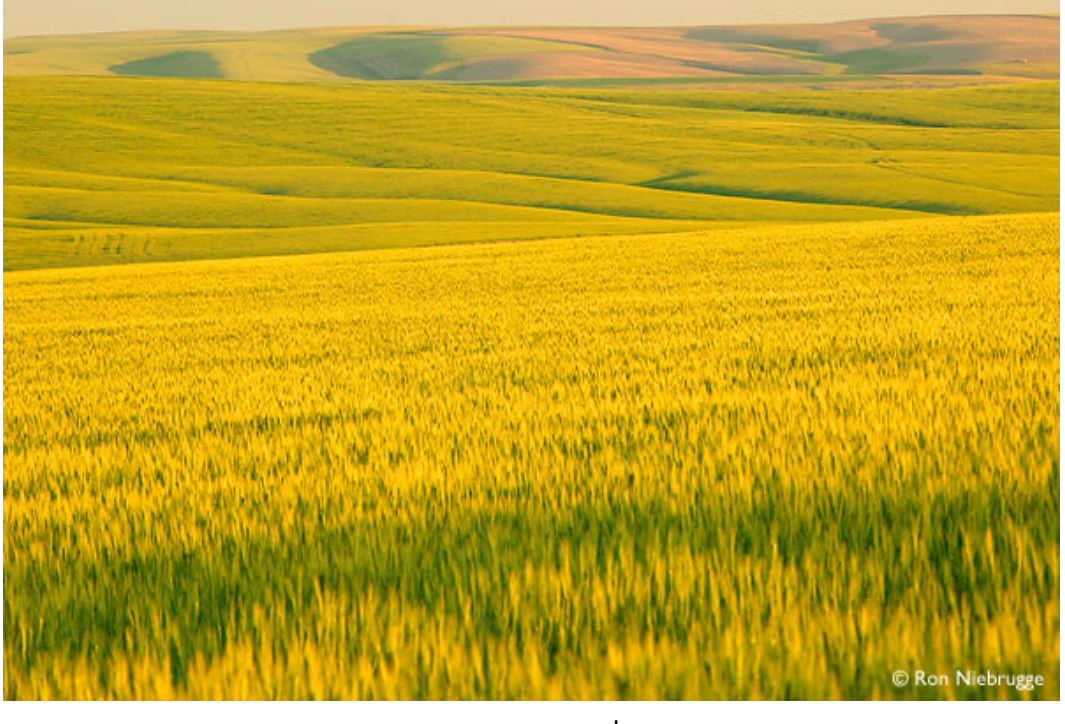
Using WRF output in the ICP



Eric Gilleland† *et al.*

†Research Applications Laboratory, EricG @ ucar. edu

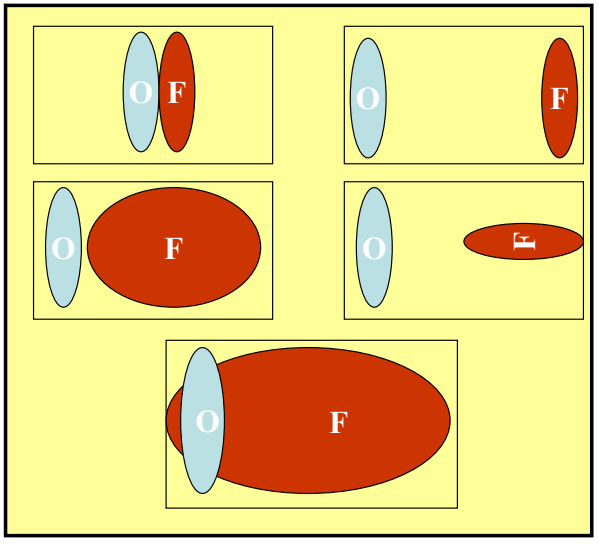
Outline

Spatial Forecast Verification Methods Inter-Comparison Project (ICP)

- Motivation and Goals for the ICP
- Summary of Methods included
- Case Studies

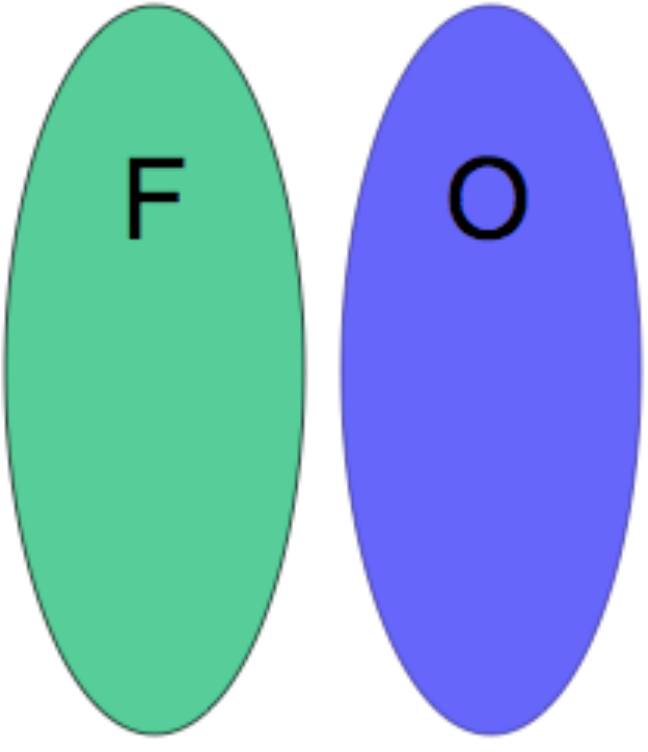
Motivation and Goals

Example

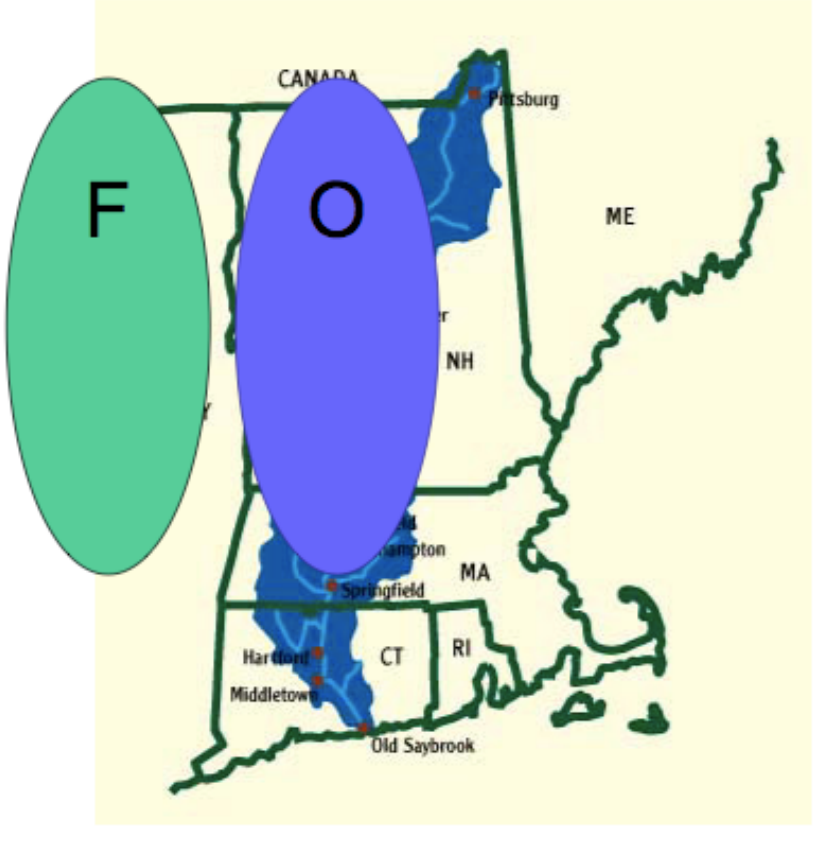


- First four forecasts have POD=0; FAR=1; CSI=0
  - i.e., all are equally "BAD"
- Fifth forecast has POD>0, FAR<1, CSI>1
- Traditional verification approach identifies "worst" forecast as the "best"

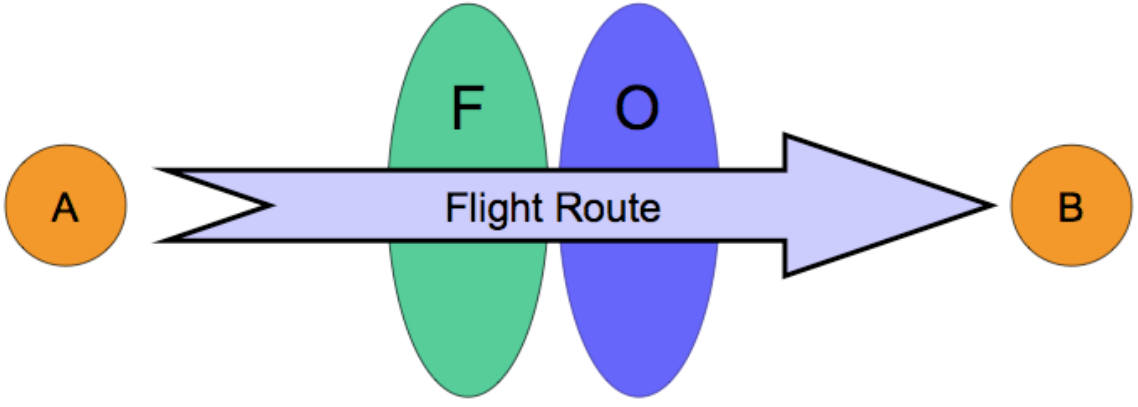
What makes a good forecast?



Motivation and Goals



Motivation and Goals



Motivation and Goals

Ultimate Goal: Set of guidelines for users

Challenges

- Comparing wide variety of methods each yielding different types of information.
- Difficult to determine true quality even for a human (subjective) observer.
- Multiple types of possible errors (displacement, intensity, coverage, etc.).

Summary of Methods

Four primary categories (not all methods fit to this exactly)

- Features-based
- Field Verification
- Neighborhood-based
- Scale Decomposition

Summary of Methods: Features-based

See web site given at end of this talk for full references.

- Davis *et al.* (2006a, 2006b) (now called MODE)
- Ebert and McBride (2000), Contiguous Rain Area (CRA)
  - Numerous modifications (e.g., Grams *et al.* (2006))
- Marzban and Sandgathe (2006a, 2006b), Cluster Analysis (CA)
- Nachamkin (2004), Composite Method
- Micheas *et al.* (2006), Cell Identification/Procrustes Shape Analysis
- Wernli *et al.* (2007), Structure, Amplitude and Location (SAL)

Summary of Methods: Field Verification

Let $\mathbf{F}$ represent the forecast field, and $\mathbf{A}$ the analysis field. The idea is to compare $\varphi(\mathbf{F})$ with $\mathbf{A}$, where $\varphi(\mathbf{F})$ is some kind of morph of $\mathbf{F}$ such that some discrepancy measure is minimized, and performance is judged by the "amount" of morphing required.

Keil and Craig (Submitted to MWR), Forecast Quality Measure (FQM)

Summary of Methods: Neighborhood-based

Traditional verification scores compared at different scales of resolution to determine highest resolution with desired *skill*.

Numerous methods proposed under this heading, see Ebert (2006) for an excellent review.

Summary of Methods: Scale Decomposition

$\mathbf{F}$ and $\mathbf{A}$ (or $\Delta = \|\mathbf{F} - \mathbf{A}\|$) decomposed via a single-band spatial filter (e.g., Fourier transforms, wavelets,etc.). For example,

$$\Delta(\mathbf{x}) = \sum a_i \varphi(\mathbf{x_i})$$

Compute traditional scores at each scale (i.e., wave number), or at each scale set all other coefficients to zero and inverse transform, then compute score on (smoothed) field.

- Briggs and Levine (1997)
- Casati *et al.* (2004), Intensity-scale
- Harris *et al.* (2001), (not performed on different scales separately)
- Mittermeier (2006), (expansion of Casati *et al.* for operational use)
- Zepeda-Arce *et al.* (2000)
- more ...

Case studies

- Various *real* cases. Beginning with NCEP Spring 2005 Program output.
  - WRF 4-km NCEP
  - WRF 4-km NCAR
  - WRF 2-km*
  - Stage II Analysis
- Known perturbations of one or more real cases (i.e., known errors).
- Simple and contrived cases (e.g., Baldwin and Kain (2006)).

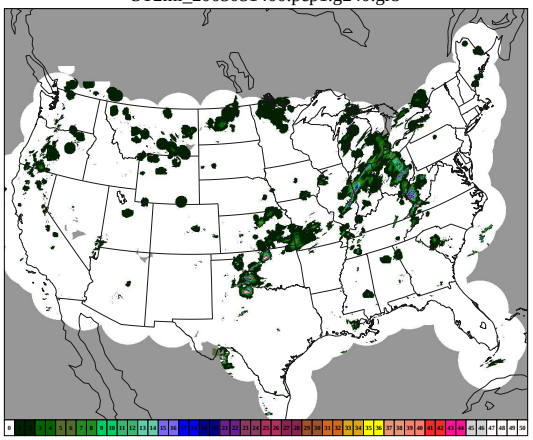
*All output and analysis are first put onto the same 4-km grid.

Case studies: Real Cases

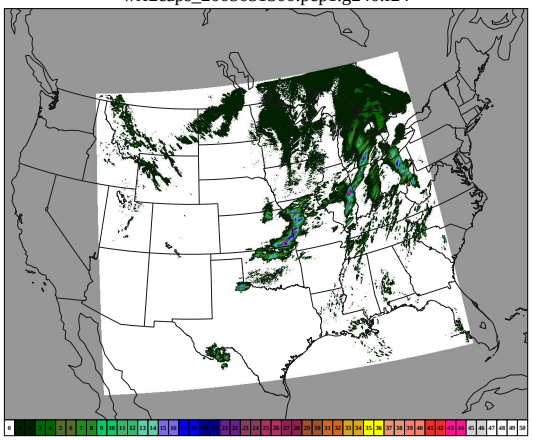
First Set

Storm Prediction Center (SPC) Spring 2005 Program Precipitation.

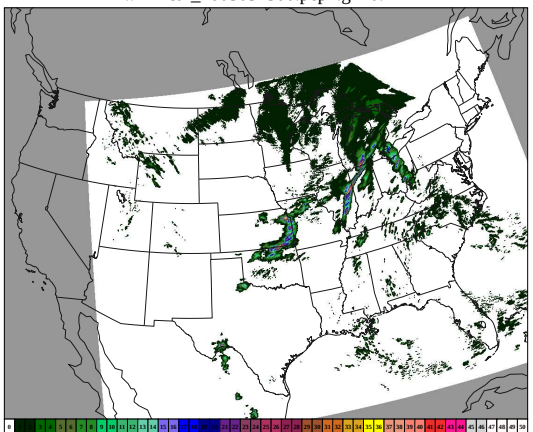
ST2ml_2005051400.pcp1.g240.grb



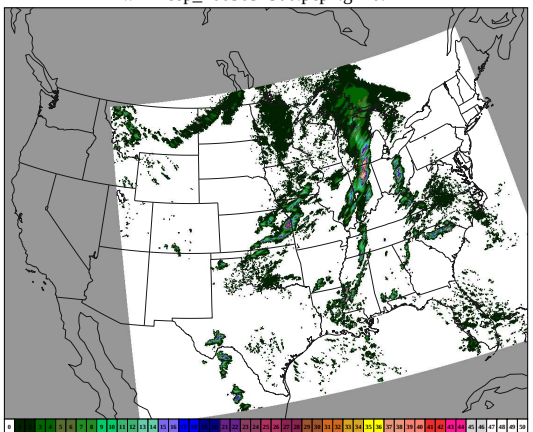
wrf2caps_2005051300.pcp1.g240.f24



wrf4ncar_2005051300.pcp1.g240.f24



wrf4ncep_2005051300.pcp1.g240.f24



Case studies: Real Cases

- Initially 9 hand-selected cases used (plan to look at about 30).
- 1-h precipitation accumulations.
- 501×601 grid.
- $\approx$ 4-km grid squares.

Case studies: Real Cases

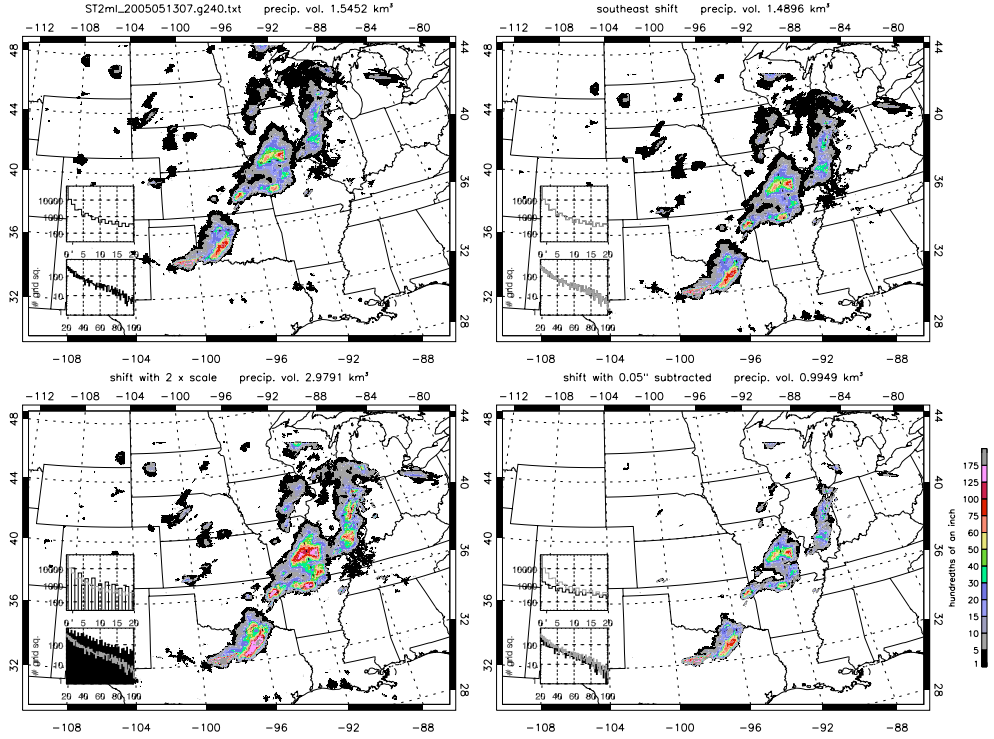
Questions to answer/Information to glean

- How does each method compare with subjective assessments?
- What is the most important aspect of forecast quality for each method?
- Information on scales that are appropriate.
- Clarification on how the forecast fails.
- Easily interpretable results?
- How to better use the forecast.
- Are there meteorological situations that one method is better equipped to handle?
- ...

Case studies: Known errors

Perturbed real cases

How does each method inform the user about various types of errors?

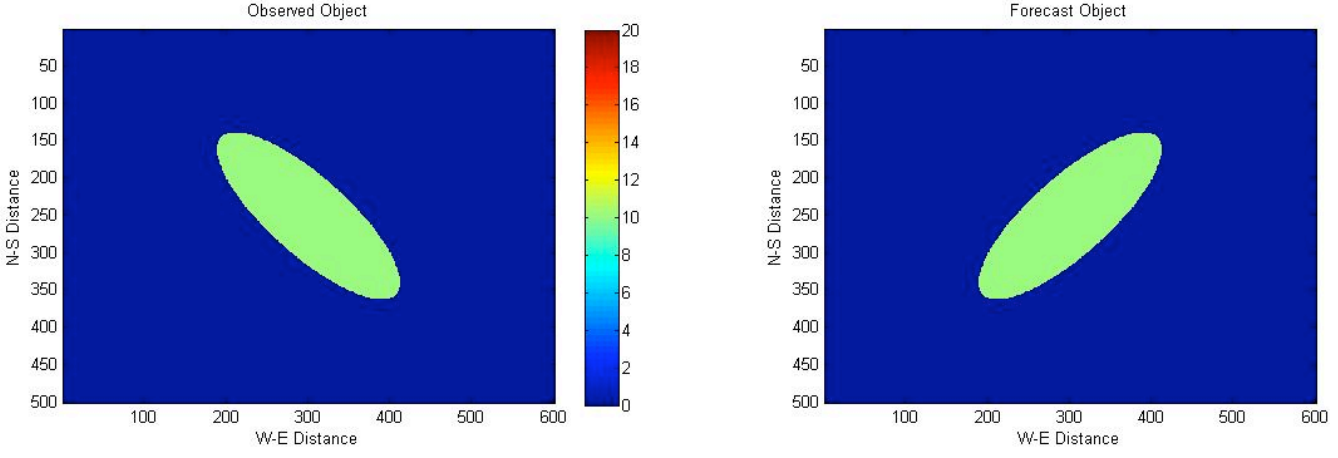


Case studies: Known errors

Simple contrived cases

Sensitivity of each method to particular types of errors (e.g., size, shape, density, smoothness,etc.).

Can the forecast be *hedged* to obtain a better score?



That's all!

Project website

http://www.rap.ucar.edu/projects/icp