The Ensemble Kalman Filter:
Theoretical formulation and practical implementation

Geir Evensen

Norsk Hydro Research Centre, Bergen, Norway
Based on Evensen 2003, Ocean Dynamics, Vol 53, No 4

The Ensemble Kalman Filter (EnKF)

- Represents error statistics using an ensemble of model states.
- Evolves error statistics by ensemble integrations.
- “Variance minimizing” analysis scheme operating on the ensemble.

- Monte Carlo, low rank, error subspace method.
- Converges to the Kalman Filter with increasing ensemble size.
- Fully nonlinear error evolution, contrary to EKF.
- Assumption of Gaussian statistics in analysis scheme.

The error covariance matrix

Define ensemble covariances around the ensemble mean

$$\boldsymbol{P}^{\mathrm{f}} \simeq \boldsymbol{P}_{\mathrm{e}}^{\mathrm{f}} = \overline{(\boldsymbol{\psi}^{\mathrm{f}} - \overline{\boldsymbol{\psi}^{\mathrm{f}}})(\boldsymbol{\psi}^{\mathrm{f}} - \overline{\boldsymbol{\psi}^{\mathrm{f}}})^{\mathrm{T}}}$$

$$\boldsymbol{P}^{\mathrm{a}} \simeq \boldsymbol{P}_{\mathrm{e}}^{\mathrm{a}} = \overline{(\boldsymbol{\psi}^{\mathrm{a}} - \overline{\boldsymbol{\psi}^{\mathrm{a}}})(\boldsymbol{\psi}^{\mathrm{a}} - \overline{\boldsymbol{\psi}^{\mathrm{a}}})^{\mathrm{T}}}$$

- The ensemble mean $\overline{\boldsymbol{\psi}}$ is the best-guess.
- The ensemble spread defines the error variance.
- The covariance is determined by the smoothness of the ensemble members.
- A covariance matrix can be represented by an ensemble of model states (not unique).

Dynamical evolution of error statistics

- Each ensemble member evolve according to the model dynamics which is expressed by a stochastic differential equation

$$d\boldsymbol{\psi} = \boldsymbol{f}(\boldsymbol{\psi})dt + \boldsymbol{g}(\boldsymbol{\psi})d\boldsymbol{q}.$$

- The probability density then evolve according to Kolmogorov's equation

$$\frac{\partial \phi}{\partial t} + \sum_i \frac{\partial (f_i \phi)}{\partial \psi_i} = \frac{1}{2} \sum_{i,j} \frac{\partial^2 \phi (\boldsymbol{g}\boldsymbol{Q}\boldsymbol{g}^T)_{ij}}{\partial \psi_i \partial \psi_j}.$$

- This is the fundamental equation for evolution of error statistics and can be solved using Monte Carlo methods.

Analysis scheme (1)

- Given an ensemble of model forcasts, $\boldsymbol{\psi}_j^{\mathrm{f}}$, defining forecast error covariance

$$\boldsymbol{P}^{\mathrm{f}} \simeq \boldsymbol{P}_{\mathrm{e}}^{\mathrm{f}} = \overline{(\boldsymbol{\psi}^{\mathrm{f}} - \overline{\boldsymbol{\psi}^{\mathrm{f}}})(\boldsymbol{\psi}^{\mathrm{f}} - \overline{\boldsymbol{\psi}^{\mathrm{f}}})^{\mathrm{T}}}.$$

- Create an ensemble of observations

$$\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{\epsilon}_j,$$

with

  - $\boldsymbol{d}$, the real observations,
  - $\boldsymbol{\epsilon}_j$, a vector of observation noise,
  - $\overline{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}} = \boldsymbol{R}_{\mathrm{e}} \simeq \boldsymbol{R}$.

Analysis scheme (2)

- Update each ensemble member according to

$$\boldsymbol{\psi}_j^{\mathrm{a}} = \boldsymbol{\psi}_j^{\mathrm{f}} + \boldsymbol{K}_{\mathrm{e}}(\boldsymbol{d}_j - \boldsymbol{H}\boldsymbol{\psi}_j^{\mathrm{f}}),$$

  where

$$\boldsymbol{K}_{\mathrm{e}} = \boldsymbol{P}_{\mathrm{e}}^{f}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}_{\mathrm{e}}^{f}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}_{\mathrm{e}})^{-1}.$$

- Thus the update of the mean becomes

$$\overline{\boldsymbol{\psi}^{\mathrm{a}}} = \overline{\boldsymbol{\psi}^{\mathrm{f}}} + \boldsymbol{K}_{\mathrm{e}}(\boldsymbol{d} - \boldsymbol{H}\overline{\boldsymbol{\psi}^{\mathrm{f}}}).$$

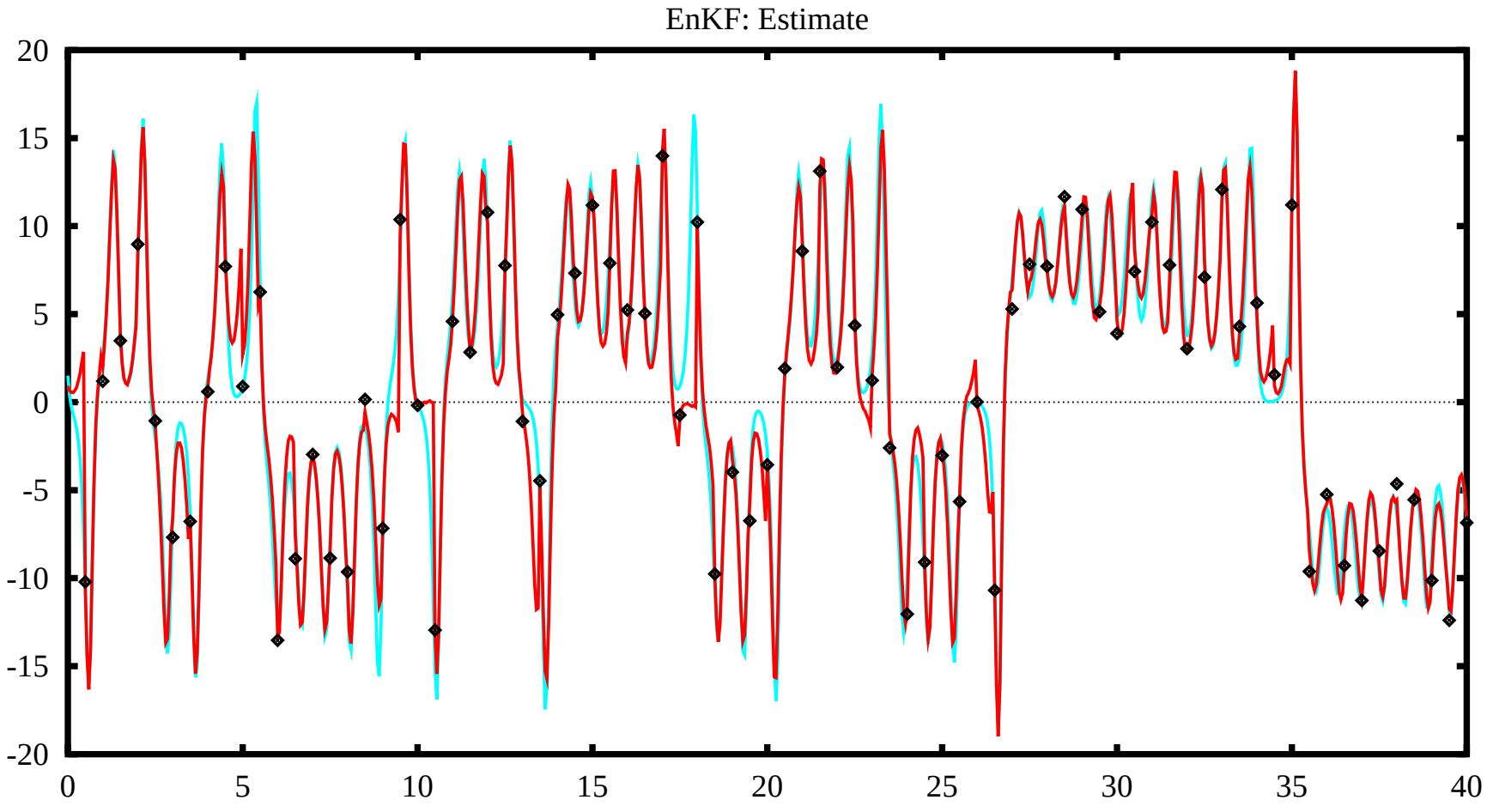
- The posterior error covariance becomes

$$\boldsymbol{P}_{\mathrm{e}}^{\mathrm{a}} = (\boldsymbol{I} - \boldsymbol{K}_{\mathrm{e}}\boldsymbol{H})\boldsymbol{P}_{\mathrm{e}}^{\mathrm{f}}.$$

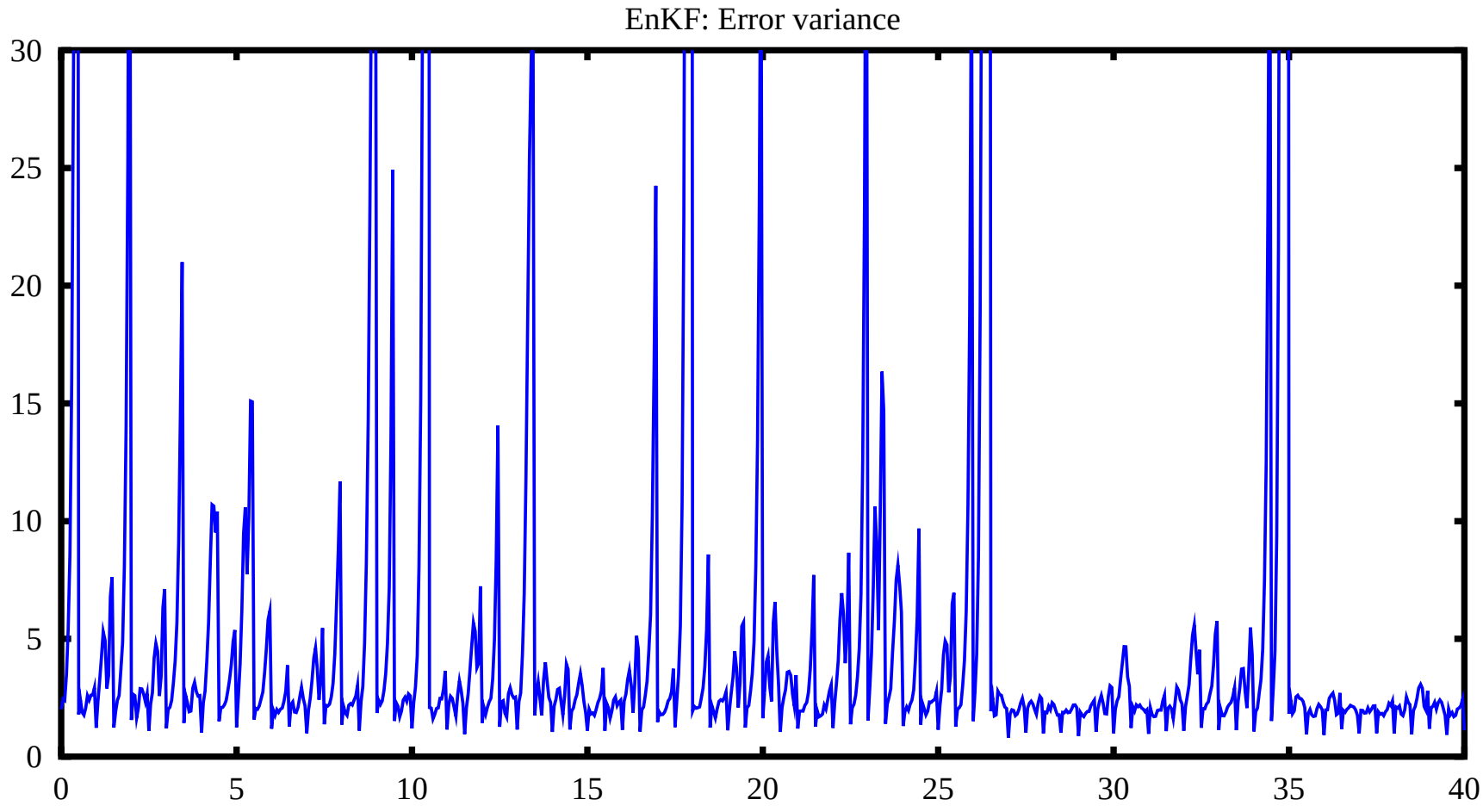
Example: Lorenz model

- Application with the chaotic Lorenz model
- Illustrates properties with higly nonlinear dynamical models.
- From Evensen (1997), MWR.

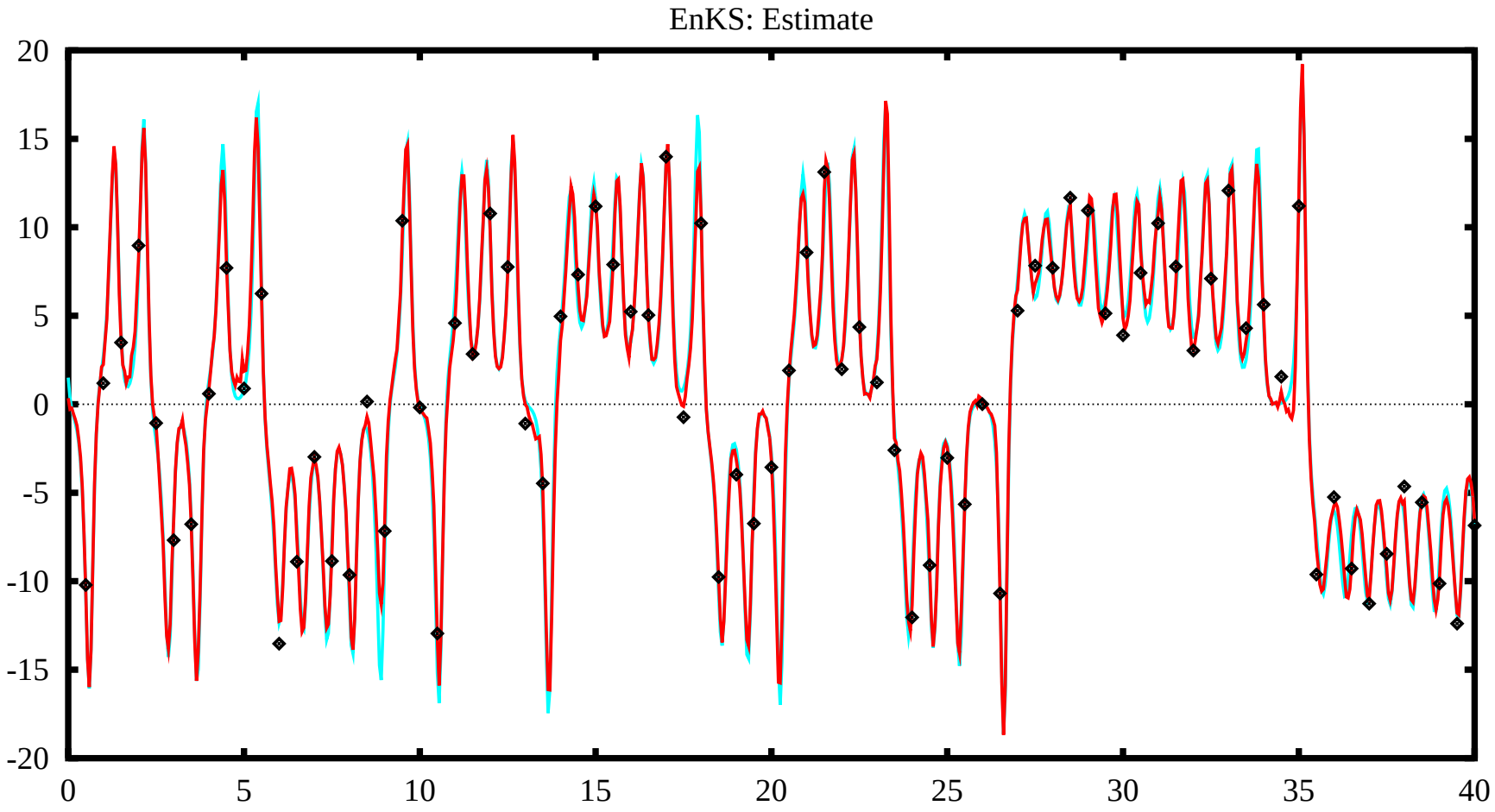
EnKF solution



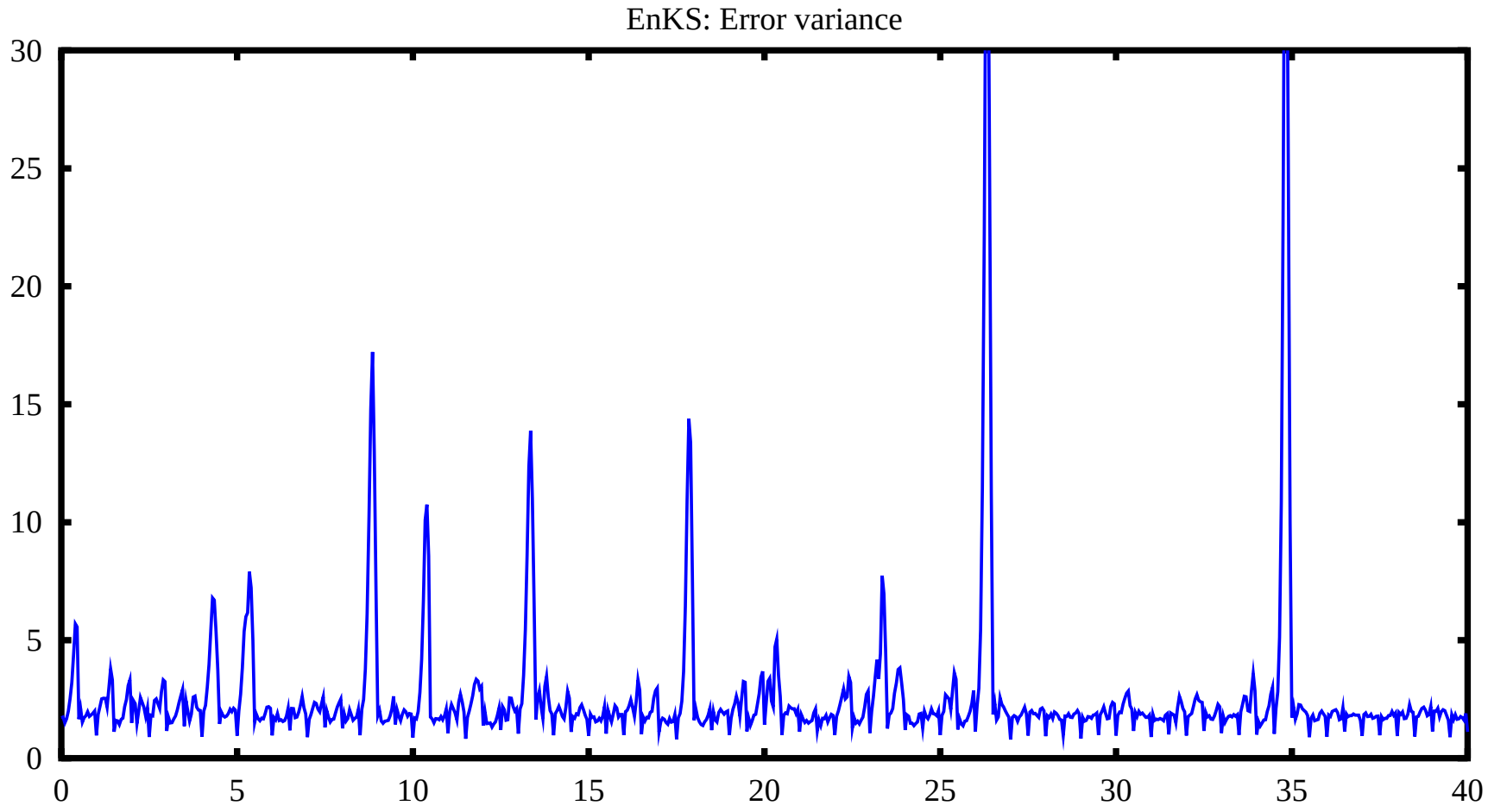
EnKF error variance



EnKS solution



EnKS error variance



Summary: Lorenz model

- The EnKF and EnKS works well with highly nonlinear dynamical models.
- There is no linearization in the evolution of error statistics.
- Methods using tangent linear or adjoint operators have problems with the Lorenz equations:
  - limited by the predictability time,
  - limited by the validity time of tangent linear operator.
- Can we expect the same to be true for high resolution ocean and atmosphere models?

Analysis equation (1)

- Define the ensemble matrix

$$\boldsymbol{A} = (\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_N) \in \Re^{n \times N}.$$

- The ensemble mean is (defining $\mathbf{1}_N \in \Re^{N \times N} \equiv 1/N$)

$$\overline{\boldsymbol{A}} = \boldsymbol{A}\mathbf{1}_N.$$

- The ensemble perturbations become

$$\boldsymbol{A}' = \boldsymbol{A} - \overline{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{I} - \mathbf{1}_N).$$

- The ensemble covariance matrix $\boldsymbol{P}_{\mathrm{e}} \in \Re^{n \times n}$ becomes

$$\boldsymbol{P}_{\mathrm{e}} = \frac{\boldsymbol{A}'(\boldsymbol{A}')^{\mathrm{T}}}{N-1}.$$

Analysis equation (2)

- Given a vector of measurements $\boldsymbol{d} \in \Re^m$, define

$$\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{\epsilon}_j, \quad j = 1, \ldots, N,$$

stored in

$$\boldsymbol{D} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_N) \in \Re^{m \times N}.$$

- The ensemble perturbations are stored in

$$\boldsymbol{E} = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \ldots, \boldsymbol{\epsilon}_N) \in \Re^{m \times N},$$

thus, the measurement error covariance matrix becomes

$$\boldsymbol{R}_{\mathrm{e}} = \frac{\boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}}{N-1}.$$

Analysis equation (3)

- The analysis equation can now be written

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{P}_{\mathrm{e}}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}_{\mathrm{e}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}_{\mathrm{e}})^{-1}(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{A}).$$

- Defining the innovations $\boldsymbol{D}' = \boldsymbol{D} - \boldsymbol{H}\boldsymbol{A}$ and using previous definitions:

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{A}'(\boldsymbol{H}\boldsymbol{A}')^{\mathrm{T}} \left((\boldsymbol{H}\boldsymbol{A}')(\boldsymbol{H}\boldsymbol{A}')^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\right)^{-1} \boldsymbol{D}'.$$

  i.e., analysis expressed entirely in terms of the ensemble

Analysis equation (4)

- Define $\boldsymbol{S} = \boldsymbol{H}\boldsymbol{A}'$ and $\boldsymbol{C} = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}$.
- Use $\boldsymbol{A}' = \boldsymbol{A}(\boldsymbol{I} - \mathbf{1}_N)$.
- Use $\mathbf{1}_N \boldsymbol{S}^{\mathrm{T}} \equiv \mathbf{0}$.

$$
\begin{aligned}
\boldsymbol{A}^{\mathrm{a}} &= \boldsymbol{A} + \boldsymbol{A}'\boldsymbol{S}^{\mathrm{T}} \left(\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}} \right)^{-1} \boldsymbol{D}' \\
&= \boldsymbol{A} + \boldsymbol{A}(\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{D}' \\
&= \boldsymbol{A} \left(\boldsymbol{I} + (\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{D}' \right) \\
&= \boldsymbol{A} \left(\boldsymbol{I} + \boldsymbol{S}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{D}' \right) \\
&= \boldsymbol{A}\boldsymbol{X}
\end{aligned}
\tag{1}
$$

Remarks on the analysis equation (1)

- Covariances only needed between observed variables at measurement locations ($\boldsymbol{H}\boldsymbol{P}_{\mathrm{e}} = \boldsymbol{S}\boldsymbol{A}'^{\mathrm{T}}$).
- $\boldsymbol{P}_{\mathrm{e}}$ never computed but indirectly used to determine $\boldsymbol{H}\boldsymbol{P}_{\mathrm{e}}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}}$.
- Analysis may be interpreted as:
  - combination of forecast ensemble members, or,
  - forecast pluss combination of covariance functions.
- Accuracy of analysis is determined by:
  - the accuracy of $\boldsymbol{X}$,
  - the properties of the ensemble error space.

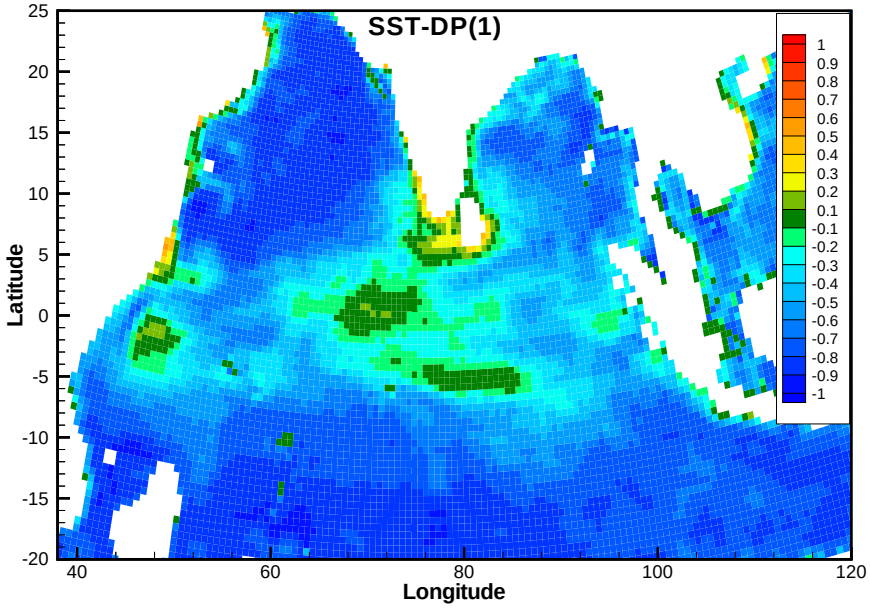
Remarks on the analysis equation (2)

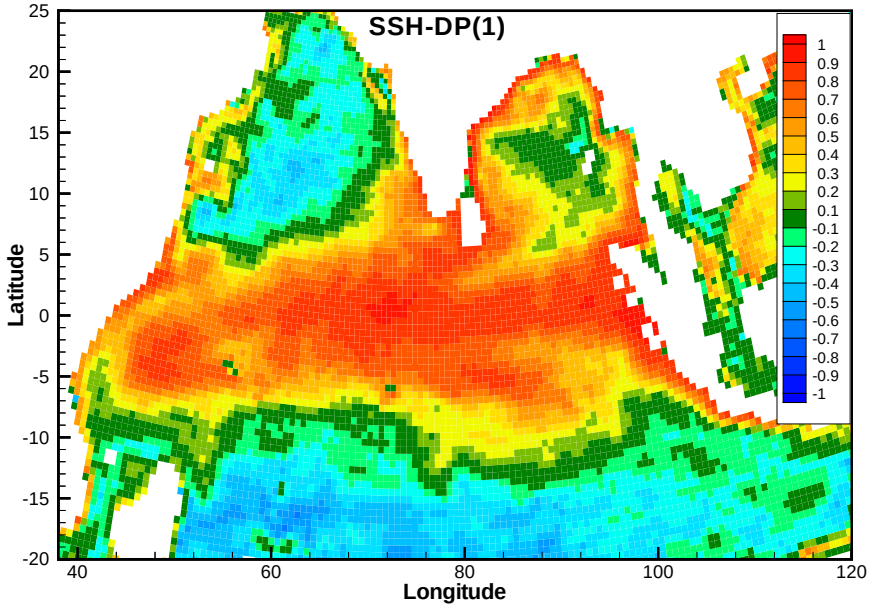
- For a linear model, any choice of $\boldsymbol{X}$ will result in an analysis which is also a solution of the model.
- Filtering of covariance functions introduces nondynamical modes in the analysis.

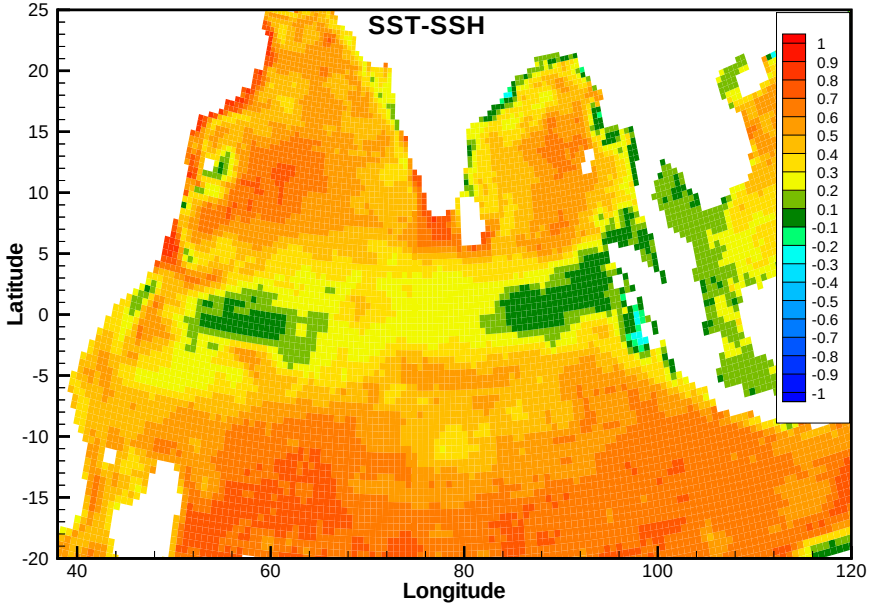
Examples of ensemble statistics

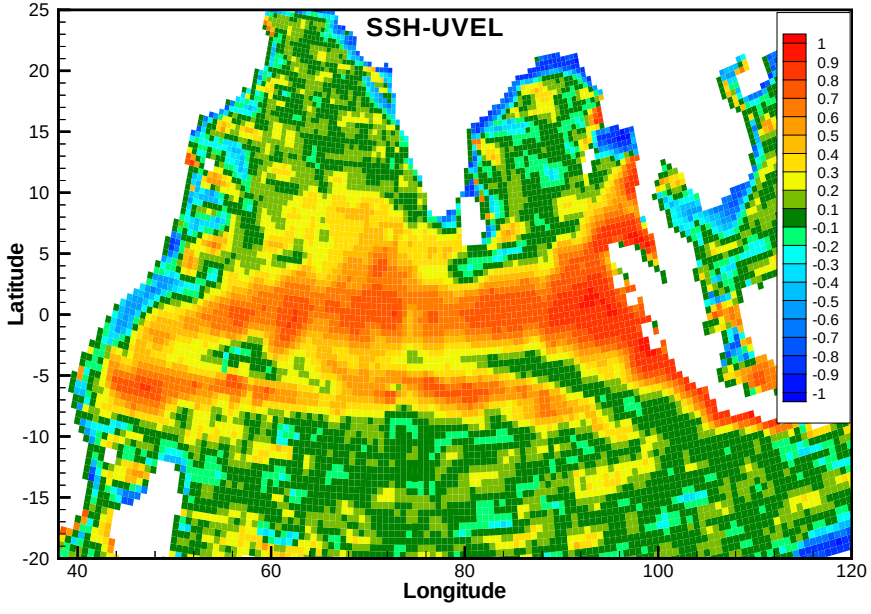
- Taken from Haugen and Evensen (2002), Ocean Dynamics.
- OGCM (MICOM) for the Indian Ocean.
- Assimilation of SST and SLA data.

Spatial correlations



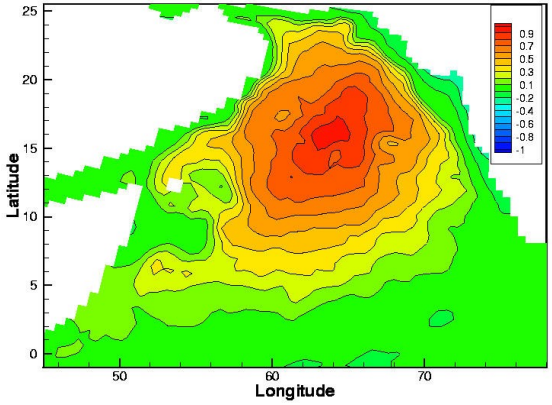




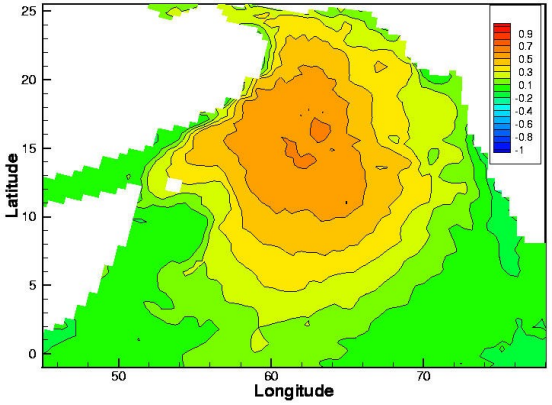


Correlation functions

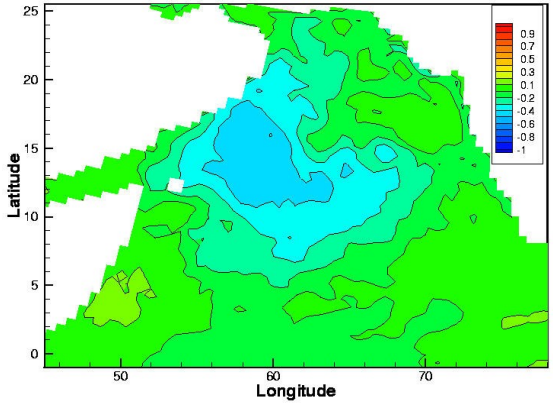
SSH–SSH



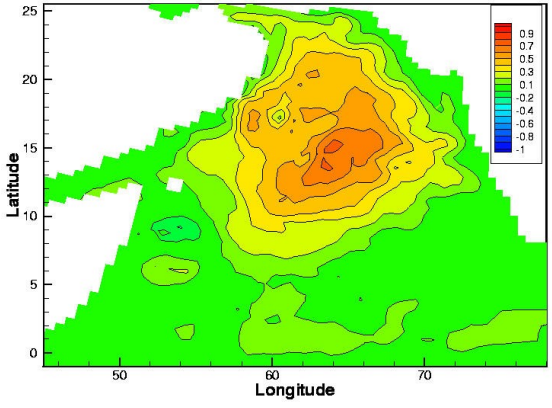
SSH–SST



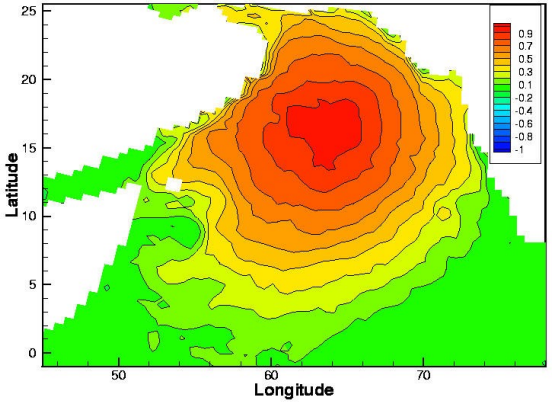
SSH–DP



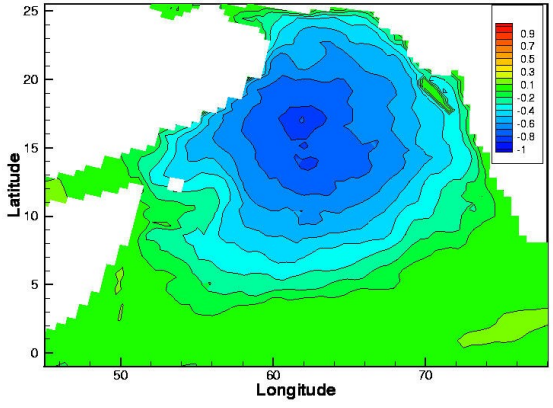
SST–SSH



SST–SST

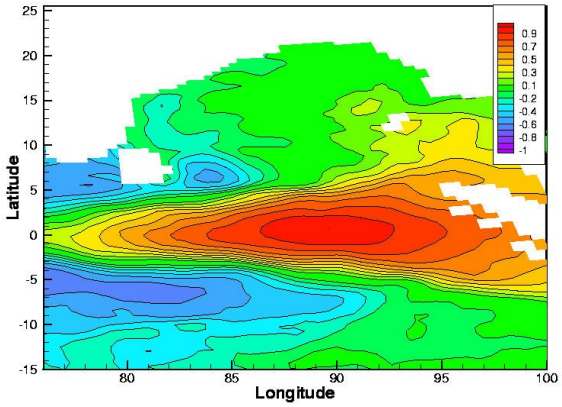


SST–DP

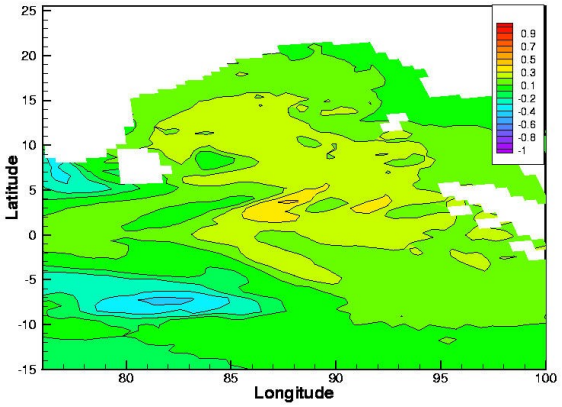


Correlation functions

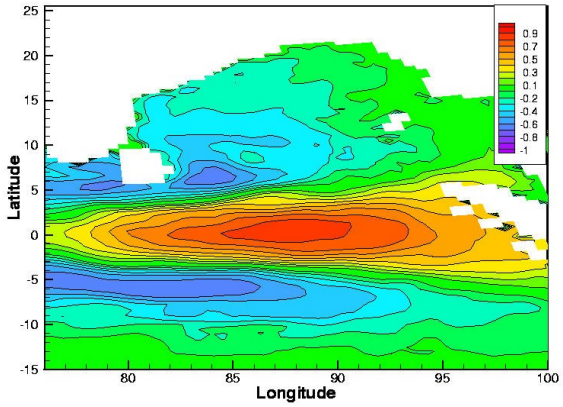
SSH–SSH



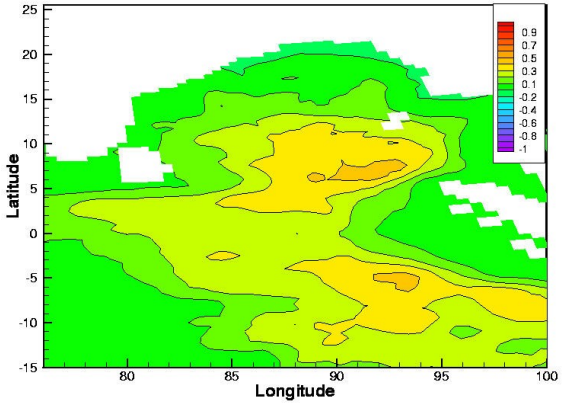
SSH–SST



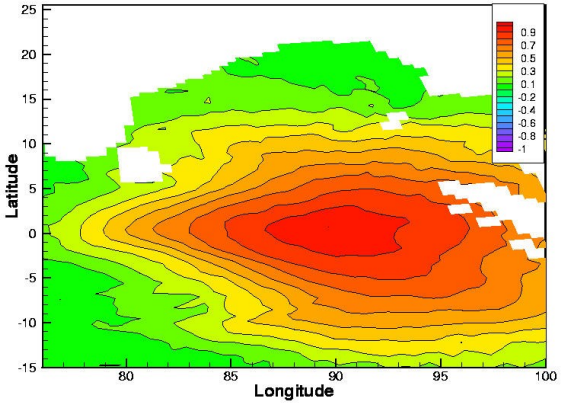
SSH–DP



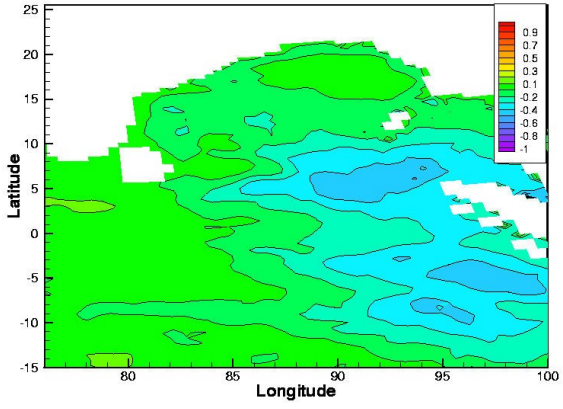
SST–SSH



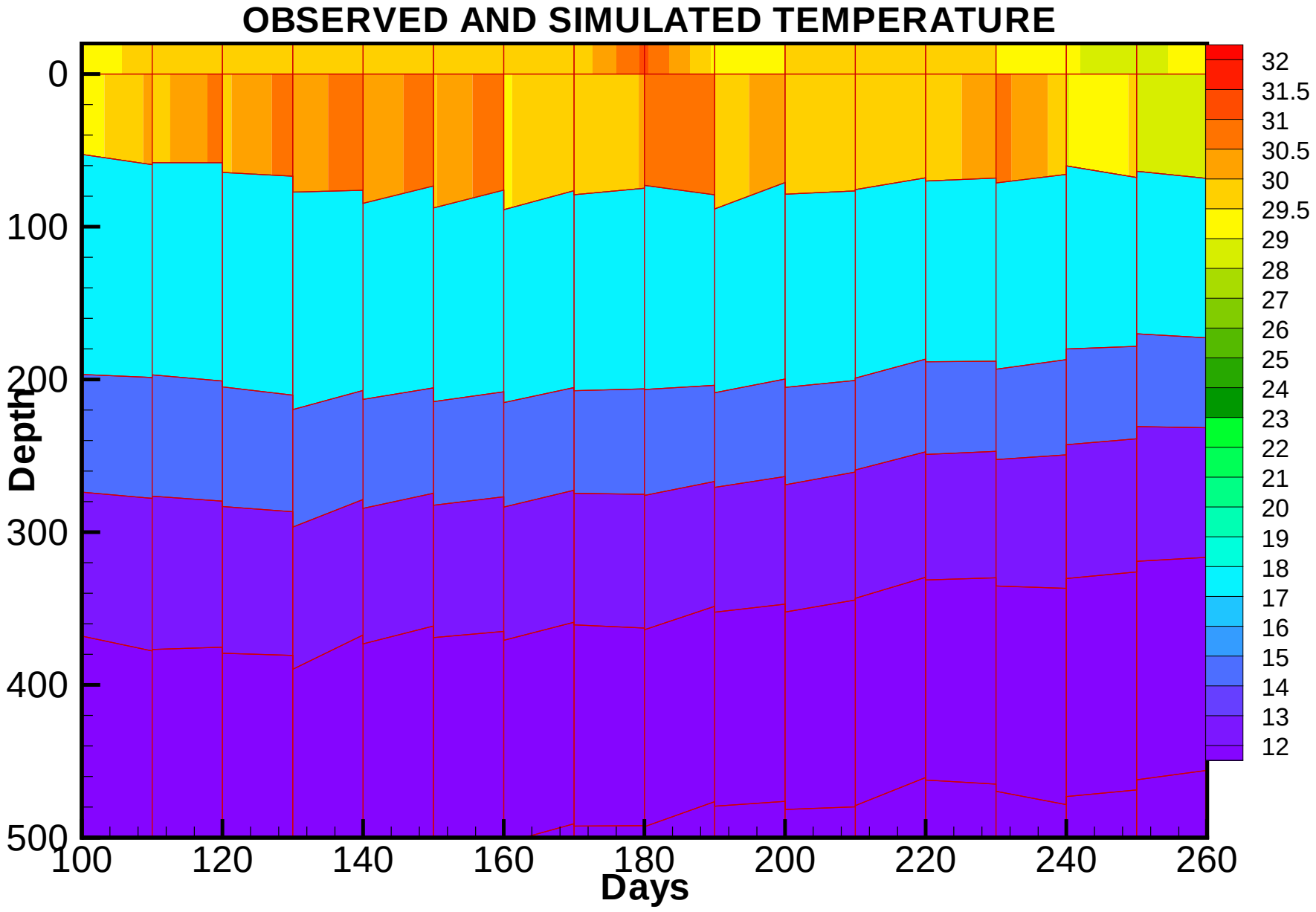
SST–SST



SST–DP



Time–Depth: Temperature



Ensemble Kalman Smoother (EnKS)

- Derived in Evensen and van Leeuwen (2000), MWR.
- Starts with EnKF solution.
- Computes updates backward in time;
  - sequentially for each measurement time,
  - using covariances in time,
  - no backward integrations.
- The analysis becomes for $t_{i-1} \leq t' < t_i \leq t_k$:

$$\boldsymbol{A}^{\mathrm{a}}_{\mathrm{EnKS}}(t') = \boldsymbol{A}_{\mathrm{EnKF}}(t') \prod_{j=i}^{k} \boldsymbol{X}(t_j)$$

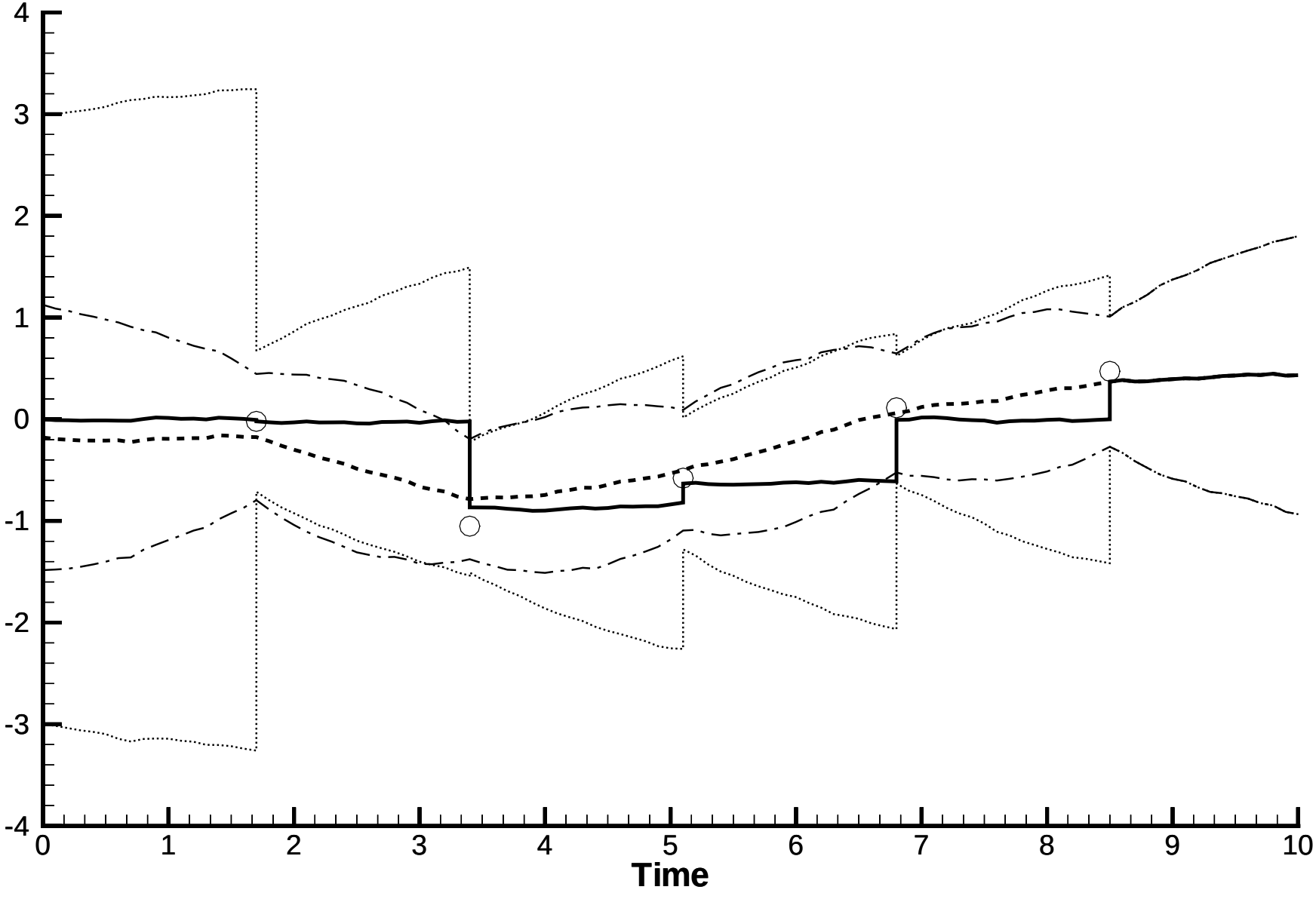
Time correlated model noise

- Most schemes assume white model noise in time.
- Augment model state to include time correlations

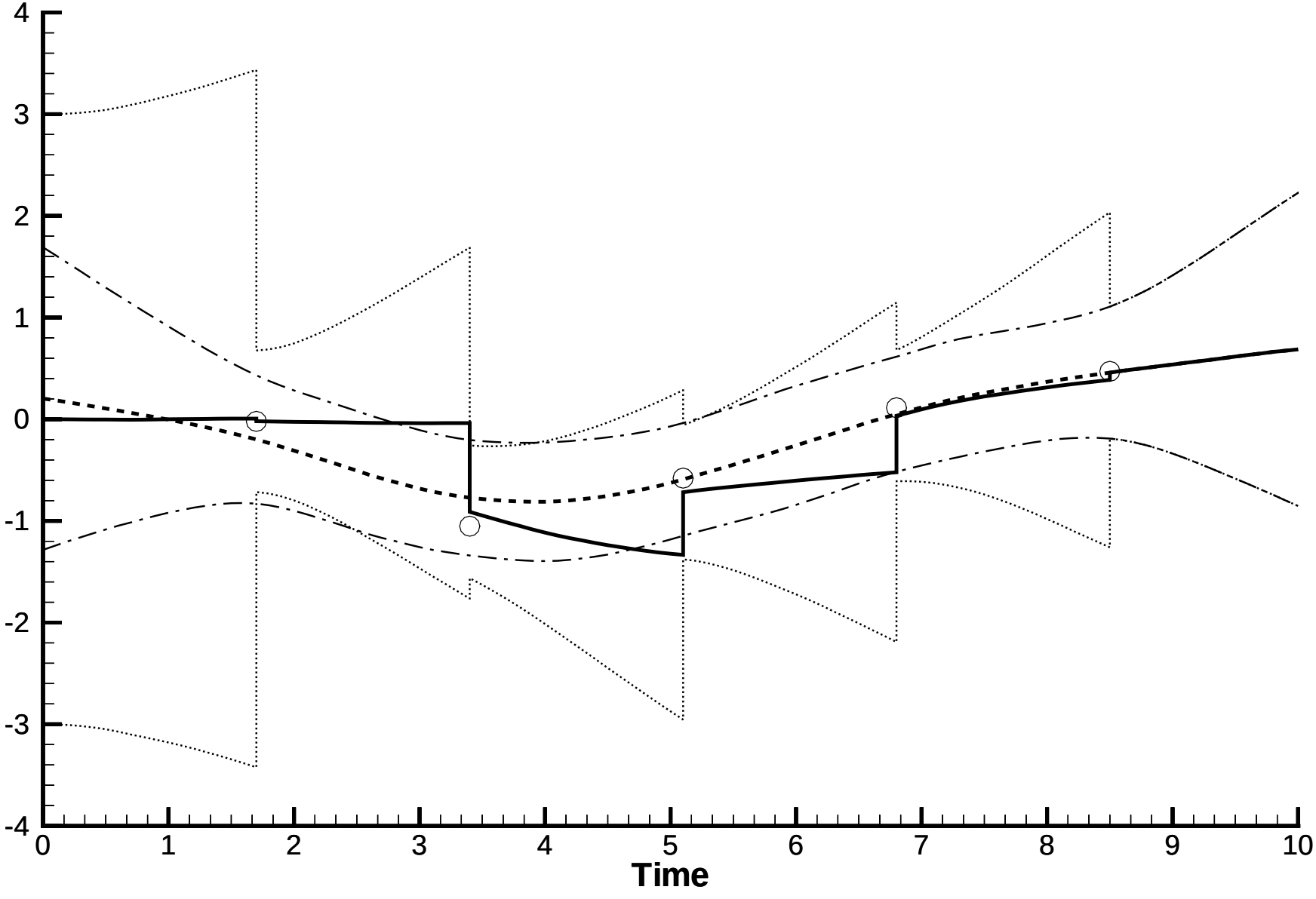
$$\begin{pmatrix} q_k \\ \psi_k \end{pmatrix} = \begin{pmatrix} \alpha q_{k-1} \\ \psi_{k-1} + \sqrt{\Delta t}\sigma\rho q_k \end{pmatrix} + \begin{pmatrix} \sqrt{1-\alpha^2} w_{k-1} \\ 0 \end{pmatrix}.$$

- White noise when $\alpha = 0$ and $\rho = 1$.

Results ($\alpha = 0$)



Results ($\alpha = 0.95$)



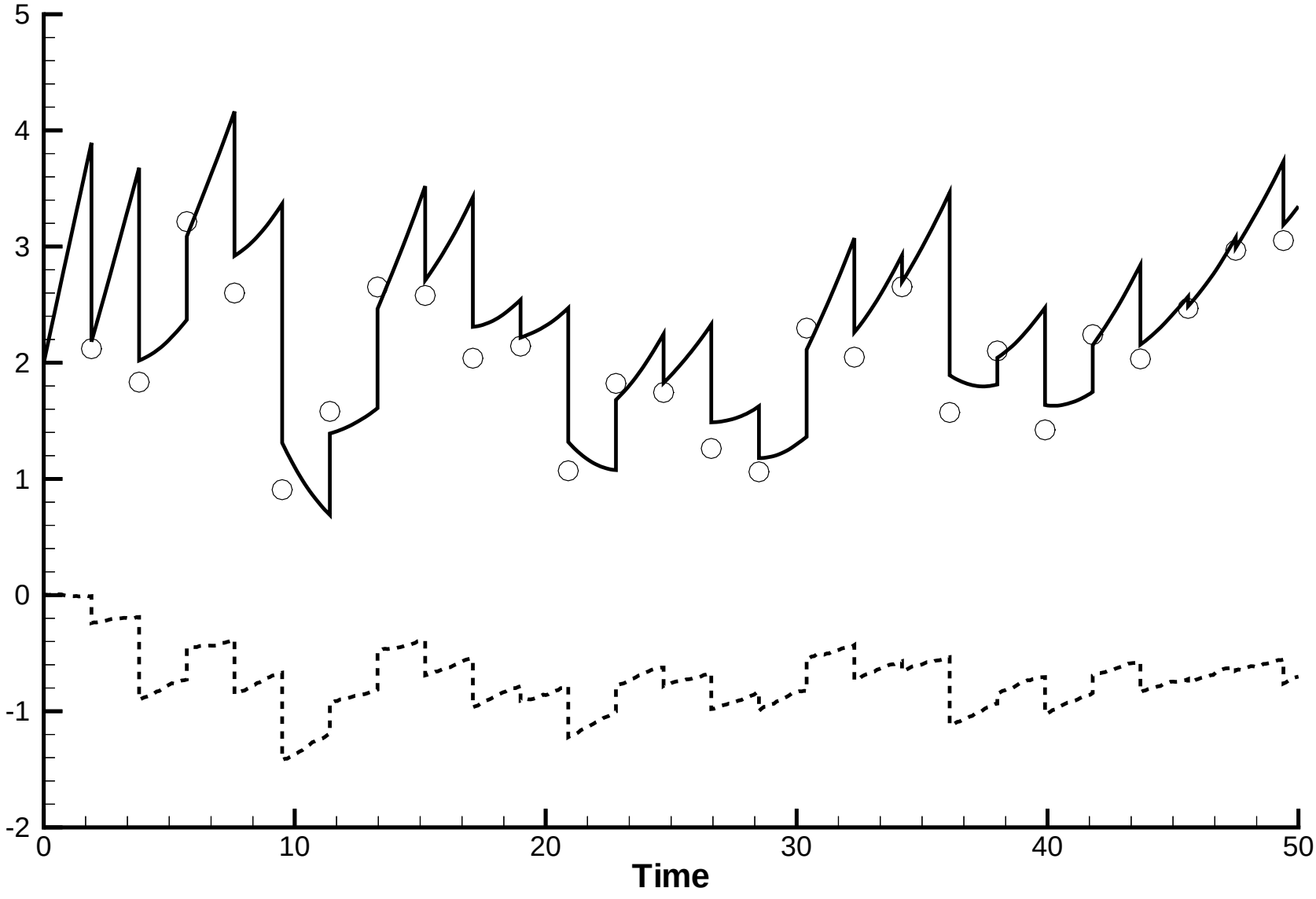
Parameter and bias estimation

- Introduces poorly known parameter β_k in model

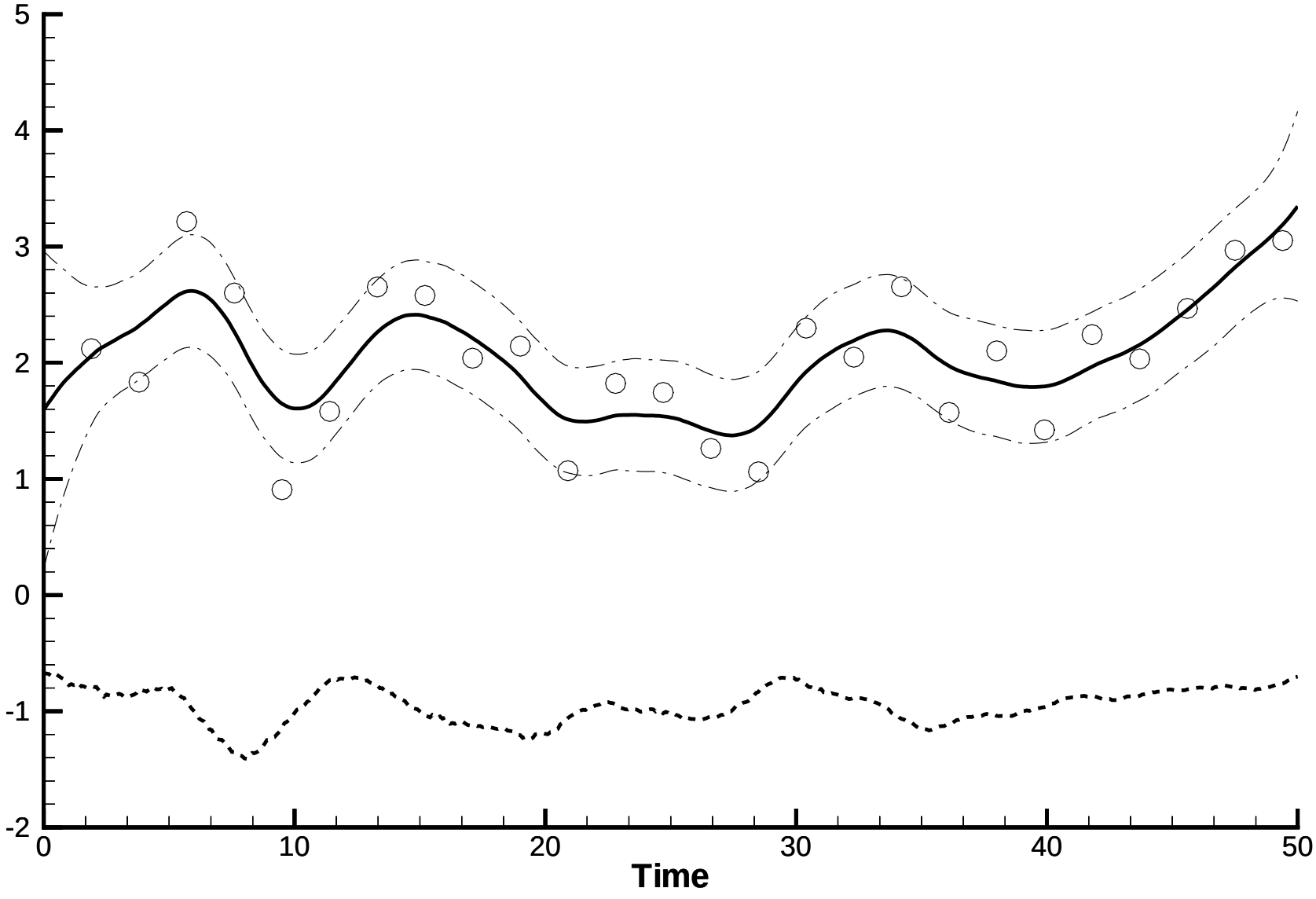
$$\begin{pmatrix} q_k \\ \beta_k \\ \psi_k \end{pmatrix} = \begin{pmatrix} \alpha q_{k-1} \\ \beta_{k-1} \\ \psi_{k-1} + (1+\beta_k)\Delta t + \sqrt{\Delta t}\sigma\rho q_k \end{pmatrix} + \begin{pmatrix} \sqrt{1-\alpha^2} w_{k-1} \\ 0 \\ 0 \end{pmatrix}.$$

- Two cases
  1. $\beta \equiv 0$.
  2. β a poorly known parameter with $\beta_0 = 0$.

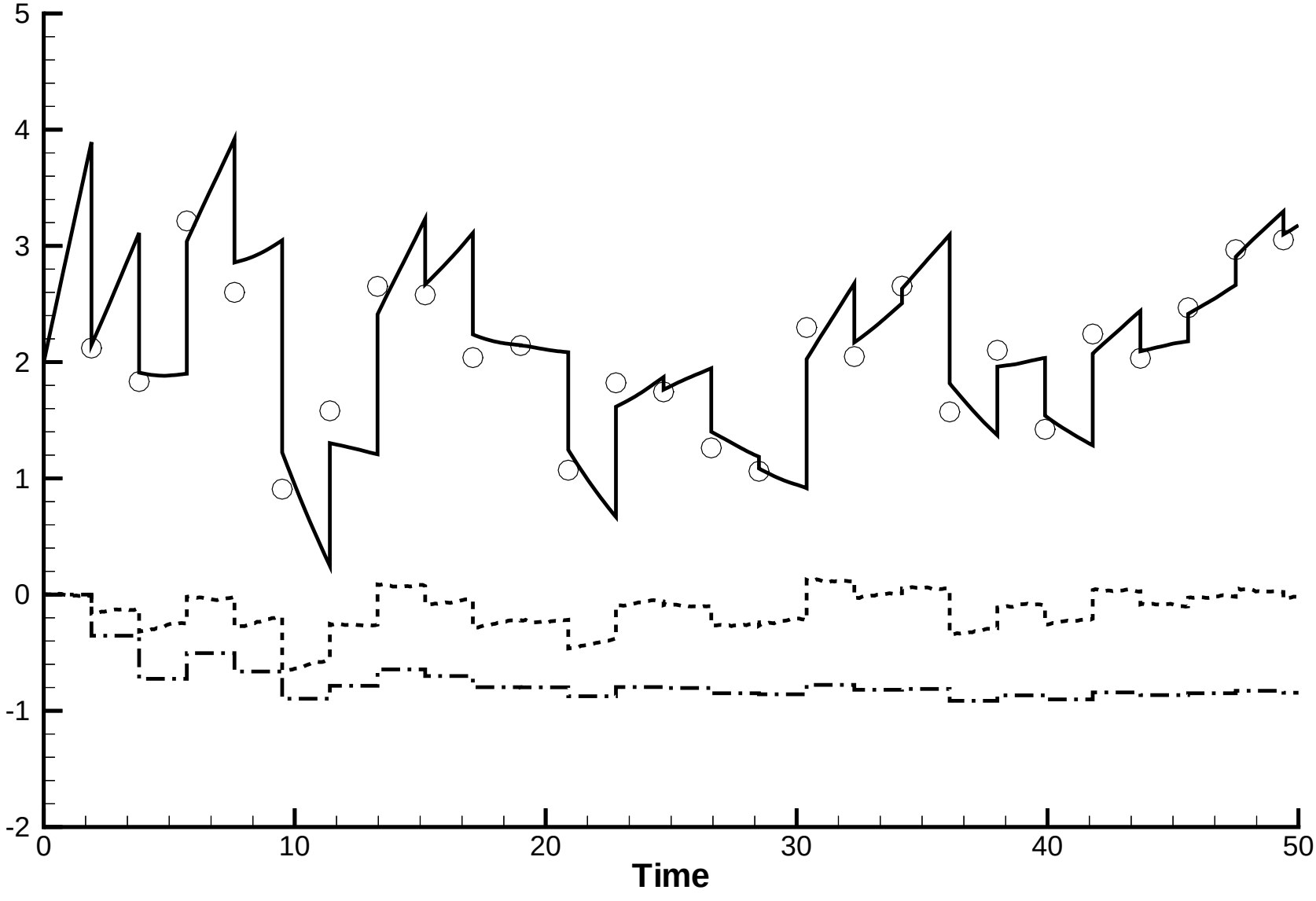
Estimate and model error, EnKF



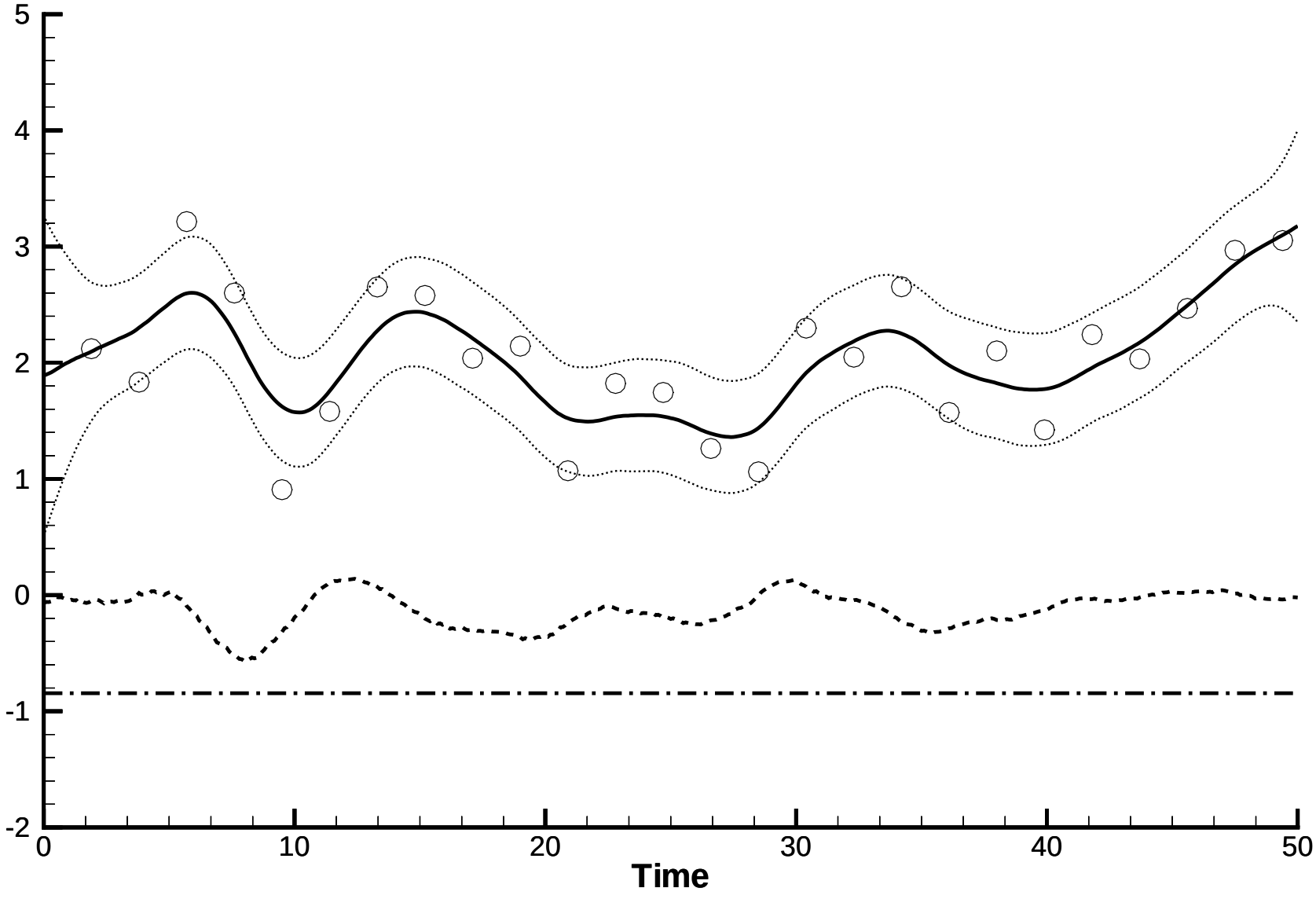
Estimate and model error, EnKS



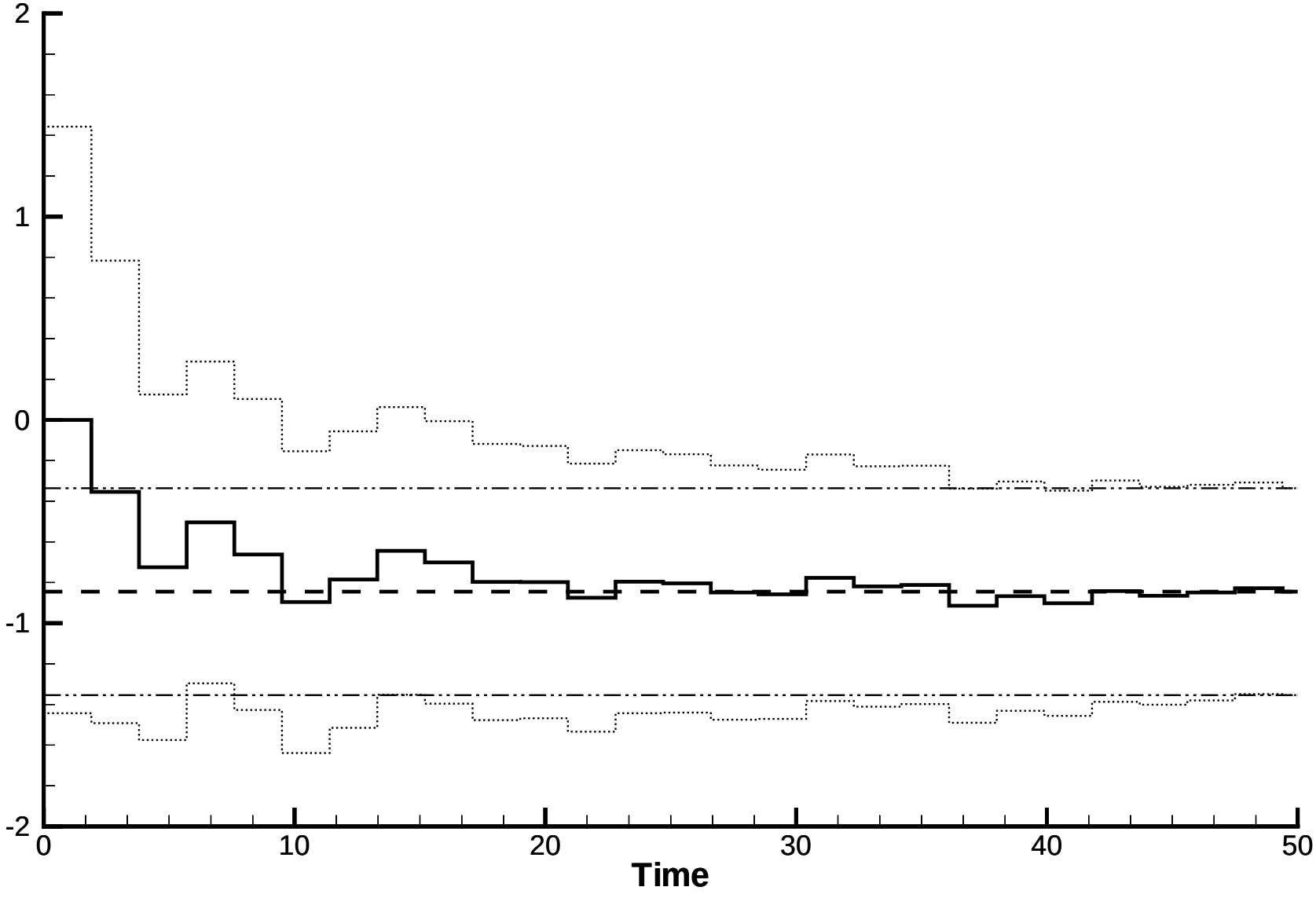
Estimate, model error and β, EnKF



Estimate, model error and β, EnKS



Estimated β and std dev



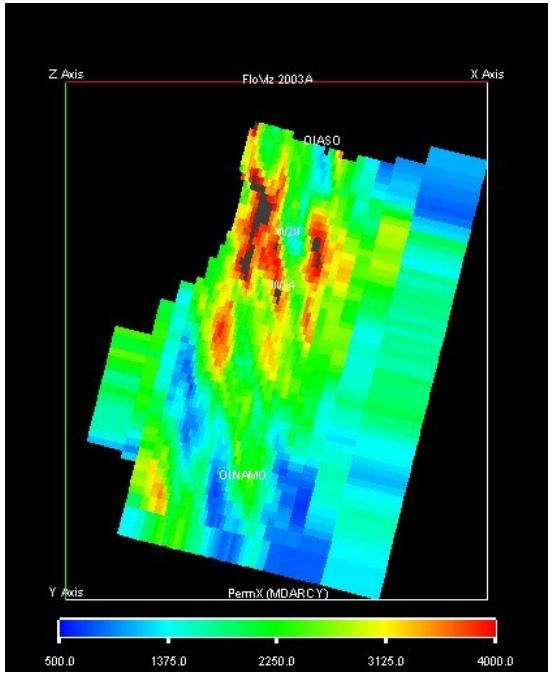
Summary

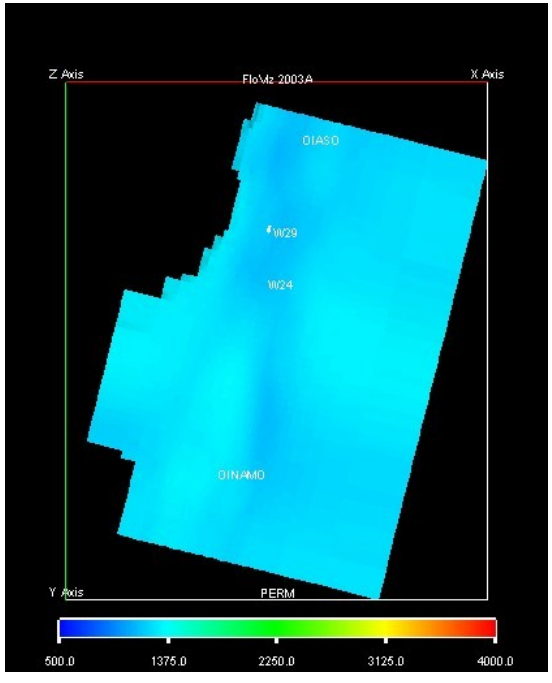
- The EnKF and EnKS can handle time correlated model errors.
- The EnKF and EnKS can be used for parameter estimation by augmenting the model state with the unknown parameters.

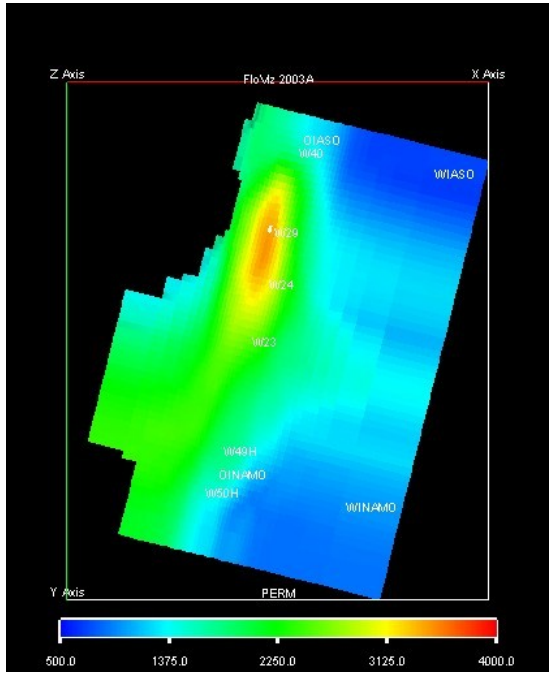
Oil reservoar application

- The EnKF has been used to estimate permeability in a reservoar simulation model using production well data.
- The greatest uncertainty is the reservoar permeability and porosity.
- The well data consists of pressures and oil, gas and water production rates.

Estimated bias and std dev







Local analysis

- May be useful when m is large or $n \gg N$.
- Analysis is computed grid point by grid point using only nearby measurements.
- Inverts many small matrices instead of one large.
- Different $\boldsymbol{X}$ for each grid point, thus, allows us to reach a larger class of solutions.
- Suboptimal solution which introduces nondynamical modes.

Nonlinear measurements

- Measurement equation

$$\boldsymbol{d} = \boldsymbol{h}(\boldsymbol{\psi}) + \boldsymbol{\epsilon}.$$

- Define ensemble of model prediction of the measurements

$$\widehat{\boldsymbol{A}} = (\boldsymbol{h}(\boldsymbol{\psi}_1), \ldots, \boldsymbol{h}(\boldsymbol{\psi}_N)), \in \Re^{\hat{m} \times N}.$$

- The analysis then becomes

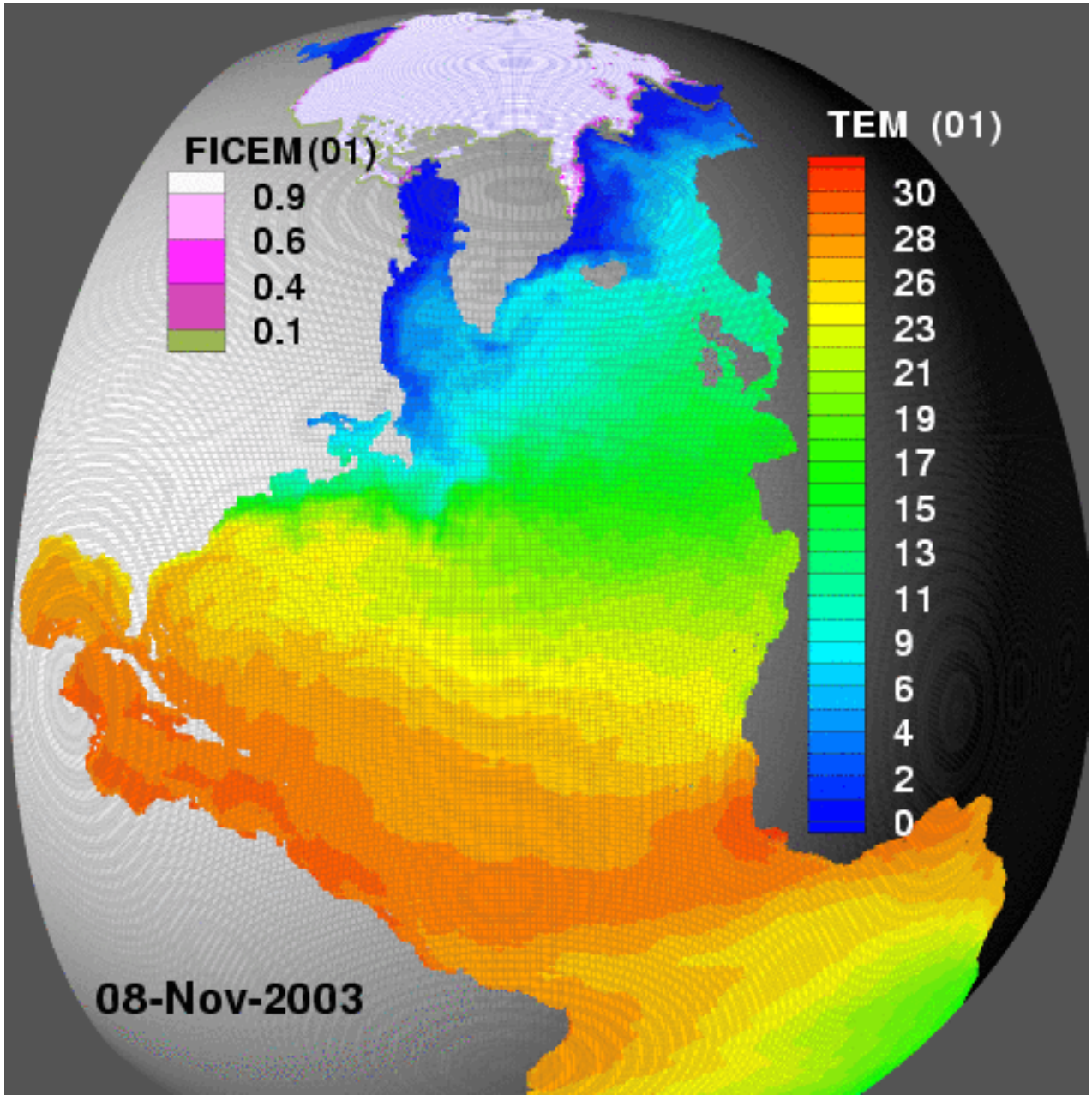
$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{A}' \widehat{\boldsymbol{A}}'^{\mathrm{T}} \left(\widehat{\boldsymbol{A}}' \widehat{\boldsymbol{A}}'^{\mathrm{T}} + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}} \right)^{-1} (\boldsymbol{D} - \widehat{\boldsymbol{A}}),$$

- Analysis based on covariances between $\boldsymbol{h}(\boldsymbol{\psi})$ and $\boldsymbol{\psi}$.

TOPAZ

- Operational ocean prediction system for the Atlantic and Arctic oceans.
- topaz.nersc.no
- Based on HYCOM
- State space is 26 500 000 unknowns.
- Ensemble size is 100.
- Assimilates SSH, SST, Ice concentration, and ARGO T&S profiles.
- Total of more than 10 000 measurements in each assimilation step.
- Uses local analysis as well as nonlinear measurements.

TOPAZ


FICEM (01)
0.9
0.6
0.4
0.1
TEM (01)
30
28
26
23
21
19
17
15
13
11
9
6
4
2
0
08-Nov-2003