Accounting for the effect of observation errors on verification

Neill Bowler

Thanks to: Beth Ebert
Outline

- Assumptions
- Root-mean-square error
- Rank Histograms
- Categorical verification (ROC)
- Verification against analysis
Assumptions

- We know the distribution of observation errors
- The errors in the observations are independent of the observed value, the forecast value and the errors in other observations (additive, uncorrelated noise)
- In this case, verification is against radio-sonde observations of wind speed at 850 hPa
- Observation errors are assumed gaussian with zero mean and standard deviation 1.6 m/s
How to estimate the observation errors?

- With difficulty!

- Differences between observations at different locations (extrapolate distance between obs to zero) – NB Ingleby, J. Atmos. Ocean. Technology, 18, 1102-1107 (2001)

- It may be possible to diagnose them from a series of assimilation cycles – G Desroziers et al, QJ, 131, 3385-3396 (2005)
Verification details

- Verification performed for 1 November 2006 to 26 January 2007 on MOGREPS global ensemble

- An in-sample bias correction has been applied to the forecast data

- Any event threshold are basic (e.g. wind speed > 10m/s) so Hamill et al’s (QJ, in press) “false skill” issue is not addressed
The effect of observation errors is

\[ RMS(f, o) = \sqrt{RMS(f, t)^2 + RMS(o, t)^2} \]

What we measureWhat we want to measure

So, we estimate the “true” RMS error by

\[ RMS_{est}(f, t) = \sqrt{RMS(f, o)^2 - RMS(o, t)^2} \]
Rank histograms

To calculate the rank histogram, rank the ensemble forecasts, and find between which members the verification falls.

If the ensemble is sampling from the distribution of forecast errors, then the rank histogram should be flat.

Remove the effect of observation errors by perturbing each ensemble member’s forecast by the observational error.

Rank histograms – 850 hPa wind speed

Near-flat when observation errors accounted for

T+72h forecasts
Contingency tables

- If the distribution of observed values (given the model forecast some event to occur) is
  \[ P_o(x \mid F = 1) \]
- then, under our assumptions about observation errors, this is related to the distribution of true values by
  \[ P_o(x \mid F = 1) = \int_{-\infty}^{\infty} P_t(y \mid F = 1)P_e(x - y)dy \]

Observation distribution

Most observations show wind speed greater than 10m/s when all members forecast this.

In both cases deconvolved distribution is slightly narrower.

Distribution of all observed wind speeds.
ROC for wind speed 850hPa > 10m/s T+72

- Raw observation data
- Corrected for observation errors
Verification against analysis gives a lower RMSE than verification against observations, corrected for their error, either:

- Our estimate of the observation error is too low
- The analysis has errors, which are correlated with forecast errors
Correlation of analysis and forecast error

- Looked at by Simmons and Hollingsworth (QJ, 2002) for 500 hPa height
- They found correlations of analysis error of around 0.5 (or less) at 1 day

- For wind speed at 850 hPa, when fitting data using an AR-1 correlation model
- Observation error = 1.6 m/s
- Analysis error = 0.6 m/s
- Correlation between analysis and forecast error

<table>
<thead>
<tr>
<th></th>
<th>T+12</th>
<th>T+24</th>
<th>T+36</th>
<th>T+48</th>
<th>T+60</th>
<th>T+72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corr</td>
<td>0.83</td>
<td>0.69</td>
<td>0.57</td>
<td>0.47</td>
<td>0.39</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Verification against analysis gives more outliers than verification against observations, corrected for errors.
Summary

- Assumptions

- Root-mean-square error
- Rank Histograms
- Categorical verification (ROC)

- Verification against analysis
Any questions?
Perturbing Forecasts vs Deconvolution

- Adding the observation error to the forecasts is treating the observation error as an error in the forecast.
- For example, one might say that the forecast is unable to represent the small-scale detail, and needs to be downsampled to the observation site – this would reduce the resolution of the ensemble forecast.
- The deconvolution approach treats the observation as being in error.
- Since rank histograms are not measuring resolution, the difference is unimportant.
- The distinction is important for categorical verification.
RMSE T250

![Diagram showing RMSE over lead time](image-url)