Estimation of model parameters using an evolutionary algorithm.

Pieter Houtekamer

June 3 2019

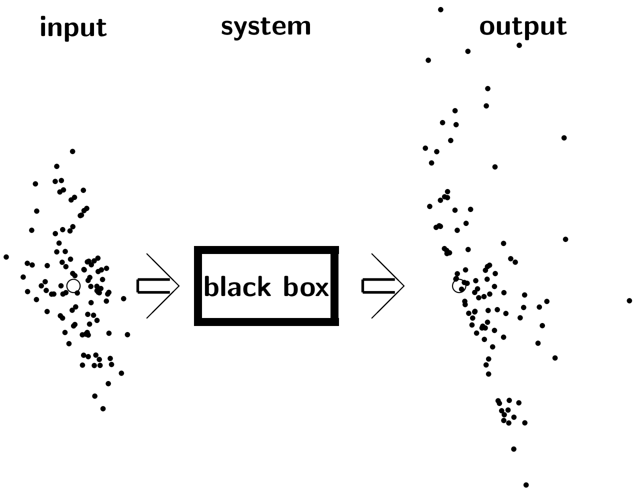
14th International EnKF workshop

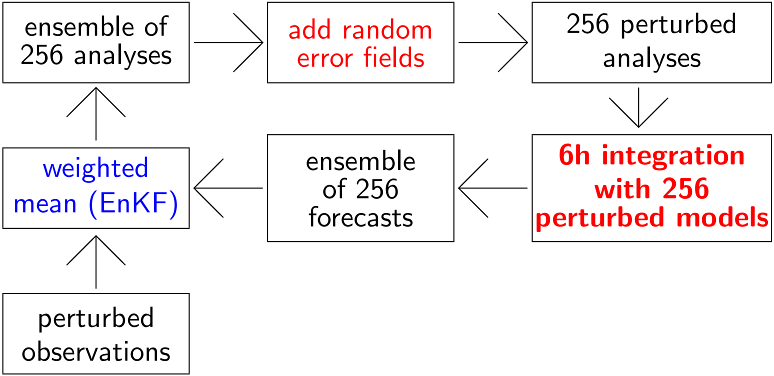
Voss, Norway

Collaborators: Xingxiu Deng, Bin He, Dominiques Jacques, Ron McTaggart-Cowan, Leo Separovic, Paul Vaillancourt, Ayrton Zadra

Contents

The Monte Carlo method





All inputs and output are in the form of ensembles. Every 6h, the new observations are perturbed randomly. Isotropic random fields are added to the analyses and different configurations of the model physics are used.

Issues

1. lack of spread
   When the error sources are sampled randomly, the spread of the output ensemble will be representative of the uncertainty in the best estimate (the ensemble mean). Unfortunetaly, the system has unknown sources of error. To obtain a realistic amount of spread, we need to add random error fields of which we cannot explain the origin.
2. ad hoc nature
   To sample model uncertainty, different parameter values or parameterization schemes are used by different members of the ensemble. Currently, the selection of model configurations is **ad hoc, difficult to maintain, and not connected with model development**.

Joint state and parameter estimation

In theory, an EnKF can estimate the model state and the model parameters simultaneously.

However, the state vector is evolving rapidly whereas model parameters have an absolute global value that may only evolve on climate time scales. To estimate the local weather conditions, an EnKF uses local observations. To estimate model parameters, localization procedures would appear inappropriate.

The number of model parameters is small and a long time period can be used in the estimation.

To estimate the state vector, we use the EnKF algorithm.

To estimate the model parameters, we use a particle filter that uses the ensembles provided by the EnKF.

Ensemble Prediction and Parameter Estimation System (EPPES)

Laine et al. (QJRMS, 2012) and Järvinen et al (QJRMS, 2012) propose the EPPES system. Here, an operational Ensemble Prediction System serves to test changes to parameter values with an evolutionary algorithm.

In Canada, the operational global medium-range ensembles consist of only 20 members. However, the operational global EnKF uses 256 members with a 6h cycle. This permits using 1024 integrations per day for parameter estimation.

In this talk, the output of the global EnKF will thus be used for the parameter estimation.

Evolutionary algorithm for model configurations a.k.a. particle filter for model parameters

1. Select 256 configurations,
2. Use the EnKF for one day to provide 4 sets of 256 background trajectories,
3. Use the about 3 000 0000 observations used by the EnKF to compute an ensemble score,
4. Over at most N=32 iterations:
   1. find the (bad) member that, when removed, improves the score the most,
   2. find the (good) member that, when removed, degrades the score the most,
   3. verify that replacement improves the score,
   4. remplace the parameters of the bad member by those of the good member while adding a small perturbation to the model parameters,
5. continue at point 2.

CRPS

To mesure the quality of the ensemble $x_i, i = 1, \cdots, N_{ens}$, given an observation y_k, we can use the CRPS (Continuous Ranked Probability Score, Gneiting et Raftering 2007, Zamo et Naveau 2018):

$$CRPS(x, y_k) = \frac{1}{N_{ens}} \sum_{i=1}^{N_{ens}} \mid x_i - y_k \mid - \frac{1}{2N_{ens}^2} \sum_{i,j=1}^{N_{ens}} \mid x_i - x_j \mid$$

For more robust results, we use the observations $y_k, k = 1, \cdots, N_{obs}$ used by the EnKF during one day. The std dev of the observation σ_k is used for normalization:

$$CRPS(x) = \frac{1}{N_{obs}} \sum_{k=1}^{N_{obs}} CRPS(x, y_k)/\sigma_k$$

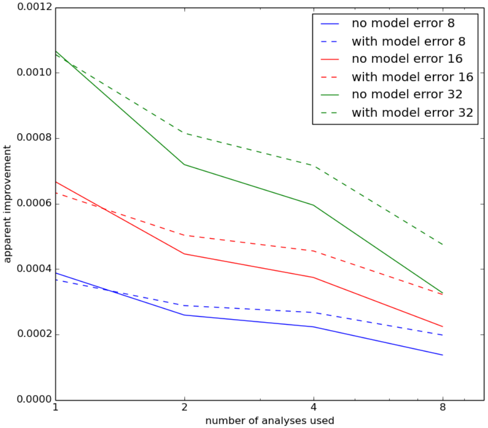
CRPS and member selection

With the CRPS, we associate a score $\mathcal{J}(1, 2, 3, \cdots, N_{ens})$ with a set of members $1, 2, 3, \cdots, N_{ens}$.
For one optimization step, we determine in a leave-one-out manner:

$$
\begin{aligned}
\mathcal{J}_1 &= \mathcal{J}(2, 3, \cdots, N_{ens}) \\
\mathcal{J}_2 &= \mathcal{J}(1, 3, 4, \cdots, N_{ens}) \\
\vdots &= \vdots \\
\mathcal{J}_{N_{ens}} &= \mathcal{J}(1, 2, \cdots, N_{ens} - 1)
\end{aligned}
$$

to obtain the worst performing member i_{bad} and the best performing member i_{good}. The procedure is continued with the set: $1, 2, i_{bad} - 1, i_{good}, i_{bad} + 1, \cdots, i_{good}, \cdots, N_{ens}$ where a small perturbation is added to the set of good parameters.

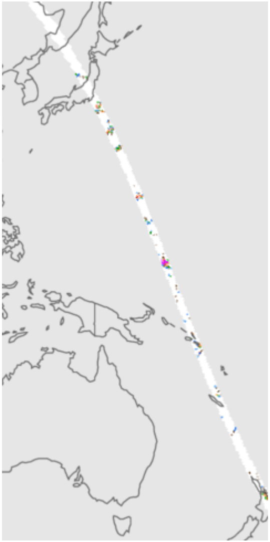
CRPS and observability



blue: replace 8 members
red: replace 16 members
green: replace 32 members
solid lines: all models are actually the same.
dashed lines: experiment with truly different configurations.

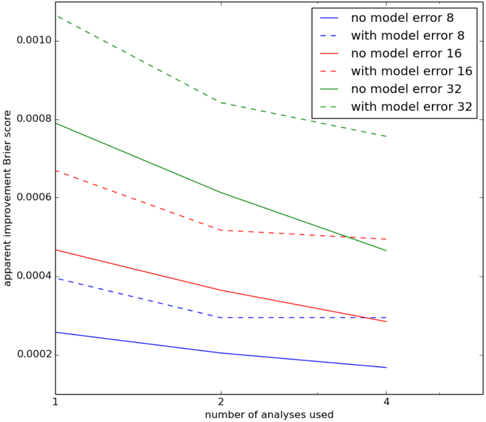
By construction, replacement of members can only improve scores. When using a 6h period, there is no observability for model parameters.

GPM-DPR observations by NASA obtained at CMC by Dominik Jacques



GPM = Global Precipitation Measurement
DPR = Dual-frequency Precipitation Radar
Swath Width = 125 km
Resolution = 5 km
Observations : precipitation rate in mm/h converted to the 39 km resolution Yin-Yang model grid.
Brier Score for a threshold of 0.3 mm/h.
About 2000 cases of precipitation are available per day.

Brier score and observability



blue: replace 8 members

red: replace 16 members

green: replace 32 members

solid lines: all models are actually the same.

dashed lines: experiment with truly different configurations.

Even when using only a 6h period, it appears relatively easy to see ensemble deficiency with respect to precipitation.

Score for the optimisation

The improvements seen in the CRPS and Brier scores are both of order 0.001. The scores have been added together for additional robustness.

The evolutionary algorithm optimizes the score $\mathcal{J}(x)$:

$$\mathcal{J}(x) = CRPS(x) + Brier(x)$$

Concretely, the optimization is to find the rearranged set of ensemble members that performed best over the preceeding day. Rearrangement involves removing and duplicating the parameters used by members. Thus, well performing model parameters become more ubiquitous.

Once daily, 32 parameter sets are replaced and a member can serve only once in a replacement.

List of model options
Paul Vaillancourt and Ayrton Zadra (2019)

1. recenter around EnKF or EnVar [0,1],
2. algorithm for deep convection: {kfc2,kfc3},
3. trigger for kfc2/3: $[0.03, 0.08]$,
4. closure for shallow convection: {equilibrium,cape},
5. evaporate detrained condensate: {false,true},
6. updraft radius kfc2/3: [1300,1700],
7. updraft radius kfc2/3 over water: [800,1300],
8. critical phase blocking height: [0.0,0.5],
9. PBL scheme: {black62, turboujo, boujo},
10. reduction factor turbulent flux: [0.5,1.0],
11. Monin Obukhov length: [5,20],
12. stability function: {beljaars91,delage97},
13. radius for ice in radiation scheme: [15,35].

The recentering

The ensemble of EnKF analyses $x^a_{i,\mathrm{EnKF}}$ is translated:

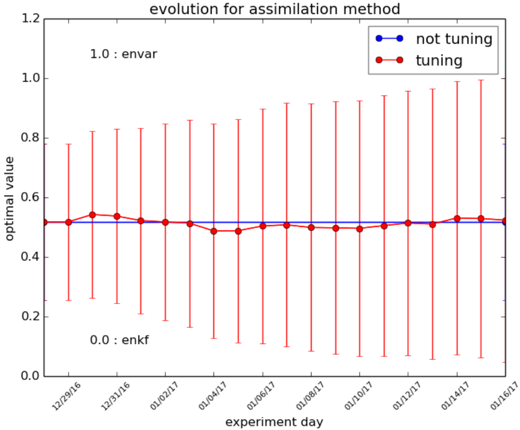
$$x^a_{i,hyb} = x^a_{i,\mathrm{EnKF}} + \gamma(x^a_{\mathrm{EnVar}} - \overline{x^a_{i,\mathrm{EnKF}}}) \tag{1}$$

Three interesting values:

- $\gamma = 0$: the EnKF analysis is not changed due to the EnVar,
- $\gamma = 1$: recenter around the EnVar analysis,
- $\gamma = 0.5$: as proposed by Penny (2014) and tested at ECMWF by Bonavita and Hamrud (2015), equal weight is given to the EnKF and the EnVar.

To obtain the best value of γ, we start the experiment with an ensemble of different values $\gamma_i, i = 1, \cdots, 256$ in the range [0,1]. The evolutionary algorithm will let the best value emerge.

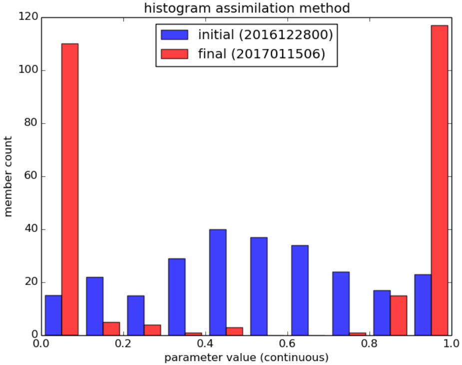
Convergence for the assimilation method (γ)



The mean and std dev for γ are given for the 2-week optimization period. The value zero gives all weight to the EnKF and one gives all weight to the EnVar.

The optimization algorithm increases the standard deviation of γ. This suggests that differences between the EnKF and EnVAR sample data assimilation uncertainty in a realistic manner.

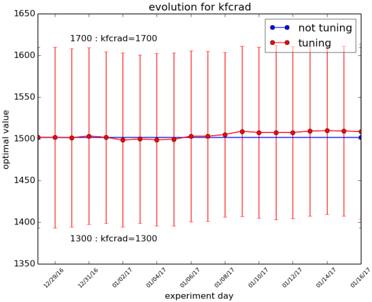
Histograms for the assimilation method (γ)

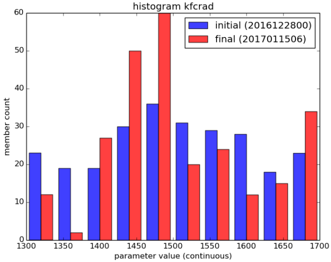


Histograms for γ at the beginning and at the end of the optimization.

The evolution is towards the distribution of the CMC-hybrid (Houtekamer, Buehner, De La Chevrotière, QJRMS, 2018) where half of the members is recentered on the EnKF and the other half on the EnVAR.

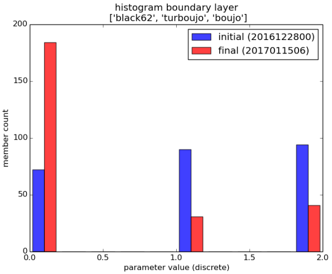
Parameter kfcrad (updraft radius for kfc over land)





The acceptable domain is [1300m, 1700m]. Prior deterministic experimentation suggested a value of 1500m. The optimization does not change the distribution.

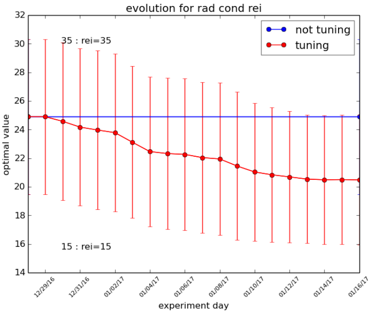
Scheme for the Planetary Boundary Layer black62, turboujo or boujo

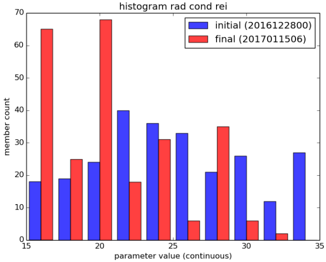


Histograms for the PBL scheme. The possible values are 0 for black62, 1 for turboujo and 2 for boujo. The target initial distribution is (33.3%,33.3%,33.3%). By chance, initially we have less members using black62.

The optimization favors the black62 scheme. Subsequent deterministic experimentation also showed slightly better results with black62 at 39 km resolution.

Parameter rad_cond_rei
radius for ice used in the radiation scheme





The acceptable domain is [15,35] μm. The deterministic models use 15 μm. We do indeed appear to evolve towards lower values.

Correlations between model parameters

Table: Correlation between parameter estimates for 2017 January 16. Only correlations with an absolute value above 0.4 are given.

	γ	kfc2/3	kfcradw	rad_cond_rei
kfc2/3	0.62			
kfcrad	-0.46	-0.58		
kfcradw	0.62	0.63		
trigger	-0.58	-0.42	*	
blocking	*	*	-0.51	
$L_{Obukhov}$	*	*	*	-0.43

Note: i) the correlation between the Kain-Fritsch parameters: kfc2/3, kfcrad, kfcradw and trigger.
ii) the undesirable correlation between γ and kfc2/3. Likely due to temperature biases, the EnKF does better with kfc2 and the EnVar does better with kfc3.

Conclusions with regard to the particle filter

- A modest number of parameters $O(10)$ can be optimized using the 256 members of the EnKF.
- After about two weeks (14x32=448 replacements), the distributions seem to stabilise.
- To confirm results, follow-up experiments with a deterministic model are needed. In medium-range forecasts, the black62 PBL scheme did well. However, the kfc3 scheme led to large biases after day 5.
- Correlations between parameters are found.

Outlook

- The theoretical basis for the methodology is weak. More work is needed.
- We are repeating the experiment with different options trying to correct a temperature bias near the surface.
- We would like to use the system to tune coupled land and atmosphere models.
- Perhaps the parameter distributions can be used in the medium-range ensemble using a stochastic parameter perturbation scheme.

Thank you