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
Approved for release; distribution is unlimited
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?

COAMPS

Liquid Water Path

12-h FCST

GOES

Point-to-Point Comparison

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

LWP (g m⁻²)
Why Spatial Correlations Matter

Contaminant dispersion forecasts

A Tale of Three Trajectories...

Low-res FCST (correlated errors)

High-res. FCST (uncorrelated errors)

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

COAMPS®

Liquid Water Path

GOES
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

<table>
<thead>
<tr>
<th>Ratio (FC:OB)</th>
<th>Percentage of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1:4</td>
<td></td>
</tr>
<tr>
<td>1:3</td>
<td></td>
</tr>
<tr>
<td>1:2</td>
<td></td>
</tr>
<tr>
<td>1:1</td>
<td></td>
</tr>
<tr>
<td>2:1</td>
<td></td>
</tr>
<tr>
<td>3:1</td>
<td></td>
</tr>
<tr>
<td>4:1+</td>
<td></td>
</tr>
</tbody>
</table>

- **Under-forecasts** (too few deep clouds)
- **Hits** (Fcst similar to obs)
  - 80% LG; 74% SM
- **Over-forecasts** (too many deep clouds)
Composites behave differently for observed events vs. predicted events.

False alarms and missed forecasts are associated with different errors.
Diagnostic Statistics

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.
Conditional Bias Difference as a Summary Score

FCST-OBS Bias (mm)

Given FC Event

Given OB Event

Scale of zero bias

Square area dimension (gpts)
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?

LWP (g m⁻²)
Fuzzy Neighborhood Method

- 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.

Roberts and Lean (2008)
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.

Scale

FSS = \( \frac{2f_{0}f_{m}}{f_{0}^2 + f_{m}^2} \)

(=1 if no bias)
**Fuzzy Neighborhood Method**

A Few Caveats

- Large scale samples exceed grid bounds
- May cause aliasing

\[ FSS_{(n)} = 1 - \frac{MSE_{(n)}}{MSE_{(n)\text{ref}}} \]

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

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

- Reference MSE not a true climatology
- Changes with each forecast
- Positive bias leads to large \( MSE_{\text{ref}} \) and improved FSS
Calculating Scale-Dependent Scores

Method:

- Identify a threshold of interest
  - \( \text{LWP} \geq 500 \text{ g m}^{-2} \)
- 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

21-hr FCSTS

Box size

Summary score:
= \sum (\text{Frac}_F - \text{Frac}_O)^2

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.