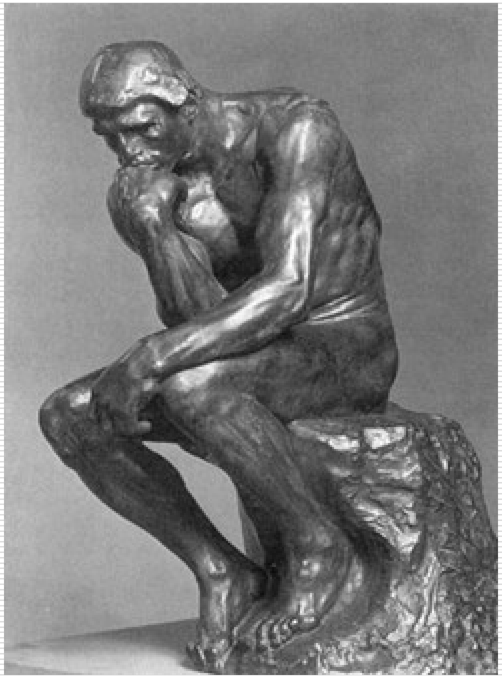
Observation uncertainty

Or… “There is no Such Thing as TRUTH”

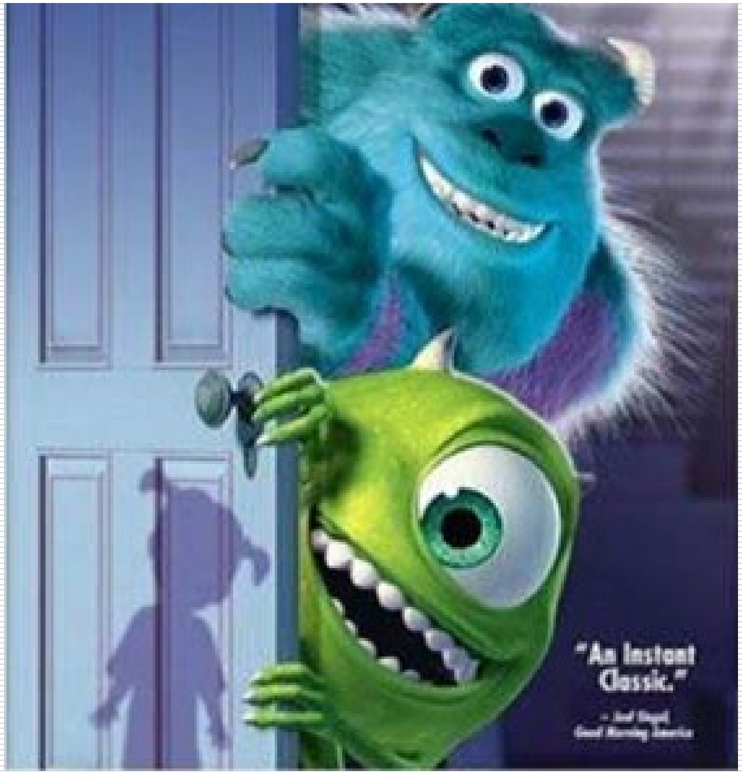


Barbara Brown
NCAR
Boulder, Colorado USA

May 2017

The monster(s) in the closet...

- What do we lose/risk by ignoring observation uncertainty?
- What can we gain by considering it?
- What can we do?



Outline

- What are the issues? Why do we care?
- What are some approaches for quantifying and dealing with observation errors and uncertainties?

Sources of error and uncertainty associated with observations

- Biases in frequency or value
- Instrument error
- Random error or noise
- Reporting errors
- Representativeness error
- Precision error
- Conversion error
- Analysis error/uncertainty
- Other?

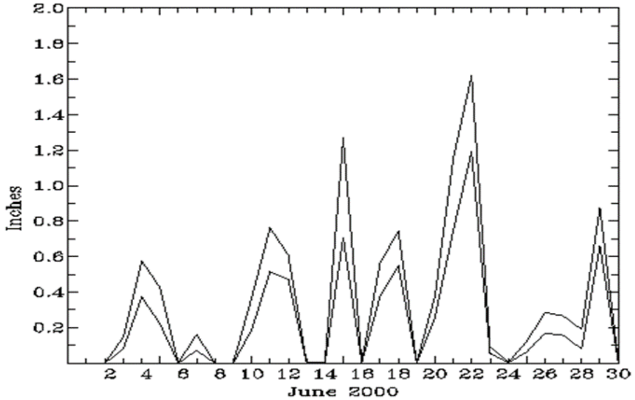


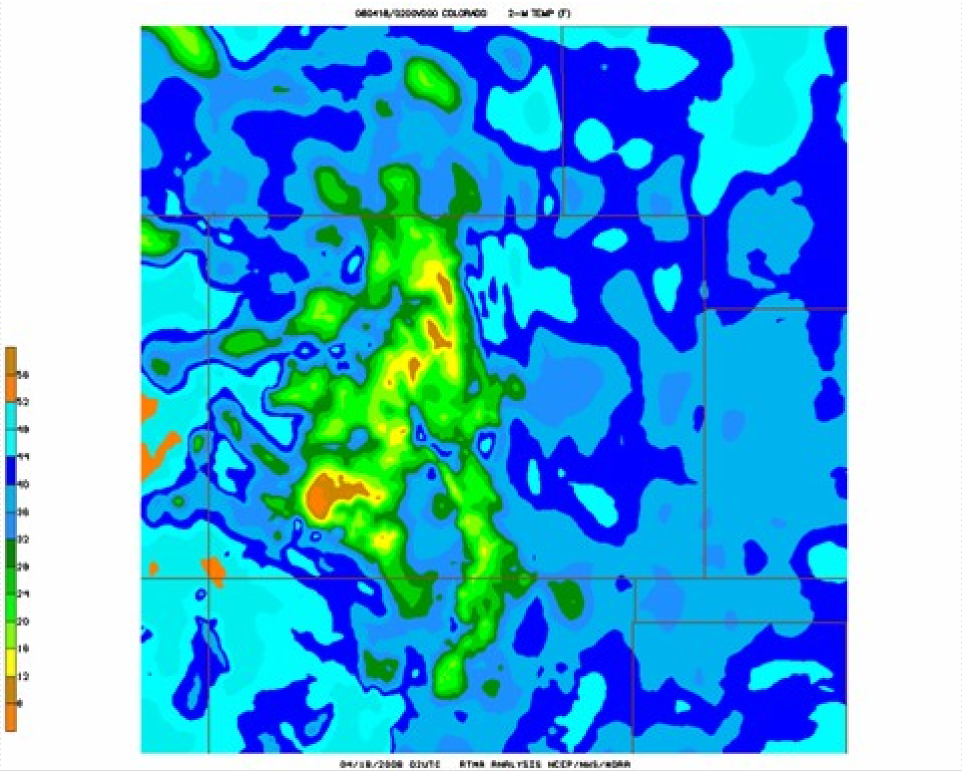
Fig 4. Daily (1200 UTC - 1200 UTC) rainfall averaged over all RFC-selected stations in Region R of Fig. 2. Upper curve shows averages computed excluding apparently missing observations; lower curve indicates comparable averages computed after converting missing observations to zero reports.

Example: Missing observations interpreted as "0's"

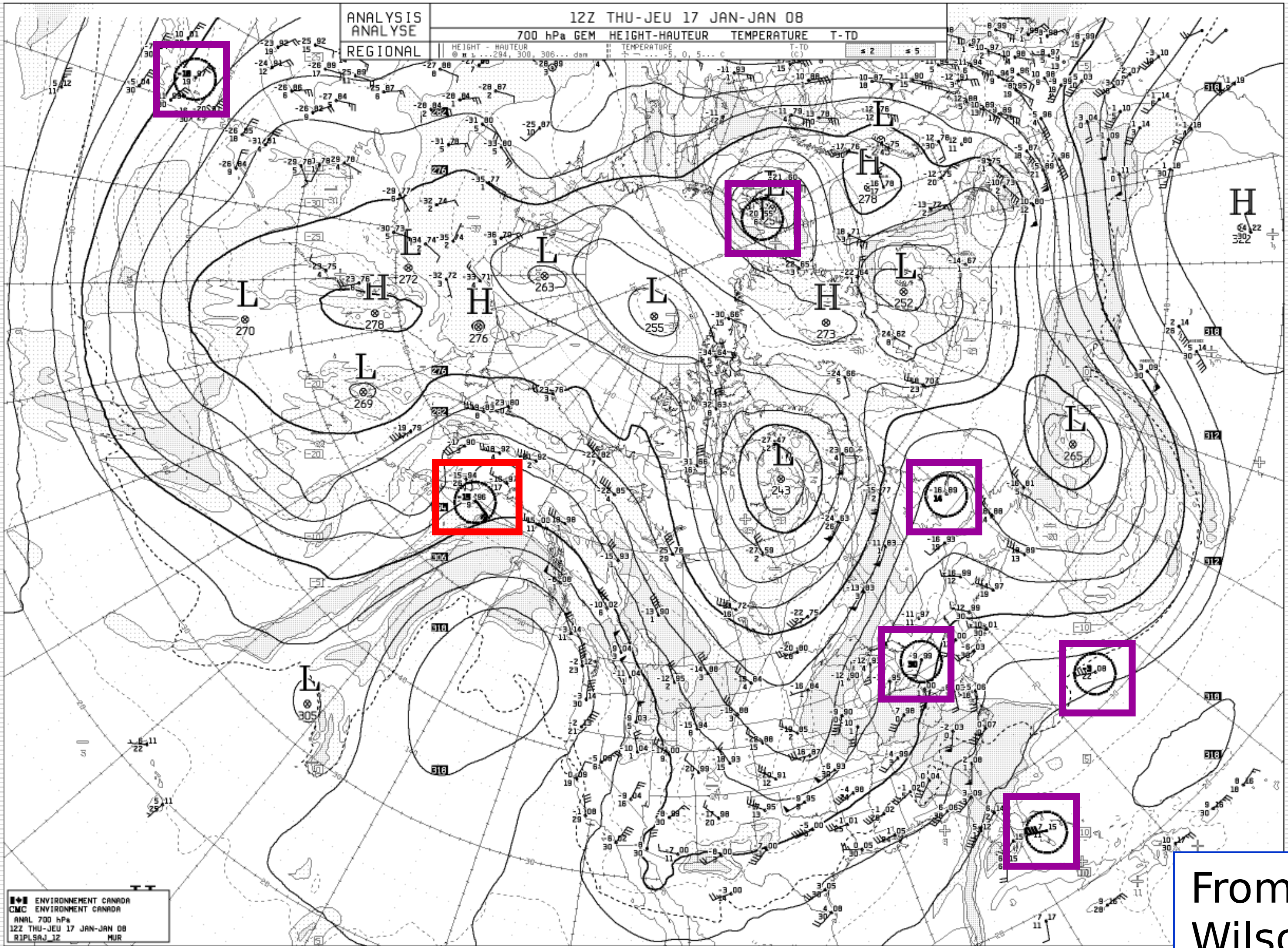
Issues: Analysis definitions

- Many varieties of analyses are available
- (How) Have they been verified? Compared?
- What do we know about analysis uncertainty?

RTMA 2 m temperature



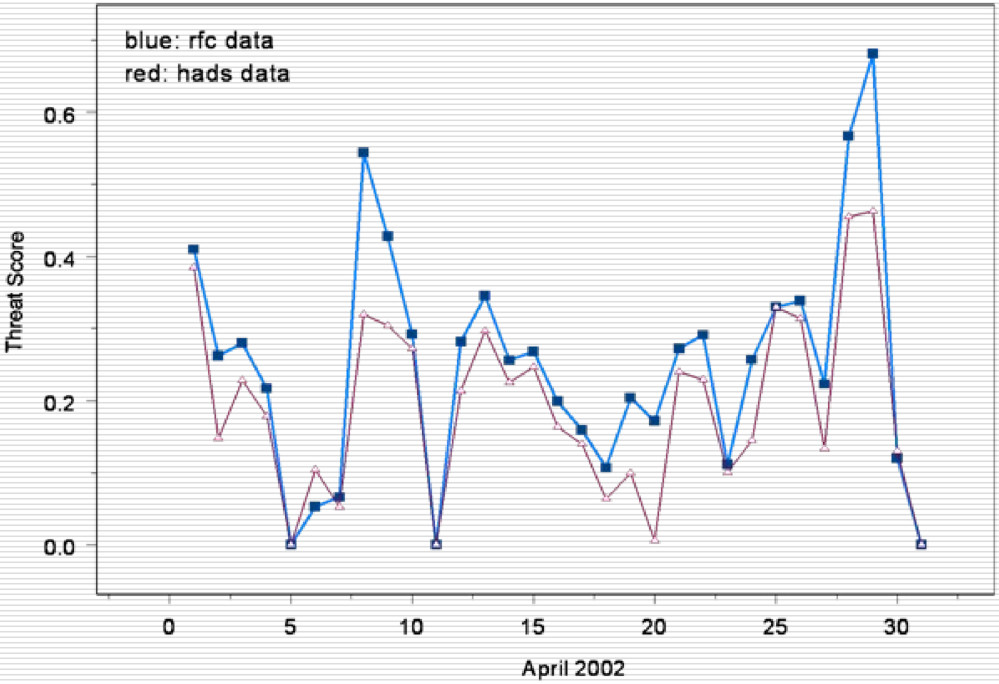
Issue – Data filtering for assimilation and QC



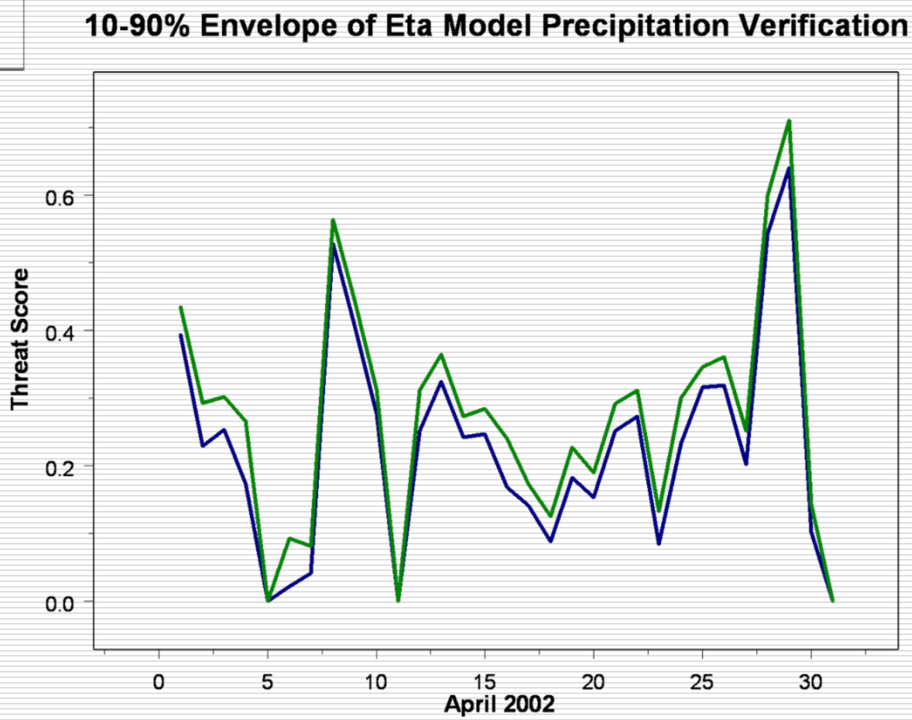
From L. Wilson

700 hPa analysis; Environment Canada; 1200 UTC, 17Jan 2008

Impacts: Observation selection



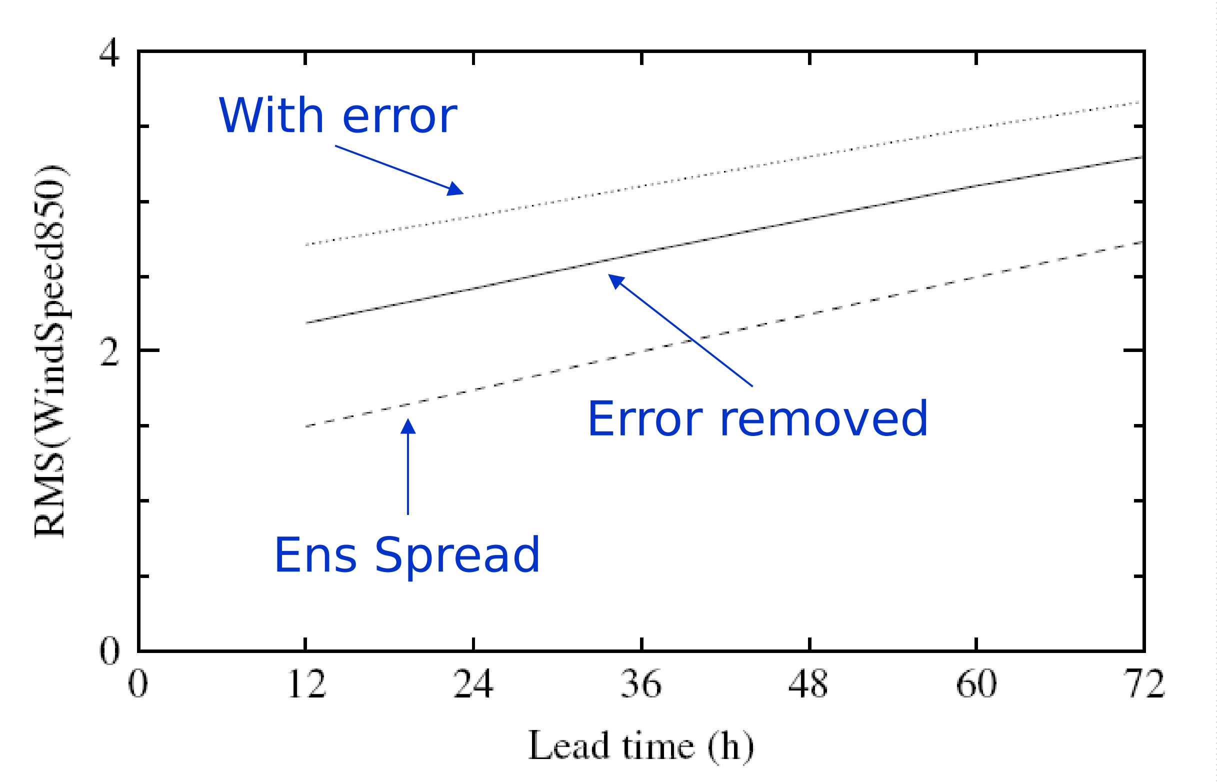
Verification with different datasets leads to different results



Random subsetting of observations also changes results

From E. Tollerud

Issue: Obs uncertainty leads to under-estimation of forecast performance



850 mb Wind speed forecasts

Assumed error = 1.6 ms^{-1}

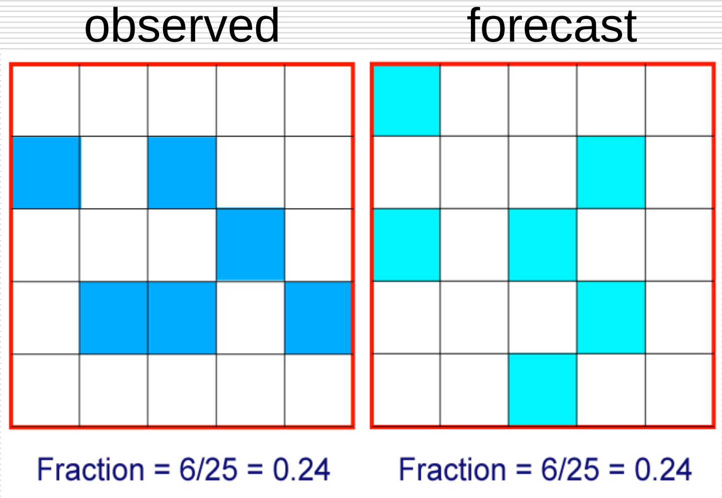
From Bowler 2008 (Met. Apps)

Approaches for coping with observational uncertainty

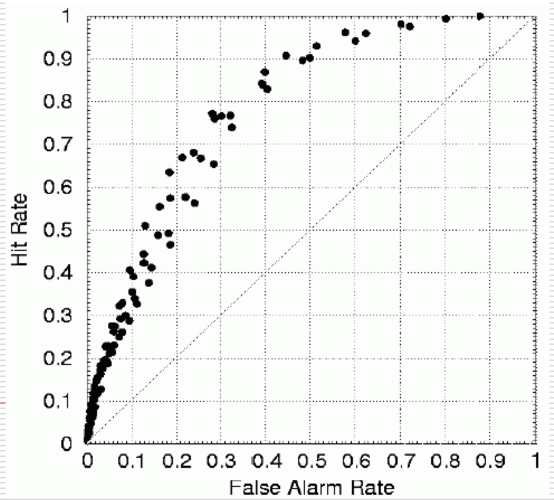
- Indirect estimation of obs uncertainties through verification approaches
- Incorporation of uncertainty information into verification metrics
- Treat observations as probabilistic / ensembles
- Assimilation approaches

Indirect approaches for coping with observational uncertainty

- Neighborhood or fuzzy verification approaches
- Other spatial methods



Vary distance and threshold



(Atger, 2001)

Direct approaches for coping with observational uncertainty

- Compare forecast error to known observation error
  - If forecast error is smaller, then
    - A **good** forecast
  - If forecast error is larger, then
    - A **bad** forecast
- Issue: The performance of many (short-range) forecasts is approaching the size of the obs uncertainty!

Direct approaches for coping with observational uncertainty

- Bowler, 2008 (MWR)
  - Methods for reconstructing contingency table statistics, taking into account errors in classification of observations
- Ciach and Krajewski (1999)
  - Decomposition of RMSE into components due to "true" forecast errors and observation errors

$$\mathrm{RMSE}_{\mathrm{o}} = \sqrt{\mathrm{RMSE}_{\mathrm{t}}^{2} + \mathrm{RMSE}_{\mathrm{e}}^{2}}$$

Where $\mathrm{RMSE}_{\mathrm{e}}$ is the RMSE of the observed values vs. the true values

Direct approaches for coping with observational uncertainty

- Candille and Talagrand (QJRMS, 2008)
  - Treat observations as probabilities (new Brier score decomposition)
  - Perturb the ensemble members with observation error
- Hamill (2001)
  - Rank histogram perturbations

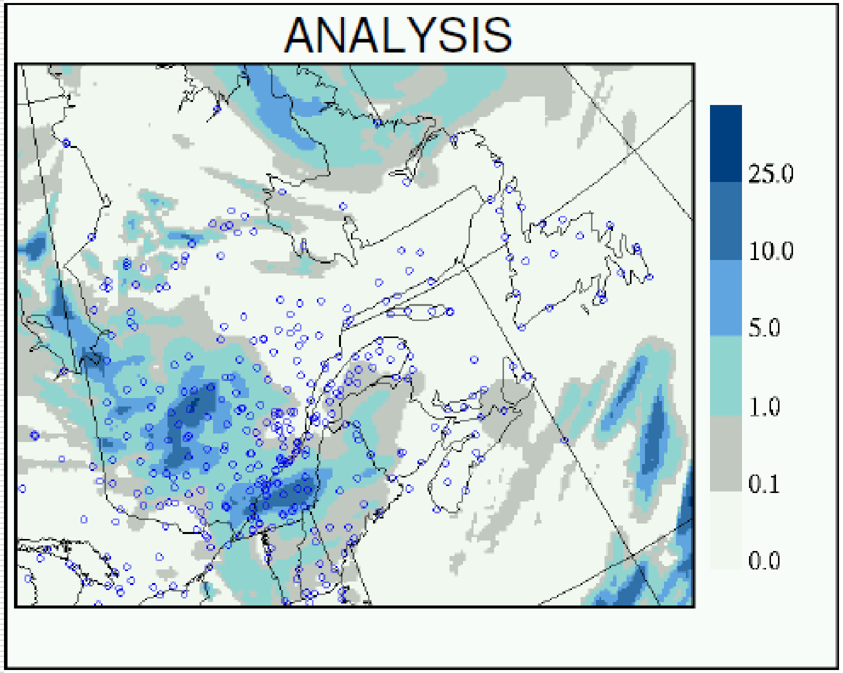
Direct approaches for coping with observational uncertainty

- B. Casati et al.
  - Wavelet reconstruction
- Gorgas and Dorninger, Dorninger and Kloiber
  - Develop and apply ensembles to represent observation uncertainty (VERA)
  - Compare ensemble forecasts to ensemble analyses

Casati wavelet approach

- Use wavelets to represent precipitation gauge analyses
- Use wavelet-based approach
  - Reconstruct a precipitation field from sparse gauges observation
  - Apply scale-sensitive verification

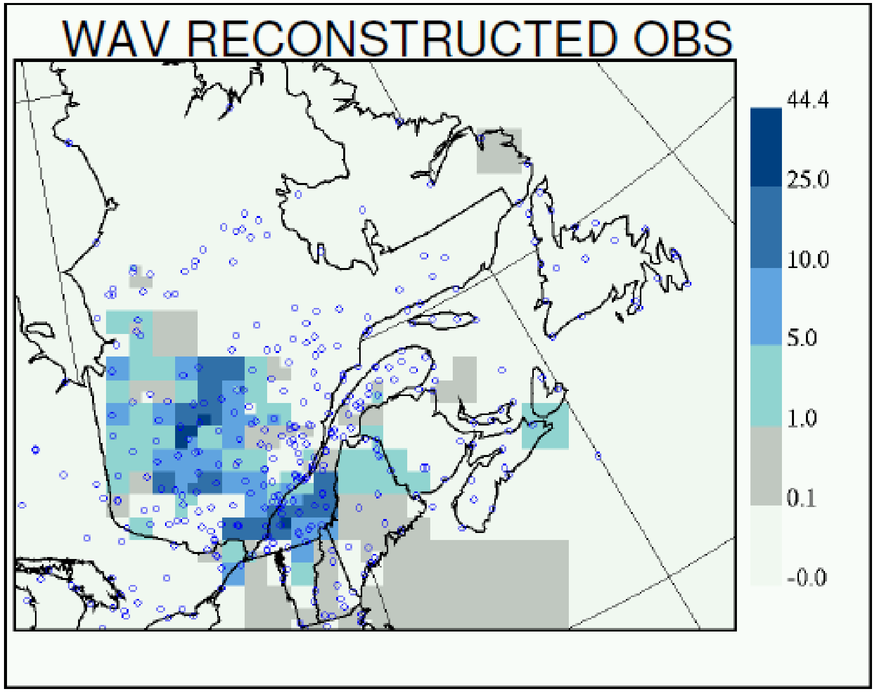
[**Recall**: Manfred Dorninger's presentation yesterday on wavelet-based intensity-scale spatial verification approach]

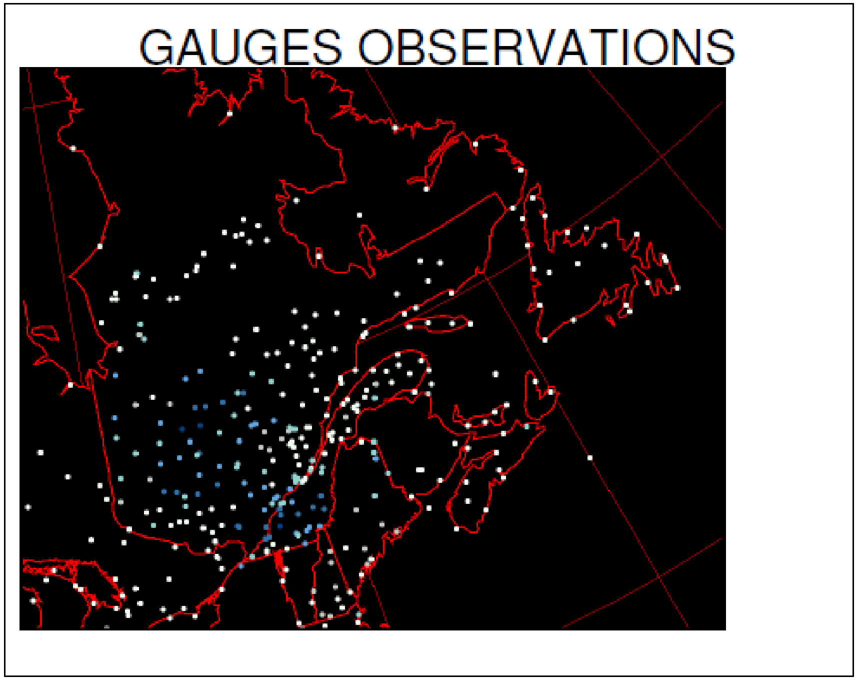


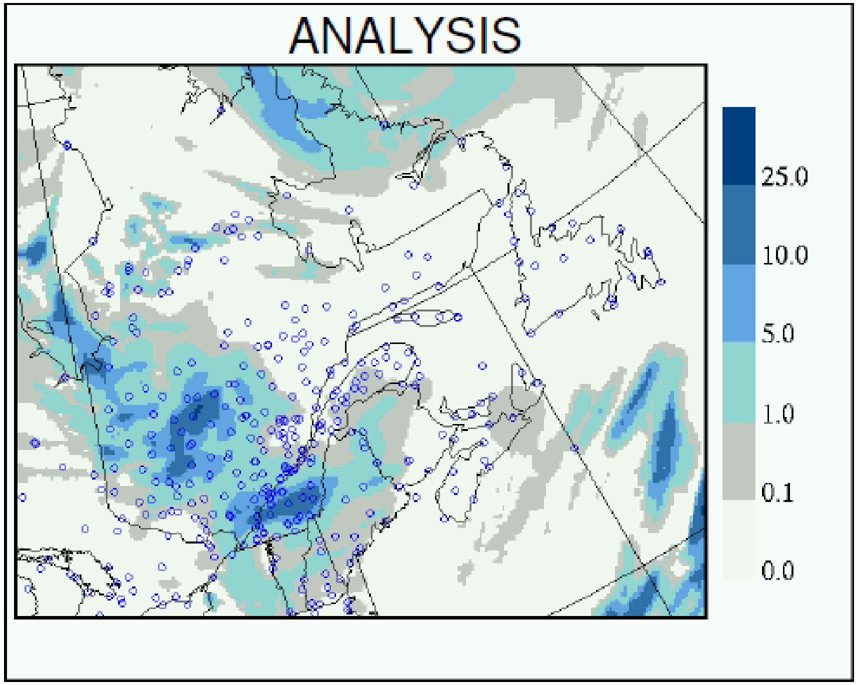
This approach…
•Accounts for existence of features and coherent spatial structure + scales
•Accounts forgauge network density
•Preserves gauge precip. values at their locations

From B. Casati

Example: 6h acc (mm)
27th Aug 2003, 6:00 UTC

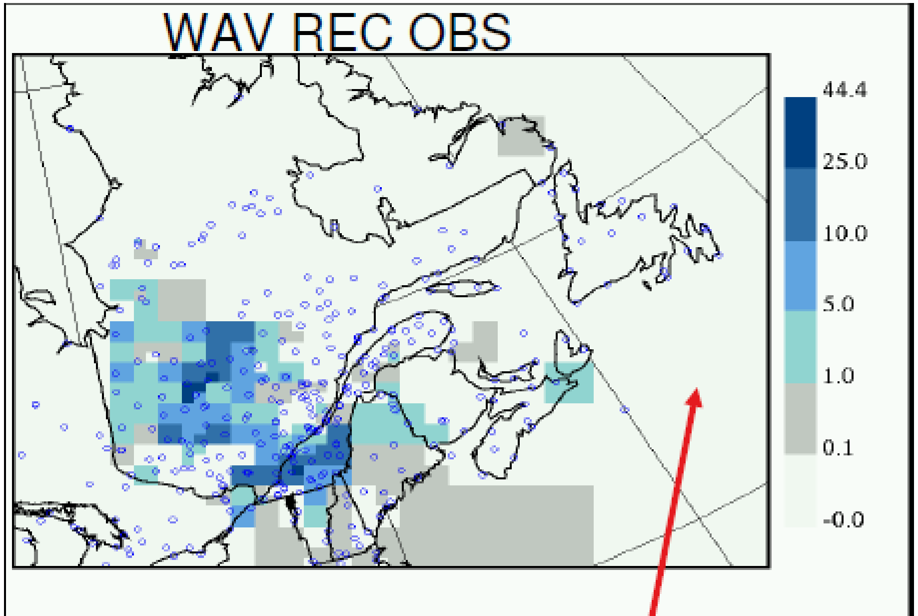


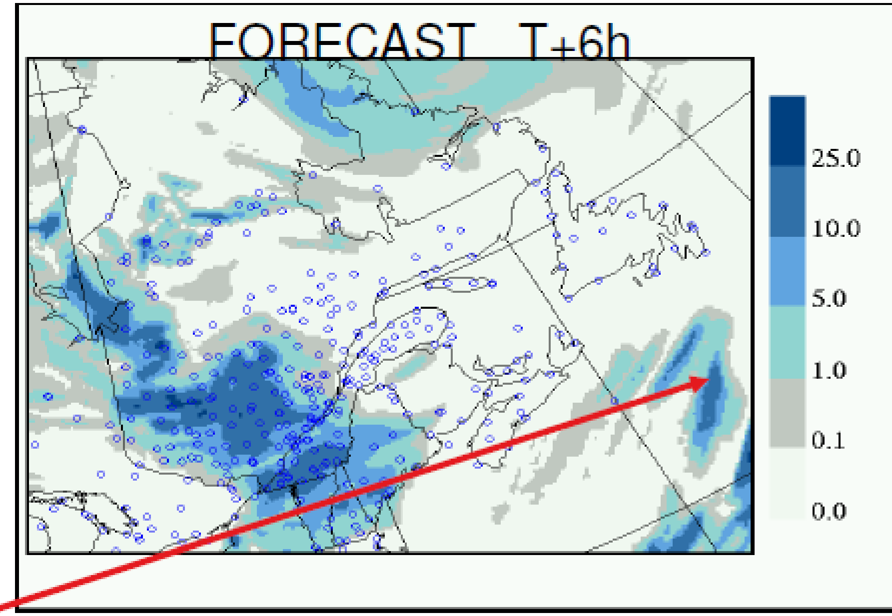




- Account for existence spatial structures on different scales
- Account for gauge network density
- Value at station location is equal to gauge value

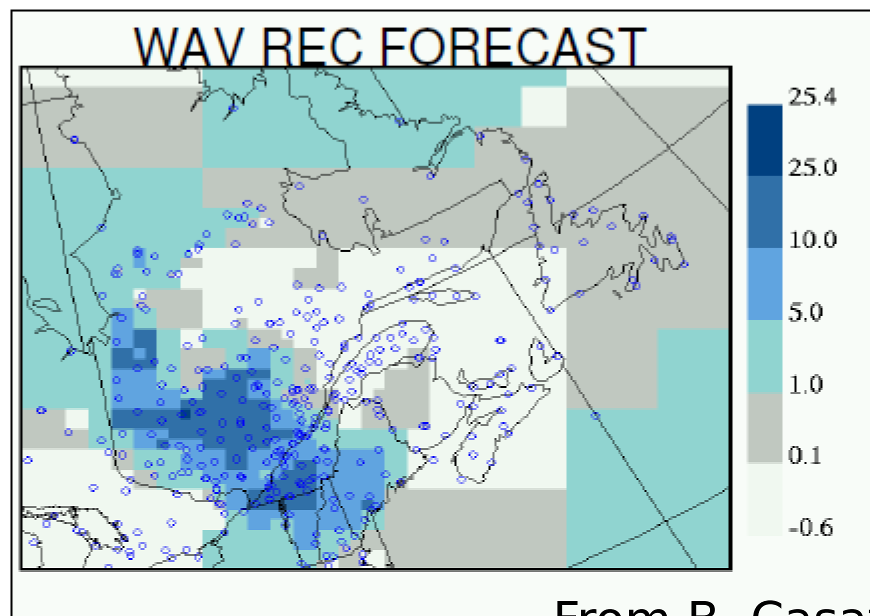
3. Representativeness and forecast filtering



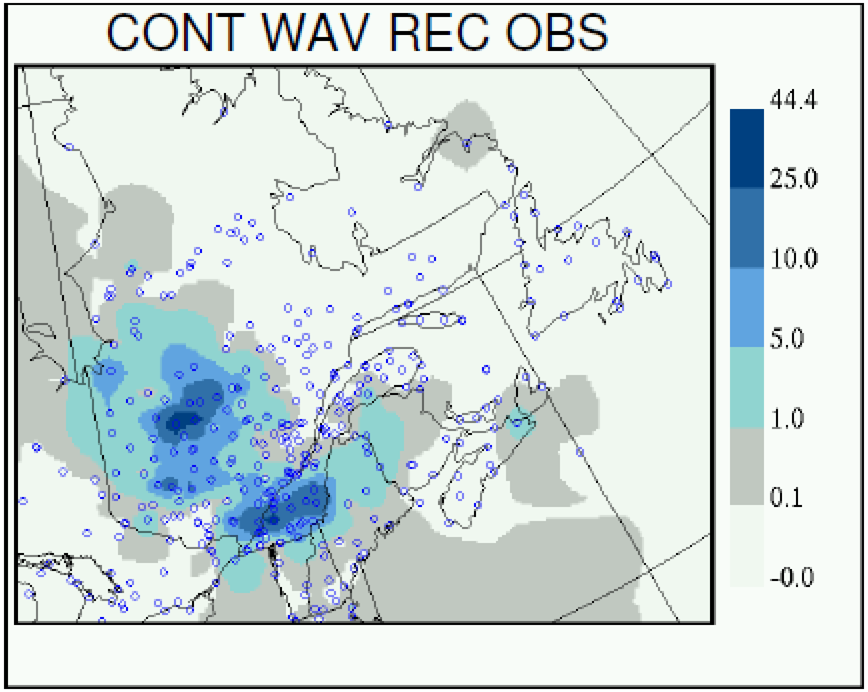


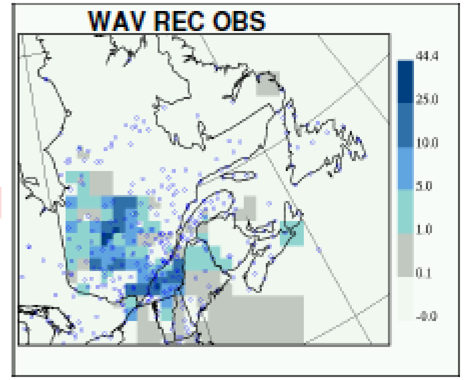
No gauges = missing obs,
but forecast has features!

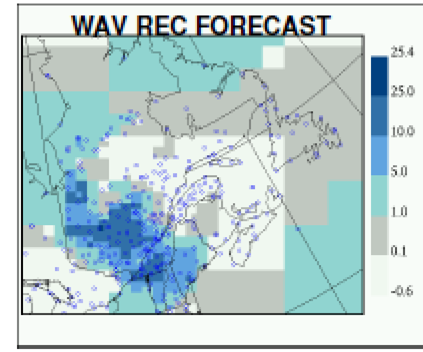
2. Decompose forecast with wavelets
3. Set to NA wavelet coefficients where no obs
4. Reconstruct forecast field



From B. Casati



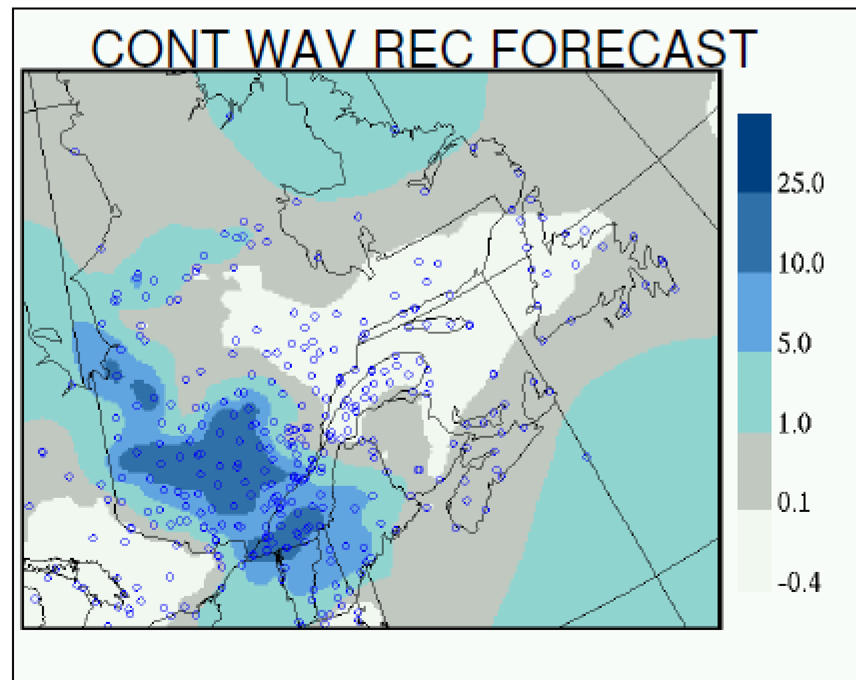




Discrete wavelets = squared areas with fix location; these are not always representative

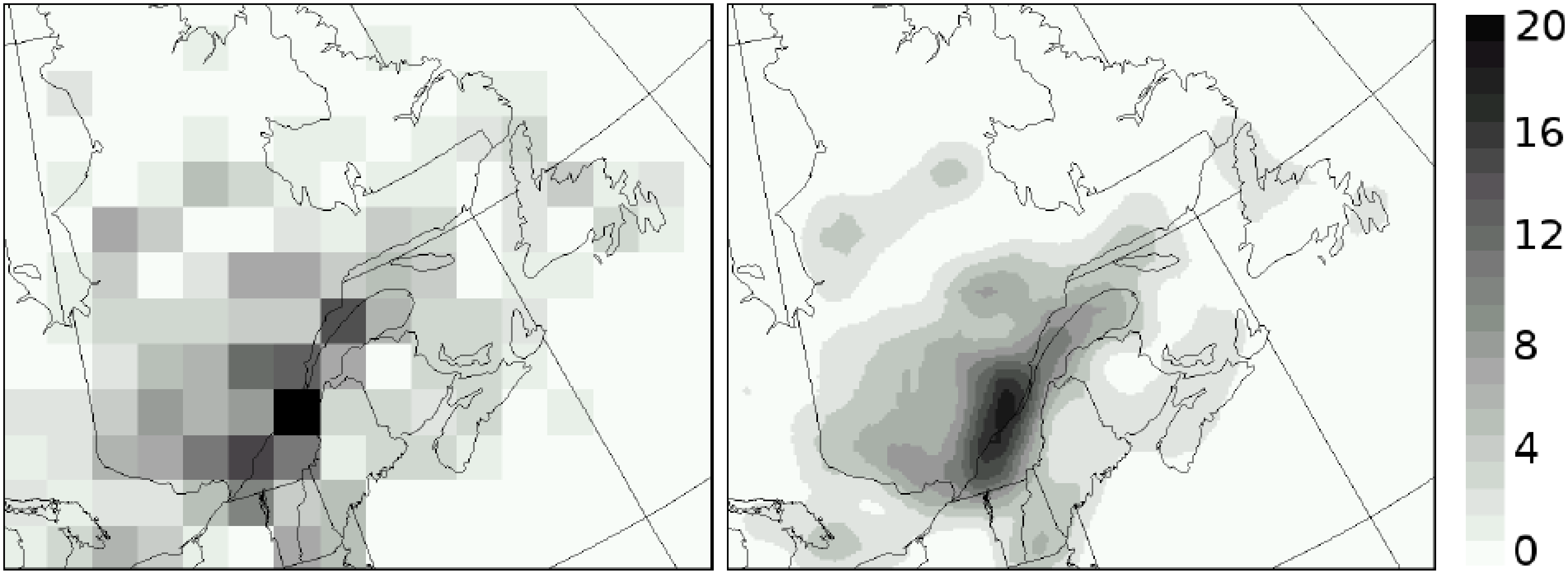
Eliminate discrete effect by moving the wavelet support and averaging

→ Continuous wavelets



From B. Casati

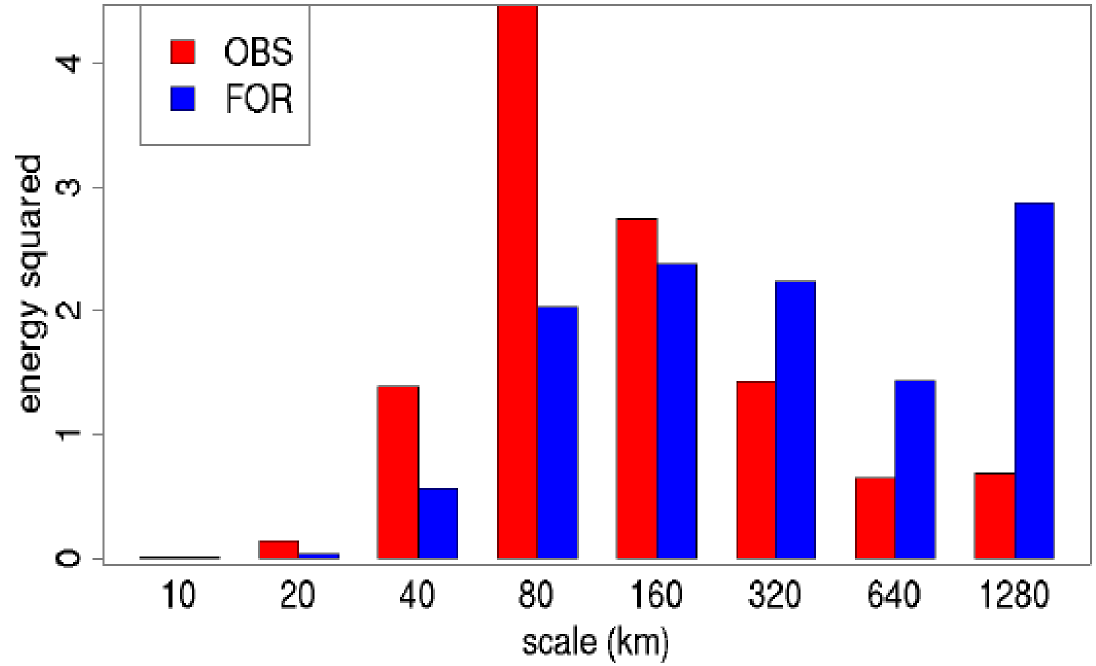
4. Confidence (uncertainty) mask

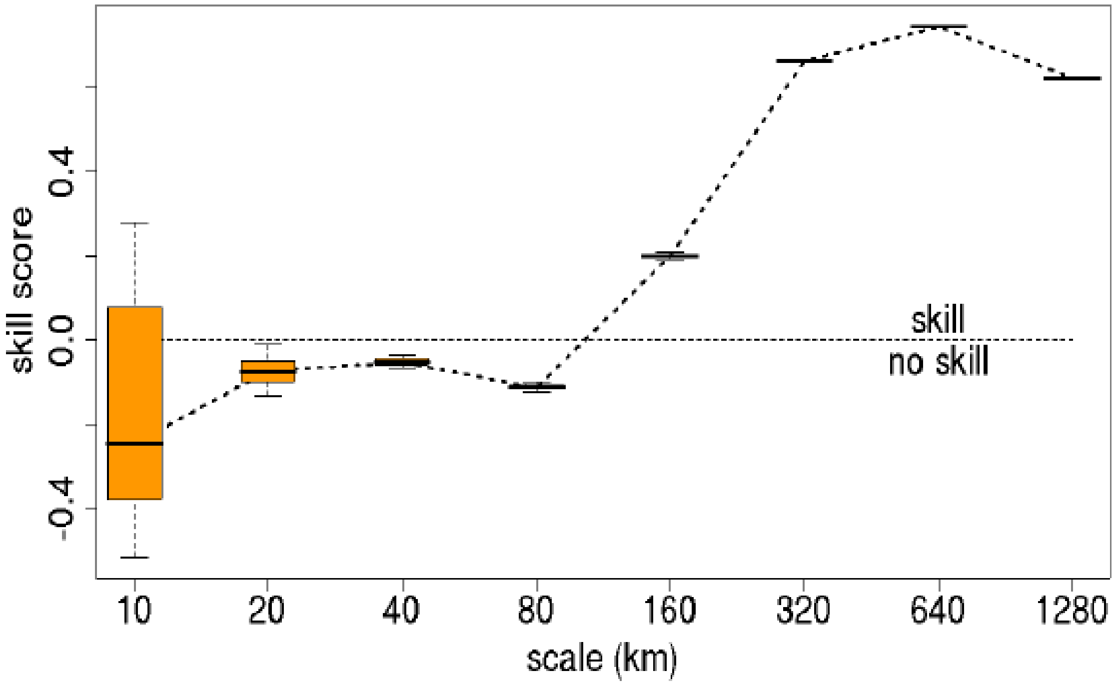


For each scale (e.g. 160 km resolution scale) provide <u>confidence/uncertainty associated to reconstructed fields</u>

large number of gauges ←→ confidence
small number of gauges ←→ uncertainty

From B. Casat





5. Verification

on different scales, but only **where obs are available**

1. Energy squared:

$$En^2(X) = \langle X^2 \rangle$$

Measures the quantity of events and their intensity at each scale => BIAS, scale structure

2. MSE Skill Score:

$$1 - \frac{MSE(Y,X)}{En^2(X) + En^2(Y)}$$

(related to correlation)

From B. Casati

VERA Application (Dorninger and Kloiber)

Data

VERA: Vienna Enhanced Resolution Analysis used to correct the surface observations and to interpolate them to a regular grid.

VERA Ensembles: Main structures in spatial fields defined by wavelet transforms, on this field perturbations gets applied. Resolution 8 km, hourly, two different set-up's ("std" and "equ-qc"), 50 members [1]

Forecast: COSMO-LEPS (CLE) provided by Arpae-SIMC Emilia-Romagna. 10 km resolution interpolated on VERA grid, every three hour, 16 members

Fig.1: Schematic plot of the investigated Area and City [2]

Verification - RMSE

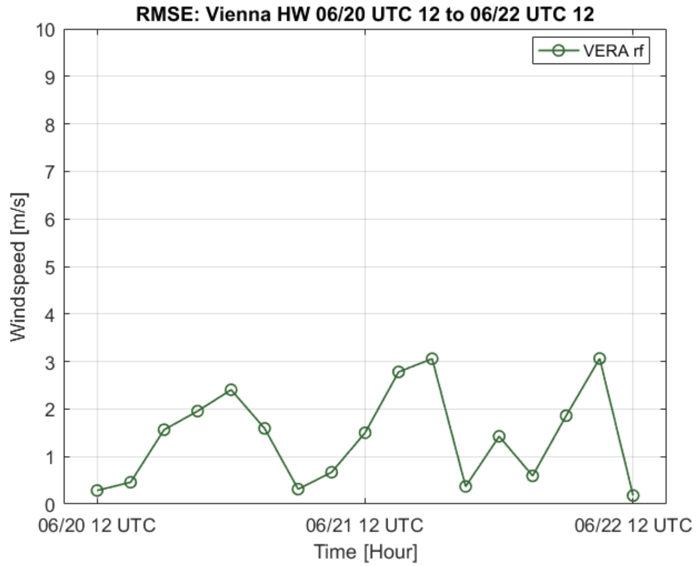


Fig.3: RMSE calculated with VERA reference and CLE mean (initial time: 06/20 12 UTC)

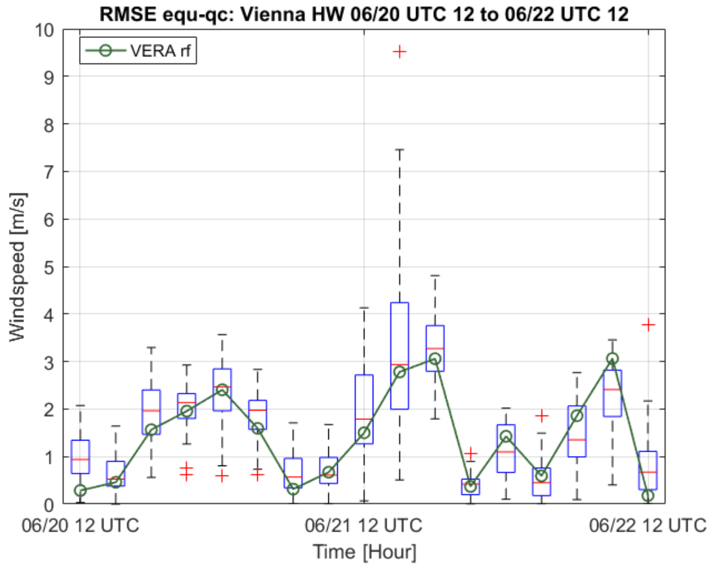


Fig.4: RMSE additionally calculated with VERA ensemble (Boxplot) and CLE mean (initial time: 06/20 12 UTC)

Dorninger and Kloiber

Verification - Time Evolution

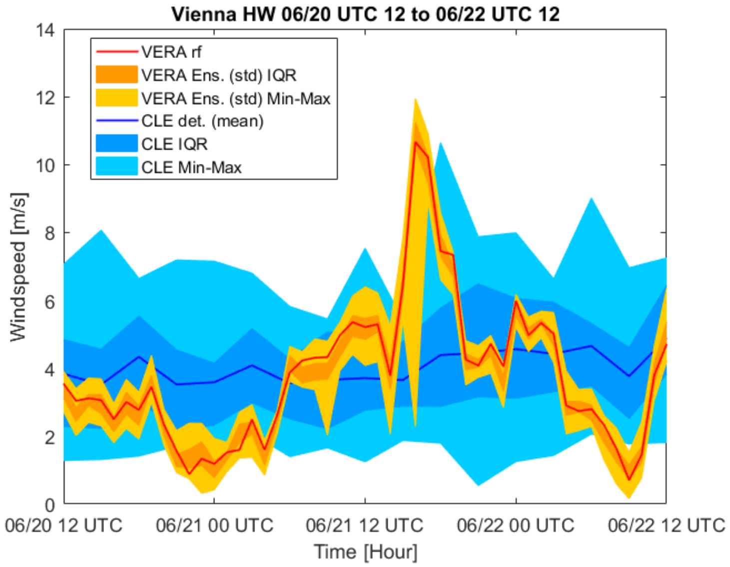


Fig.5: Time series of VERA Ensemble (std) and all CLE runs (initial time: 06/20 12 UTC)

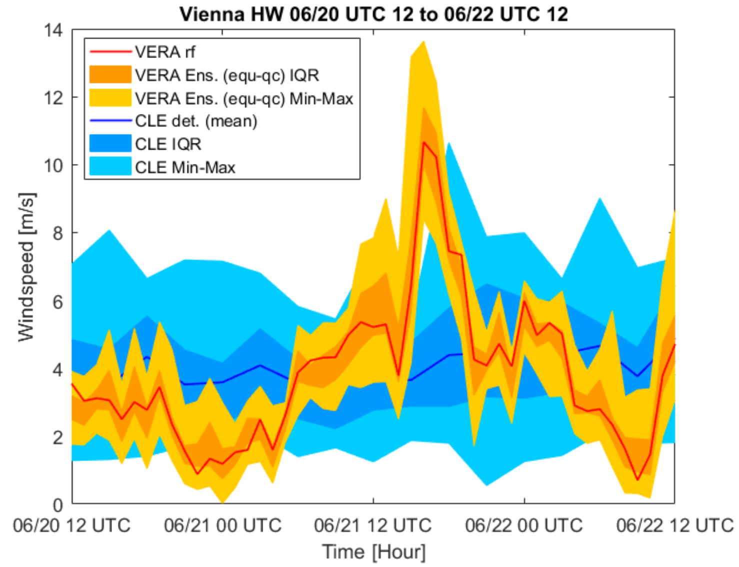
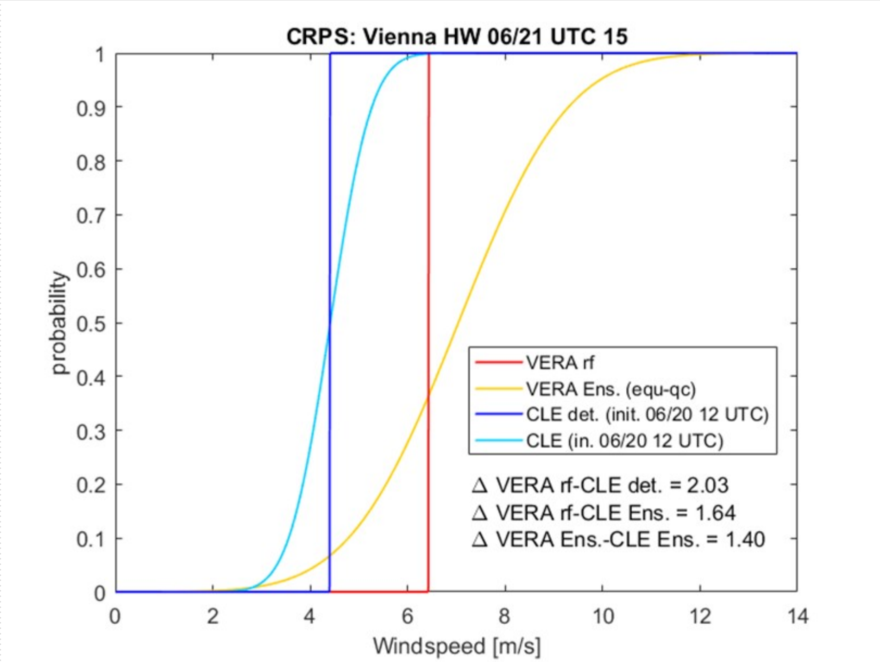


Fig.6: Time series of VERA Ensemble (equ-qc) and all CLE runs (initial time: 06/20 12 UTC)

Dorninger and Kloiber

Comparing observation ensemble to forecast ensemble (Dorninger and Kloiber)

- CRPS
- Modified ROC
- Distance metrics
- Distribution measures



Summary and conclusion

- Observation uncertainties can have large impacts on verification results
- Obtaining and using meaningful estimates of observational error remains a challenge
- Developing “standard” approaches for incorporating this information in verification progressed in recent years – but still a distance to go… room for new researchers!

DISCUSSION / COMMENTS / QUESTIONS