What determines the success of ensemble data assimilation methods?

Femke C. Vossepoel

Max Frei and Nikolaj Mücke

TU Delft Available from

https://github.com/nmucke/non_gaussian_data_assim.git

Outline

- How data assimilation methods are evaluated
- Reconstruction of posterior pdf: Kullback-Leibler divergence
- State- and parameter estimation
  - Entropy
  - χ^2 value
- Forecasting
- Summary

We start from Bayes' theorem

$$f(\mathbf{z}|\mathbf{d}) \propto f(\mathbf{d}|\mathbf{z}) f(\mathbf{z}). \tag{1}$$

Gaussian prior and likelihood gives

$$f(\mathbf{z}|\mathbf{d}) \propto \exp\left\{-\mathcal{J}(\mathbf{z})\right\}, \tag{2}$$

with cost function defined as

$$\mathcal{J}(\mathbf{z}) = \frac{1}{2}\left(\mathbf{z} - \mathbf{z}^{\mathrm{f}}\right)^{\mathrm{T}} \mathbf{C}_{zz}^{-1} \left(\mathbf{z} - \mathbf{z}^{\mathrm{f}}\right) + \frac{1}{2}\left(\mathbf{g}(\mathbf{z}) - \mathbf{d}\right)^{\mathrm{T}} \mathbf{C}_{dd}^{-1} \left(\mathbf{g}(\mathbf{z}) - \mathbf{d}\right). \tag{3}$$

How data assimilation methods are being evaluated

- Fit to observations (RMSE)
- Forecasting skill
- Fit of estimated posterior probability density function to ’true’ posterior pdf

Kullback-Leibler divergence

Outline

- How data assimilation methods are evaluated
- Reconstruction of posterior pdf: Kullback-Leibler divergence
- State- and parameter estimation
  - Entropy
  - χ^2 value
- Forecasting
- Summary

State- and parameter estimation

- xxx

Entropy

χ^2 value

Example with (coupled?) multiscale Kuramoto-Sivashinsky model

$$\tau_a \frac{\partial A}{\partial t} = -\frac{\lambda_a}{2}\frac{\partial A^2}{\partial x} - \lambda_a^2 \frac{\partial^2 A}{\partial x^2} - \frac{\lambda_a^4}{2}\frac{\partial^4 A}{\partial x^4} + \alpha_{oa}(O - A), \tag{4}$$

$$\tau_o \frac{\partial O}{\partial t} = -\frac{\lambda_o}{2}\frac{\partial O^2}{\partial x} - \lambda_o^2 \frac{\partial^2 O}{\partial x^2} - \lambda_o^4 \frac{\partial^4 O}{\partial x^4} + \omega_{ao}(A - O), \tag{5}$$

Scale difference by setting domain size of 32 for Atmos and 256 for Ocean on a 1024 nodes grid.
$\tau_o > \tau_a$: ocean evolves on a slower scale than the atmosphere.

We apply ESMDA in synthetic data-assimilation experiments. For a comparison of ES, IES, and ESMDA results, see **?**

Outline

- How data assimilation methods are evaluated
- Reconstruction of posterior pdf: Kullback-Leibler divergence
- State- and parameter estimation
  - Entropy
  - χ^2 value
- Forecasting
- Summary

Forecasting

Summary

- Summary or conclusions