Observation impact diagnostics in an Ensemble Data Assimilation System

1. Introduction: Forecast impact diagnostics developed so far

**Motivation**
Assessing the impact of observations in data assimilation systems is generally extremely expensive:
- Envis experiments (OSE).

**Aim**
- Find diagnostic tools to indicate impact of observation subsets on analysis/forecast.
- Identify where observation impact is sub-optimal/positive.

Cost-function
- gives impact of obs assimilated at time $t = 0$.
- can be written as sum over the individual observations.

Time evolution
- Different DA Systems use different approaches for computing the time evolution.
- 4D Var (Langland and Barker 2004)
- use linear (adjoint) model
- Ensemble Kalman Filter (EKF) use ensemble

Here: Ensemble Kalman Filter Verification with observations

2. Optimality condition

Consider the cost function $J$ for different initial conditions

If $J$ has a minimum for one finds

- $\Delta x = 0$ should be small --- "optimality condition"
- $\Delta x$ should be positive (and large) --- "potential benefit"

3. Our System

- LETKF (Hunt et al. 2007) - 40 ensemble members
- verification vs obs. (Sommer and Weissmann 2014)
- time evolution via analysis ensemble (Kalnay et al. 2012)

weights are computed as:

4. Results - What is shown?

So far only results for $t=0$ ("impact on analysis").

Statistics have been computed for different cost-function components separately:

5. Results: Impact on analysis ($t=0$)

Tems verified by GPSRO

- Excellent correspondence/consistency between Terms and GPSRO
- ($J_t$) clearly positive everywhere
- ($J_t$) much smaller than ($J_t$ - $\Delta x$)
- optimality condition largely fulfilled.
- Many positively
- weight on TERMS in assimilation could be slightly increased in tropical regions

AMSU-A verified by GPSRO

- ($J_t$) generally positive, strong potential
- ($J_t$) clearly negative
- Obs have too strong weight
- AMSU-A observation errors are too small
- Increased observation errors have been tested for operational implementation
- But: Positive impact on forecast only after reduced thinning of AMSU-A

Data from individual AMSU-A channels ($\chi_t$) - normalised

The correspondence between AMSU-A channels and GPSRO is positive or neutral for most channels and most latitudes ($\chi_t$) is mostly positive.

In some regions the dark shaded areas are negative or neutral but the green curve (bias removed a posteriori) is clearly positive. This indicates that AMSU and GPSRO have opposite bias.

Significant bias problems occur for:
- ch.4 (everywhere)
- chs.9+10 (esp. towards poles).

Ch. 9+10 are the lowest channels used over land. As seen on the right below bias problems are much stronger over land than over sea.

6. Outlook: Assessing impact of individual observations

Work so far:
Cost-function $J$ gives impact of all observations assimilated at $t=0$.
Interpretation of different components (corresponding to individual observations) is suggestive but not rigorous ($\chi_t - \Delta x$) depends on all observations used in assimilation).

More rigorous approach:
Cost-function for data denial experiment. Replace: $\chi$ $\rightarrow$ $\chi_{den}$ (analysis not using $\chi$)

7. Discussion

The interpretation of observation impact diagnostics is often not trivial. Statistical significance is a big issue (particularly for large forecast lead times). To facilitate the interpretation and to differentiate between model and DA issues the work presented here has been (so far) restricted to $t=0$ (impact on the analysis).

It is explained that different parts of the cost function $J = \frac{1}{2} (\chi - \chi_{true})^T R^{-1} (\chi - \chi_{true})$ should be considered (interpreted) separately.

Examples are given for how the diagnostics could be linked to the following observational problems:
- The use of too small observation errors in the DA system for AMSU-a
- Biases of AMSU-A channels $\leftrightarrow$ Bias of GPSRO

Biais problems only show up if the bias is opposite to the bias which the verifying obs (here GPSRO) have with respect to the model.

The diagnostics reveal inconsistencies. The separation into contributions from different observations is, however, not rigorous. Particularly for strongly overlapping observations the interpretation in terms of impact on analysis is problematic. A method is under development to show the "decentral impact" for individual observations (e.g. a single AMSU channel). See sec. 6.

References

DWD, German Meteorological Service, Data assimilation section
Author: Olaf Stiller, Olaf.Stiller@dwd.de