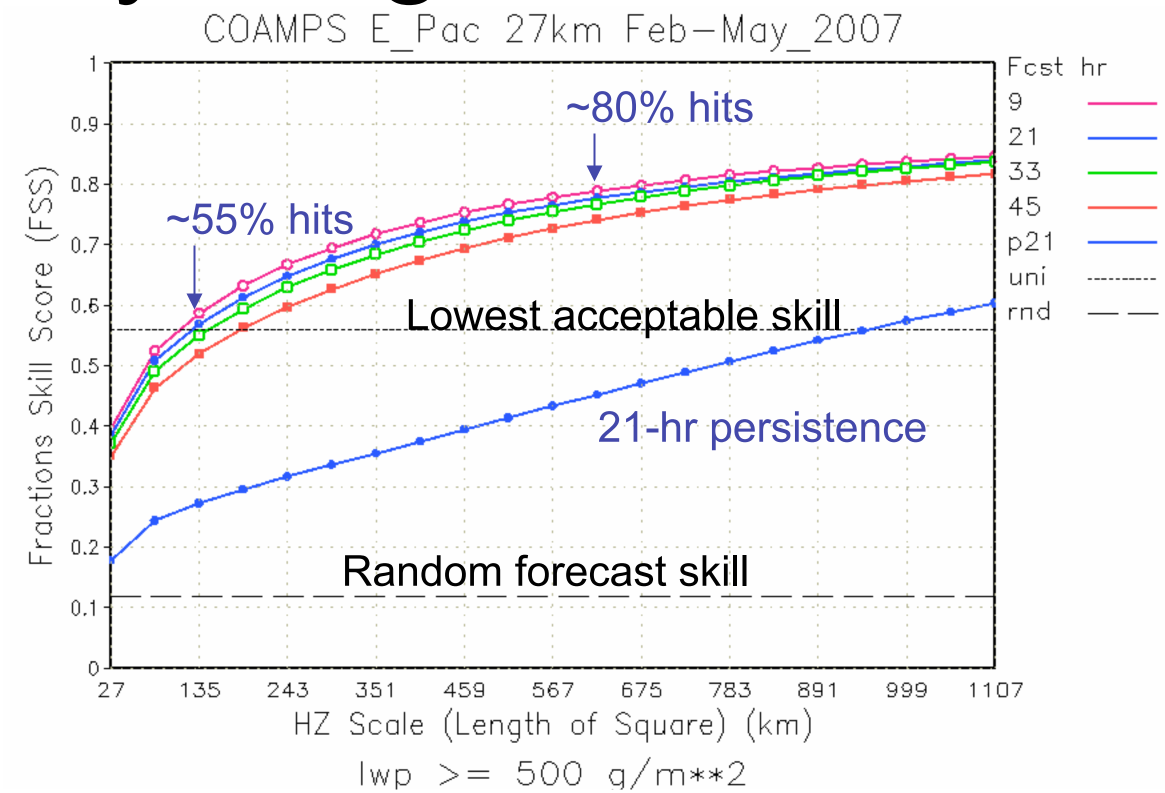
- Collect samples at all points at multiple scales
- Determine the smallest scale to possess an acceptable error

Method:

- Identify a threshold of interest
  - LWP ≥ 500 g m^{-2}
- Create binary field based on the threshold
- Examine the observed and predicted fractional coverage at each scale

Roberts and Lean (2008)

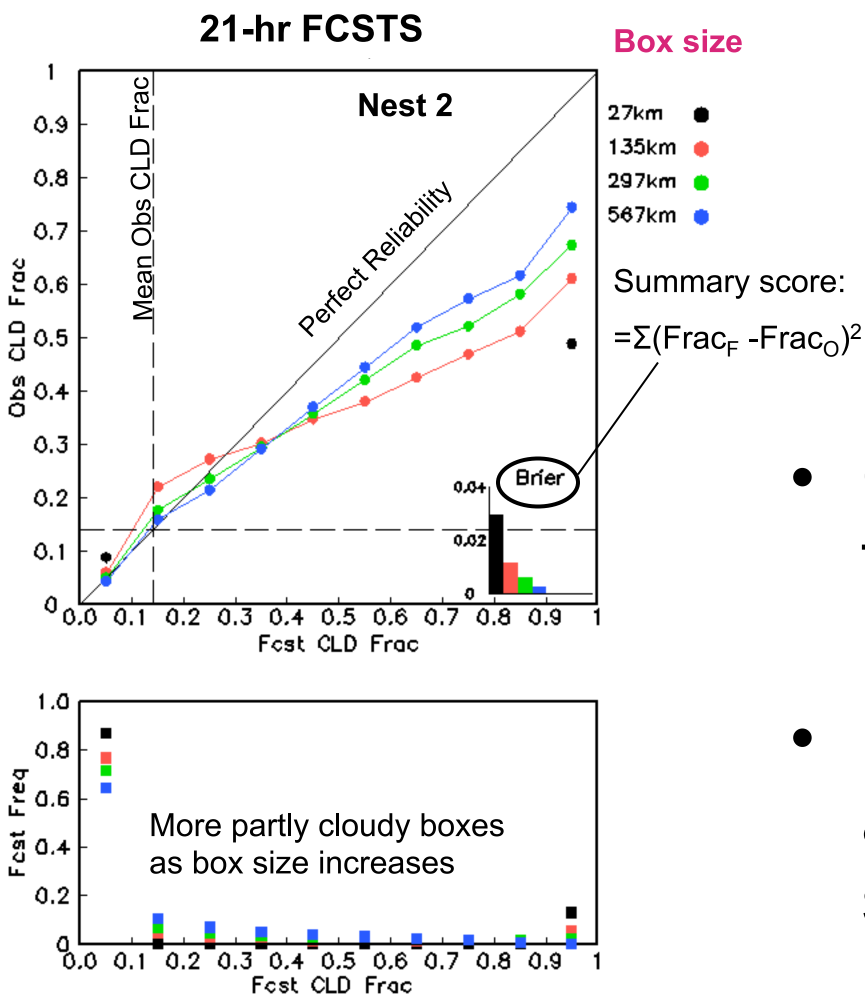
Fuzzy Neighborhood Method



- Forecast uncertainty is a function of sampling area.
- Point forecasts are often wrong.

Forecast Uncertainty

Scale-Based Statistics



Results are best plotted using reliability diagrams.

- Compute average cloud fractions over increasingly large areas.
- Forecasts for large areas are more reliable but less spatially certain.

Neighborhood Method Recap

- Strengths
  - Simple (much like composite method)
  - Summarizes grid-total performance in one score
  - Good for large, complex entities
  - Skill implicit in the scores
- Weaknesses
  - No phase error information
  - Initially difficult to gauge

Summary

- The characterization of mesoscale uncertainty is a challenging problem.
- Traditional methods do not account for spatial error correlation.
  - Good forecasts can have large errors.
  - Low spatial error correlations are desired.
- Composites provide useful performance information.
  - Focused on specific events.
  - Hard to characterize the entire forecast.
- Fuzzy verification provides useful statistical information.
  - Entire forecast solution is characterized.
  - Uncertainty is quantified over a range of scales.
  - Threshold value is required.