Wetter und Klima aus einer Hand

Deutscher Wetterdienst

Proposal Transition Density Choices in a Simple 1-D Particle Filter

Andreas Rhodin German Weather Service andreas.rhodin@dwd.de

Standard SIR Particle Filter

Approximate Pdf p by an ensemble ψ_k:

$$p(\psi) \approx \frac{1}{N}\sum_{k=1}^{N} \delta(\psi - \psi_k)$$

Cycled data assimilation in 3 steps:

1. Integrate the Kolmogorov equations in a Monte Carlo approach:

$$\psi_{b\,k}^{n} = f(\psi_k^{n-1}) + \beta_k \quad \text{with : model} f, \text{ random model error } \beta$$

2. Analysis step: assign weights $w_{a\,k}$ according to observation error pdf p_o.

$$p_a(\psi) \approx \sum_k w_{a\,k}\,\delta(\psi - \psi_{b\,k}) \quad \text{with : } w_{a\,k} = \frac{p_o(\psi_{b\,k})}{\sum_k p_o(\psi_{b\,k})}$$

3. Resampling step (not considered here): replace $w_{a\,k}$, $\psi_{b\,k}^n$ by ψ_k^n with equal weights.

Setup

p_b Gaussian background pdf, $\sigma_b = 1$, $\overline{\psi_b} = 3$
p_o Gaussian observation pdf, $\sigma_o = 1$, $\overline{\psi_o} = 7$
p_a Gaussian analysis pdf, $\sigma_a = .7071$, $\overline{\psi_a} = 5$

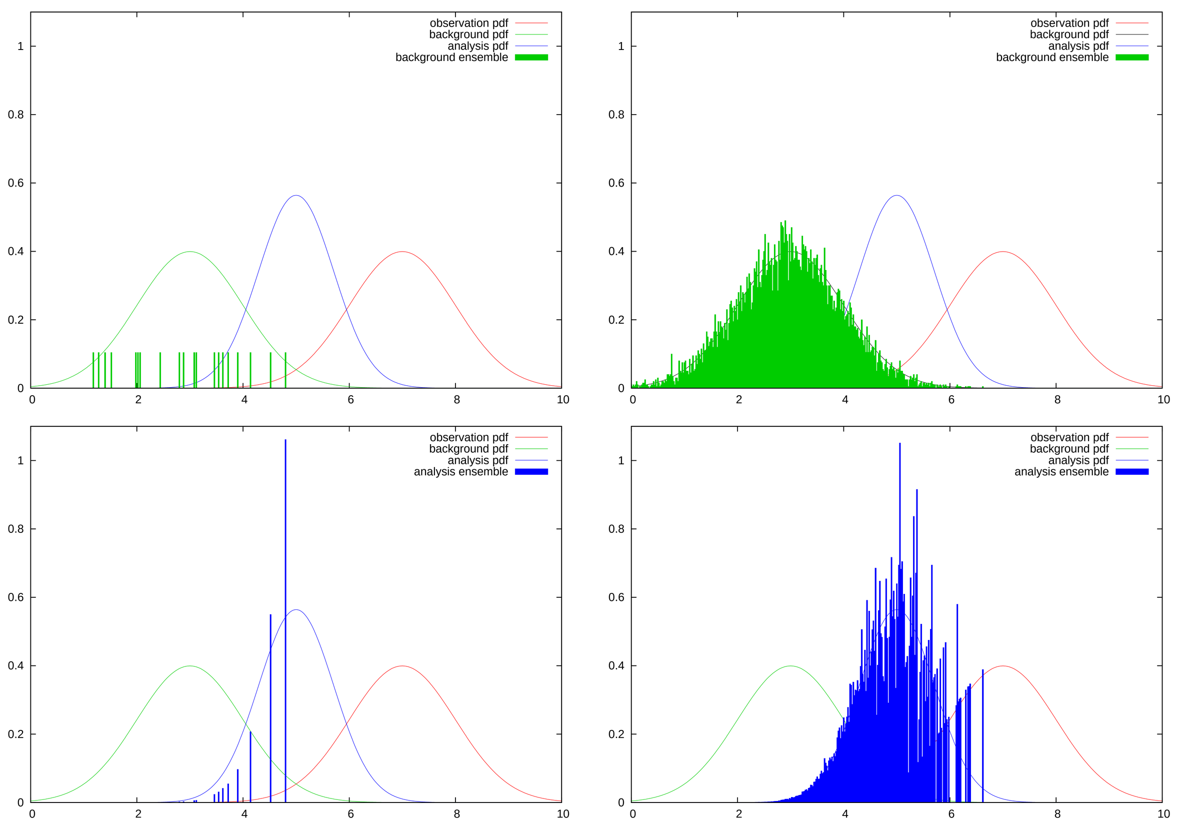
f Persistence model

Visualisation

Steps 1,2 of the data assimilation cycle

top: Background ensemble $\psi_{b\,k}$
bottom: Analysis ensemble $w_{a\,k}\,\psi_{b\,k}$

left: Ensemble size $N = 20$
right: Ensemble size $N = 10\,000$



Optimal Proposal Transition Density

Method:

Modify random perturbation term [Leeuwen 2010]:

$$\psi_{b\,k}^{n} = f(\psi_k^{n-1}) + (\hat{\beta} + K\,[o^n - \psi^{n-1}])$$

Perturbation from proposal transition density p_q, not model error pdf p_m.
Leads to modified ensemble representation of background pdf:

$$p_b(\psi) \approx \sum_k w_{b\,k}\,\delta(\psi - \psi_{b\,k}) \quad \text{with : } w_{b\,k} = \frac{p_m(\psi_k^n|\psi_k^{n-1})}{p_q(\psi_k^n|\psi_k^{n-1}, d^n)}$$

Goal:

Almost equal weights $w_k = w_{a\,k}\,w_{b\,k}$ in the analysis ensemble.

Optimal Choice:

$\hat{\beta}$ from Gaussian pdf with specified stdev $\sigma_q = \frac{1}{\sigma_m^{-1} + \sigma_o^{-1}}$.

Kalman Gain term $K = \left(1 + \frac{\sigma_o^2}{\sigma_{b_K}^2}\right)^{-1}$ with $\sigma_{b_K} = \sigma_m$

Idealistic Setup

Large model uncertainty $\sigma_m \approx \sigma_b$:

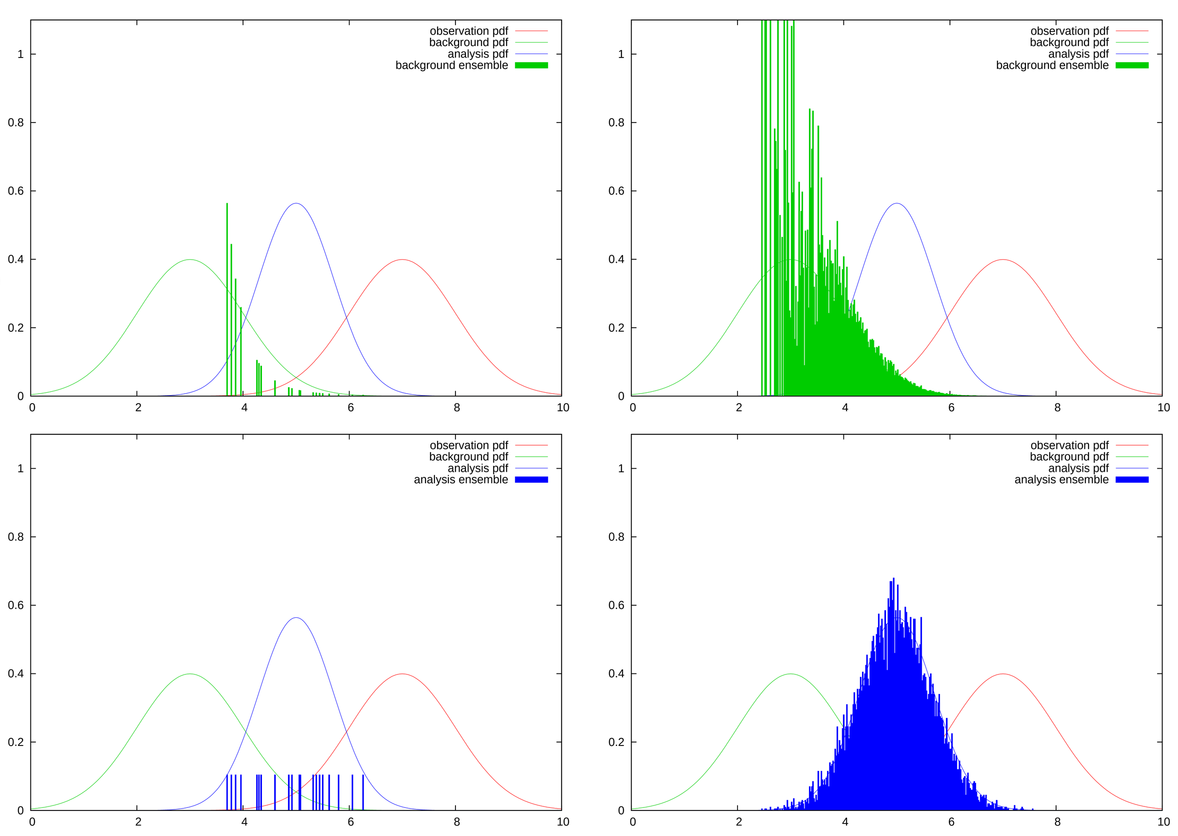
p^{n-1} Delta Previous pdf, $\sigma_p = 0$, $\overline{\psi_b} = 3$
p_m Gaussian model error pdf, $\sigma_m = 1$, $\overline{\psi_m} = 0$
p_b Gaussian background pdf, $\sigma_b = 1$, $\overline{\psi_b} = 3$

p_q Gaussian proposal pdf, $\sigma_q = .7071$
Kalman gain background error, $\sigma_{b_K} = 1$

Visualisation

top: Background ensemble $w_{b\,k}\,\psi_{b\,k}$
bottom: Analysis ensemble $w_{a\,k}\,w_{b\,k}\,\psi_{b\,k}$

left: Ensemble size $N = 20$
right: Ensemble size $N = 10\,000$



Realistic Setup

Above idealistic setup is highly unrealistic.
Realistic setup: **Moderate model uncertainty $\sigma_m < \sigma_b$:**

p^{n-1}	Gaussian Prior pdf,	$\sigma^{n-1} = .7071$
p_m	Gaussian model error pdf,	$\sigma_m = .7071$
p_b	Gaussian background pdf,	$\sigma_b = 1$
p_q	Gaussian proposal pdf,	$\sigma_q = .5773$
	Kalman gain background error,	$\sigma_{b_K} = .7071$

Enforce Equal Weights

Same setup as above, but:

- Modify proposal density for ensemble members with large weight:
  - multiply p_q by small factor ϵ in the center of the distribution
  - multiply p_q by by factor > 1 in the tails of the distribution
- Effect:
  - practically no ensemble members in the center of the distribution p_q
  - all ensemble members in the tail of the distribution p_q
  - reduced weights in the tails: equal weights for all members
  - **But:** Sampling of incorrect pdf for $N << 1/\epsilon$
    Even worse statistical sampling for $N >> 1/\epsilon$

Visualisation

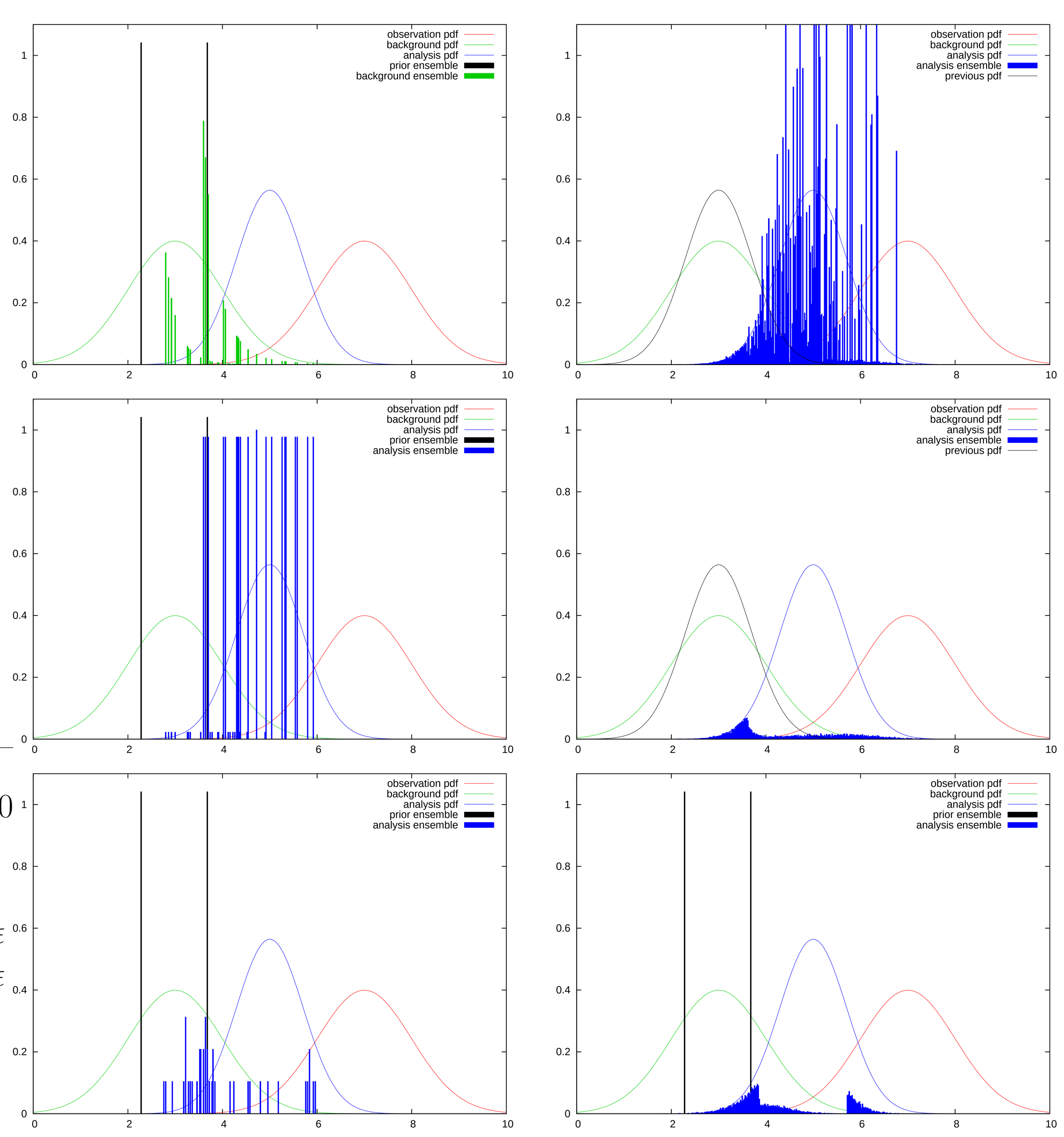
left: **Realistic setup**

For better visual assessment we chose a 'bimodal' previous ensemble ψ^{n-1} with ensemble size $N = 40$

all: **Previous ensemble** ψ_k^{n-1}
top: Background ensemble $w_{b\,k}\,\psi_{b\,k}$
center: Analysis ensemble $w_{a\,k}\,w_{b\,k}\,\psi_{b\,k}$
bottom: Analysis ensemble Enforce Equal Weights

right: **Enforce Equal weights** $N = 10\,000$

bottom: Analysis ensemble 'bimodal' ψ^{n-1}, $N \ll 1/\epsilon$
center: Analysis ensemble Gaussian ψ^{n-1}, $N \ll 1/\epsilon$
top: Analysis ensemble Gaussian ψ^{n-1}, $N \gg 1/\epsilon$



Assessment of Statistical Accuracy

The analysis ensemble provides properties of the state, e.g:

$$\overline{g(\psi)} = \int g(\psi)\,p(\psi)\,d\psi \approx \sum_k w_k\,g(\psi_k)$$

As an example we take the ensemble mean:

$$\overline{\psi} = \int \psi\,p(\psi)\,d\psi \approx \sum_k w_k\,\psi_k$$

With the statistical uncertainty (stdev):

$$\sigma_{\overline{\psi}} = \frac{1}{\sqrt{N}}\sigma_{w\psi}$$

Results

σ_m	σ_q	σ_{b_K}	$\sigma_{w\psi}$	comment
1.	1.	0.	3.493	no proposal
1.	0.707	1.	0.707	idealistic setup
0.707	0.577	0.707	1.763	realistic setup
0.707	0.577	0.707	14.833	Eq.Wgh. $N >> 1/\epsilon$

$\sigma_{w\psi}$ is derived for $N = 100\,000$.

Conclusions

- Idealistic setup ($\sigma_m \approx \sigma_b$): using a proposal transition density helps a lot
- Realistic setup ($\sigma_m < \sigma_b$) : using a proposal transition density helps a bit
- Enforcing almost equal weights: only converges to correct pdf for large ensemble size
- Statistical accuracy of the results should be assessed

References: van Leeuwen PJ. (2010, Nonlinear data assimilation in geosciences: an extremely efficient particle filter, QJRMS 136 p1991-1999, DOI:10.2001/qj.699