Simulation and diagnosis of error contributions in DA cycling

Loïk Berre, Gérald Desrozières, Benjamin Ménétrier
7th WMO DA symposium
Florianopolis, 14 September 2017
What does contribute to forecast errors?

Several contributions, with different « ages »:

- Forecast errors arise from analysis errors and model errors.

- Analysis errors result from background errors and observation errors.

- DA cycling: background errors depend on previous background errors, previous observation errors and previous model errors, and so on.

- Goal of this study: simulate these different error contributions, diagnose their amplitude and evolution during the cycling.

- Motivations: get knowledge of error dynamics in DA cycling, develop error simulation and estimation methods.
Outline

- Expansion of forecast error contributions
- Old *versus* recent error contributions
- Observation *versus* model error contributions
- Conclusions
What does contribute to forecast errors? (linear expansion at cycling step $t_i$)

$$
\begin{align*}
\varepsilon_i^f &= \varepsilon_i^a + \epsilon_i^m \\
&= \mathbf{M}_i(I - \mathbf{K}_i \mathbf{H}_i)\varepsilon_i^b \\
&= \mathbf{M}_i(I - \mathbf{K}_i \mathbf{H}_i)[\mathbf{M}_{i-1}(I - \mathbf{K}_{i-1} \mathbf{H}_{i-1})\varepsilon_{i-1}^b + \varepsilon_{i-1}^m] + \mathbf{M}_i \mathbf{K}_i \varepsilon_i^o + \epsilon_i^m \\
&= \mathbf{T}_2 \varepsilon_{i-1}^b \\
&= \ldots \\
&= \mathbf{T}_{i+1} \varepsilon_0^b \\
&+ \sum_{j=0}^{i} \mathbf{T}_{i-j}(\mathbf{M}_j \mathbf{K}_j \varepsilon_j^o + \epsilon_j^m) \\
\end{align*}
$$

where, for $j < i$,

$$
\mathbf{T}_{i-j} = \prod_{k=j+1}^{i} \mathbf{M}_k(I - \mathbf{K}_k \mathbf{H}_k)
$$

and $t_0$ is the beginning of the considered cycling period.

(e.g. El Ouaraini and Berre 2011)
Old and recent error contributions

\[
\varepsilon_i^f = T_{i+1} \varepsilon_0^b + \sum_{j=0}^{i} T_{i-j}(M_j K_j \varepsilon_j^o + \varepsilon_j^m)
\]

with \( T_{i-j} = \prod_{k=j+1}^{i} M_k(I - K_k H_k) \) (= cycling operator).

How do these 3 error contributions compare and how do they evolve during the cycling?
Simulation of error contributions of old background and recent observations

- Baseline ensemble DA experiment (EDA): Arpege 4D-Var (global NWP), observation perturbations and multiplicative inflation, warm start on 9 January 2017 from operational EDA; 6h cycling; same $B_j$ (provided by operational EDA) for all xp’s.

- To quantify contributions of $\varepsilon^b_0$ and $\varepsilon^o_j$, variants of this EDA baseline are run, from 9 to 22 January 2017 (2 weeks):
  - OLD_Eb: only $\varepsilon^b_0$ contributes
  - RECENT_Eo: only $\varepsilon^o_j$ contributes, from $t_0$ until current time $t_i$
  - OLD+RECENT_Eo: only $\varepsilon^o_j$ contributes, from $t_{-24}$ (6 days before $t_0$) to $t_i$

- Evolution of global variance of error contributions for temperature (500 hPa) from corresponding ensemble spread $^2$. 


Evolution of old background error contribution (spread$^2$ of OLD_Eb xp)

$$\text{Var}(T_{i+1} \varepsilon_0^b)$$

with

$$T_{i+1} = \prod_{k=0}^{i} M_k (I - K_k H_k)$$

$$T_{i+1} \simeq (M(I - KH))^{i+1}$$

Old background errors are dampened by successive DA steps (~ 4-day period).
Evolution of recent observation error contributions (spread² of RECENT_Eo xp)

Recent observation errors are accumulated and dampened by successive DA steps. Convergence like a power series.
Evolution of old & recent observation error contributions

Total (old+recent) contribution is stable.
Evolution of old & recent observation error contributions

The total (old+recent) contribution is stable: compensation between damping of old errors and accumulation of recent errors.
Contributions to forecast error variance

\[ \varepsilon_i^f = T_{i+1} \varepsilon_0^b + \sum_{j=0}^{i} T_{i-j} M_j K_j \varepsilon_j^o + \sum_{j=0}^{i} T_{i-j} \varepsilon_j^m \]

with

\[ \text{Var}(T_{i+1} \varepsilon_0^b) \approx 0 \quad \text{for } i \gtrsim \tau^T, \]
where \( \tau^T \approx 4 \text{ days} \) is the timescale over which old errors vanish,

\[ \text{Cov}(\varepsilon_0^b, \varepsilon_j^o) = 0 \quad \text{for time uncorrelated random observation errors}, \]

\[ \text{Cov}(T_{i+1} \varepsilon_0^b, T_{i-j} \varepsilon_j^m) \approx 0 \quad \text{when } i \gtrsim \max(\tau^T, \tau^m) \]
where \( \tau^m \) is the correlation timescale of random model errors.

for \( i \gtrsim \max(\tau^T, \tau^m) \):

\[ \text{Var}(\varepsilon_i^f) = \text{Var}(\sum_{j=0}^{i} T_{i-j} M_j K_j \varepsilon_j^o) + \text{Var}(\sum_{j=0}^{i} T_{i-j} \varepsilon_j^m) \]
Diagnosis of recent model error contributions

Contributions to forecast error variance:

$$\text{Var}(\varepsilon_i^f) = \text{Var}\left(\sum_{j=0}^{i} T_{i-j} M_j K_j \varepsilon_j^o\right) + \text{Var}\left(\sum_{j=0}^{i} T_{i-j} \varepsilon_j^m\right)$$

which leads to the following estimation approach (e.g. at day 4, considering $max(\tau^T, \tau^m) \simeq 4$ days):

- $\text{Var}(\varepsilon_i^f)$ estimated by innovation-based diagnostics (e.g. Desrozières et al 2005);
- $\text{Var}\left(\sum_{j=0}^{i} T_{i-j} M_j K_j \varepsilon_j^o\right)$ estimated by EDA with observation perturbations only;
- $\text{Var}\left(\sum_{j=0}^{i} T_{i-j} \varepsilon_j^m\right) = \text{Var}(\varepsilon_i^f) - \text{Var}\left(\sum_{j=0}^{i} T_{i-j} M_j K_j \varepsilon_j^o\right)$
Diagnosis of model error contributions versus observation error contributions

\[ \sigma\left( \sum_{j=0}^{i} T_{i-j} M_j K_j \varepsilon_j^o \right) \]

\[ \sigma\left( \sum_{j=0}^{i} T_{i-j} \varepsilon_j^m \right) \]
Conclusions

- Linear forecast error expansion to diagnose (at different ages) background, observation and model error contributions.

- Global forecast error variance tends to be stable: compensation between damping of old errors (by successive analyses, within 4 days) and accumulation of recent errors (like a power series).

- Observation error contributions are significant, and model error contributions seem to be even larger.

- Extend this study to spatial correlation aspects, regional variations, etc. Possible use for calibration of model error representations.
Thank you for your attention