The combined parameter and state estimation problem

Geir Evensen

Hydro Research Centre, Bergen, Norway,
Nansen Environmental and Remote Sensing Center, Bergen, Norway

based on
“The combined parameter and state estimation problem”
submitted to Computational Geosciences

ICMS Workshop on Parameter Estimation in Continuous Time Models,
5-9 December 2005, Edinburgh

- Combined parameter and state estimation

  "Find the pdf of the parameters and the model solution conditional on a set of measurements."

- Bayesian formulation.
- Ensemble solutions.

Bayesian formulation

Model equations and measurements:

$$
\begin{aligned}
\frac{\partial \boldsymbol{\psi}}{\partial t} &= \boldsymbol{g}(\boldsymbol{\psi}, \boldsymbol{\alpha}) + \boldsymbol{q}, \\
\boldsymbol{\psi}|_{t_0} &= \boldsymbol{\Psi}_0 + \boldsymbol{a}, \\
\boldsymbol{\psi}|_{\partial \mathcal{D}} &= \boldsymbol{\psi}_b + \boldsymbol{b}, \\
\boldsymbol{\alpha} &= \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}', \\
\mathcal{M}(\boldsymbol{\psi}, \boldsymbol{\alpha}) &= \boldsymbol{d} + \boldsymbol{\epsilon}.
\end{aligned}
$$

Bayes theorem becomes:

$$
\boxed{f(\boldsymbol{\psi}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto f(\boldsymbol{\psi} | \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) f(\boldsymbol{\alpha})\, f(\boldsymbol{d} | \boldsymbol{\psi}, \boldsymbol{\alpha}).}
$$

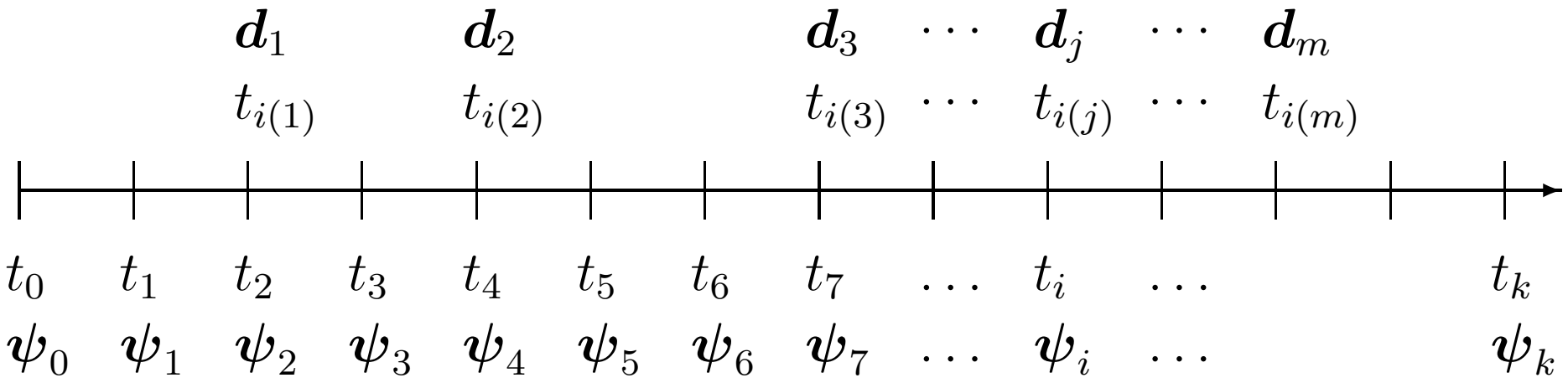
Gaussian priors

$$f(\boldsymbol{\psi}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto \exp\left\{-\frac{1}{2}\mathcal{J}[\boldsymbol{\psi}, \boldsymbol{\alpha}]\right\},$$

$$\begin{aligned}
\mathcal{J}[\boldsymbol{\psi}, \boldsymbol{\alpha}] = & \left(\frac{\partial \boldsymbol{\psi}}{\partial t} - \boldsymbol{g}(\boldsymbol{\psi}, \boldsymbol{\alpha})\right)^{\mathrm{T}} \bullet \boldsymbol{W}_{qq} \bullet \left(\frac{\partial \boldsymbol{\psi}}{\partial t} - \boldsymbol{g}(\boldsymbol{\psi}, \boldsymbol{\alpha})\right) \\
& + (\boldsymbol{\psi}_0 - \boldsymbol{\Psi}_0)^{\mathrm{T}} \circ \boldsymbol{W}_{aa} \circ (\boldsymbol{\psi}_0 - \boldsymbol{\Psi}_0) \\
& + (\boldsymbol{\psi} - \boldsymbol{\psi}_b)^{\mathrm{T}} \star \boldsymbol{W}_{bb} \star (\boldsymbol{\psi} - \boldsymbol{\psi}_b) \\
& + (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^{\mathrm{T}} \circ \boldsymbol{W}_{\alpha\alpha} \circ (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0) \\
& + (\boldsymbol{d} - \mathcal{M}[\boldsymbol{\psi}])^{\mathrm{T}} \boldsymbol{W}_{\epsilon\epsilon} (\boldsymbol{d} - \mathcal{M}[\boldsymbol{\psi}]).
\end{aligned}$$

- MLH solution, hard to solve and to compute error estimates.

Discretisation in time



Assume

- Model is first order Markov process.

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) \propto$$

$$f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_b) f(\boldsymbol{\psi}_0) \prod_{i=1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}, \boldsymbol{\psi}_b).$$

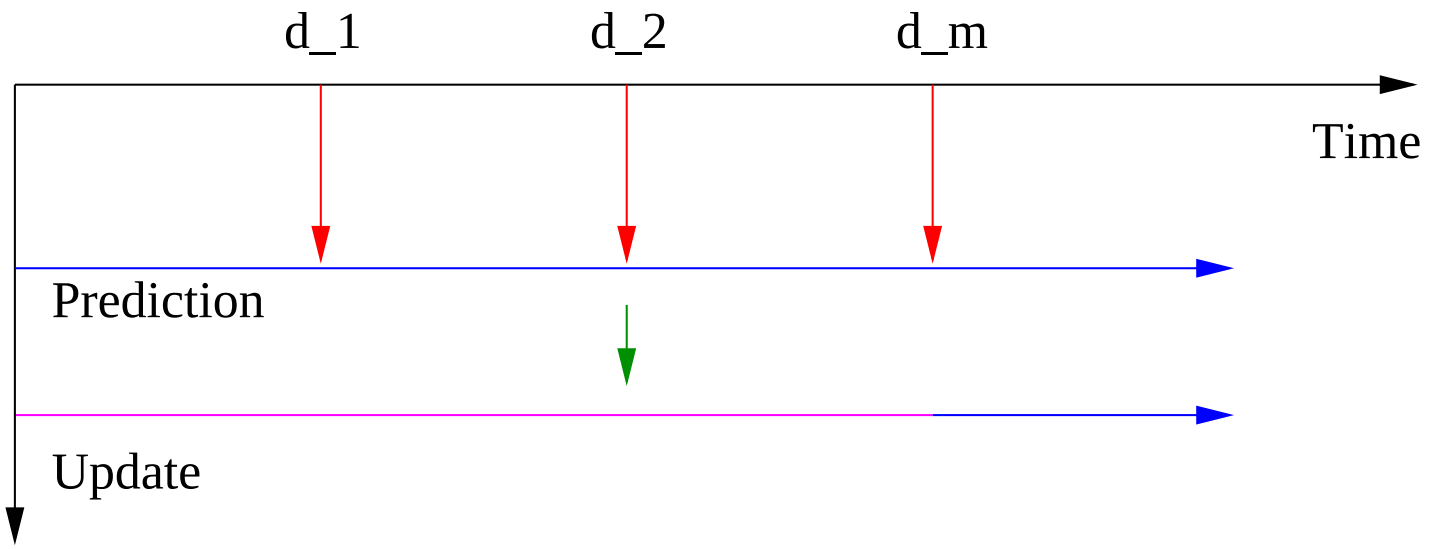
- Measurement errors are independent in time

$$f(\boldsymbol{d} | \boldsymbol{\psi}, \boldsymbol{\alpha}) = \prod_{j=1}^{m} f(\boldsymbol{d}_j | \boldsymbol{\psi}_{i(j)}, \boldsymbol{\alpha}).$$

Bayes for discrete state

- Bayes theorem then becomes

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) \prod_{j=1}^{m} f(\boldsymbol{d}_j | \boldsymbol{\psi}_{i(j)}, \boldsymbol{\alpha}),$$



Rewrite as:

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b)$$

$$\prod_{i=1}^{i(1)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_1 | \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha})$$

$$\prod_{i=i(1)+1}^{i(2)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_2 | \boldsymbol{\psi}_{i(2)}, \boldsymbol{\alpha}) \cdots$$

$$\prod_{i=i(m-1)+1}^{i(m)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_m | \boldsymbol{\psi}_{i(m)}, \boldsymbol{\alpha})$$

$$\prod_{i=i(m)+1}^{k} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha})$$

Sequential processing of measurements

- First update

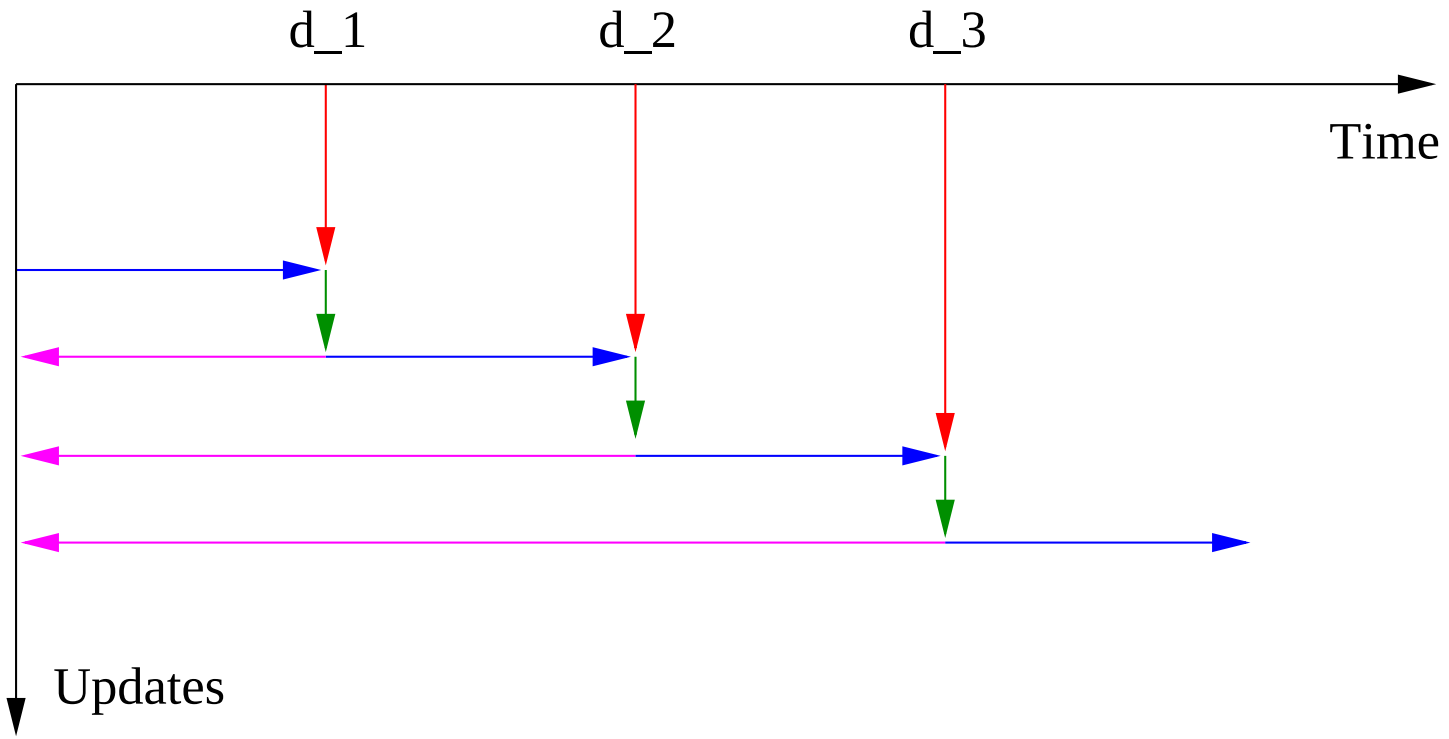
$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1) \propto f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{i(1)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_1 | \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha})$$

- Second update

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(2)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1, \boldsymbol{d}_2) \propto f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1) \prod_{i=i(1)+1}^{i(2)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_2 | \boldsymbol{\psi}_{i(2)}, \boldsymbol{\alpha})$$

Summary

- Independent measurements processed sequentially in time.
- Sequence of inverse problems.
- Solution of one sub-problem is prior for next.
- Hard to solve using traditional variational methods.
- Well suited for ensemble methods.



Ensemble methods

ES: Ensemble Smoother

EnKS: Ensemble Kalman Smoother

EnKF: Ensemble Kalman Filter

- Ensemble representation for pdfs.
- Ensemble prediction for time evolution of pdfs.
- Linear ensemble analysis scheme:
  - "Variance minimizing".
  - Assumes Gaussian pdf for model prediction.
  - No resampling!

Variance minimizing analysis

Given

- First guess, $\boldsymbol{\psi}^{\mathrm{f}}$, and prediction error covariance, $\boldsymbol{C}^{\mathrm{f}}_{\psi\psi}$.
- Measurements, $\boldsymbol{d}$, and error covariance, $\boldsymbol{C}_{\epsilon\epsilon}$.

The variance minimizing analysis is minimum of

$$\mathcal{J}[\boldsymbol{\psi}^{\mathrm{a}}] = \left(\boldsymbol{\psi}^{\mathrm{a}} - \boldsymbol{\psi}^{\mathrm{f}}\right)^{\mathrm{T}} \bullet \boldsymbol{C}^{-1}_{\psi\psi} \bullet \left(\boldsymbol{\psi}^{\mathrm{a}} - \boldsymbol{\psi}^{\mathrm{f}}\right) + \left(\boldsymbol{d} - \mathcal{M}[\boldsymbol{\psi}^{\mathrm{a}}]\right)^{\mathrm{T}} \boldsymbol{C}^{-1}_{\epsilon\epsilon} \left(\boldsymbol{d} - \mathcal{M}[\boldsymbol{\psi}^{\mathrm{a}}]\right).$$

Ensemble analysis equations

- Define randomized measurements $\boldsymbol{D} = \boldsymbol{d} + \boldsymbol{E}$.
- Variance minimizing ensemble update becomes

$$
\begin{aligned}
\boldsymbol{A}^{\mathrm{a}} &= \boldsymbol{A} + \mathcal{M}^{\mathrm{T}}\left[\boldsymbol{C}^{\mathrm{e}}_{\psi\psi}\right] \\
&\quad \left(\mathcal{M}^{\mathrm{T}}\left[\mathcal{M}\left[\boldsymbol{C}^{\mathrm{e}}_{\psi\psi}\right]\right] + \boldsymbol{C}_{\epsilon\epsilon}\right)^{-1}\left(\boldsymbol{D} - \mathcal{M}\left[\boldsymbol{A}\right]\right) \\
&= \boldsymbol{A} + \boldsymbol{A}'\mathcal{M}^{\mathrm{T}}\left[\boldsymbol{A}'\right] \\
&\quad \left(\mathcal{M}\left[\boldsymbol{A}'\right]\mathcal{M}^{\mathrm{T}}\left[\boldsymbol{A}'\right] + \boldsymbol{E}\boldsymbol{E}^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{D} - \mathcal{M}\left[\boldsymbol{A}\right]\right) \\
&\vdots \\
&= \boldsymbol{A}\boldsymbol{X}
\end{aligned}
$$

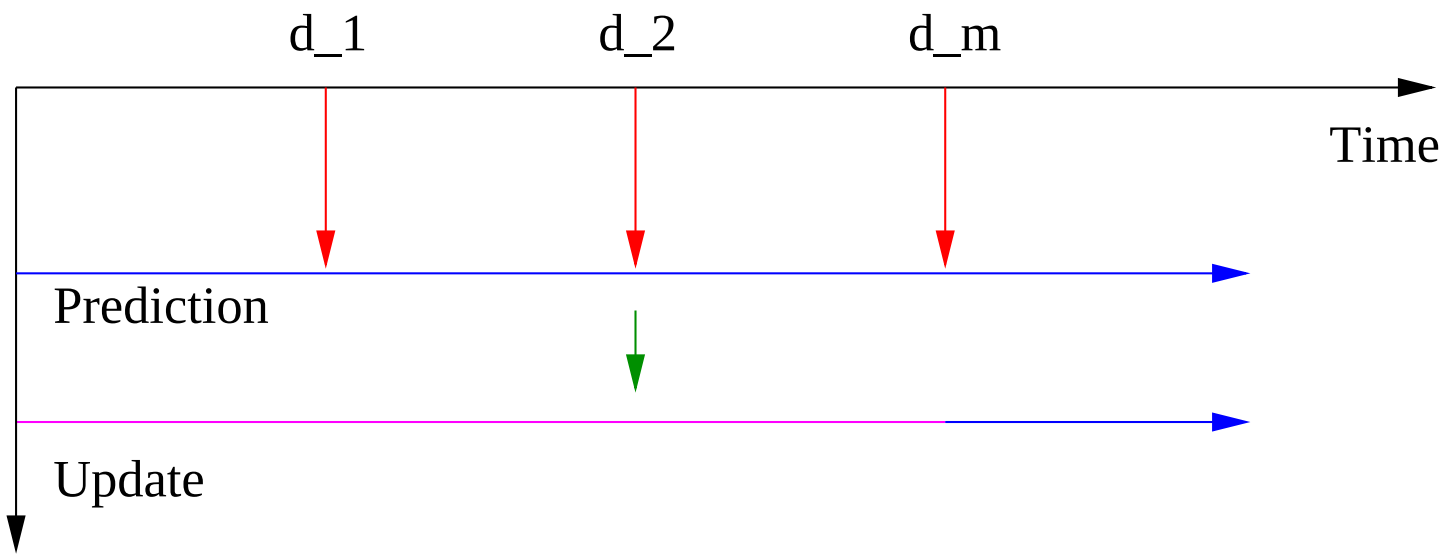
$\Rightarrow$ Updated ensemble with correct mean and covariance!

ES: The Ensemble Smoother

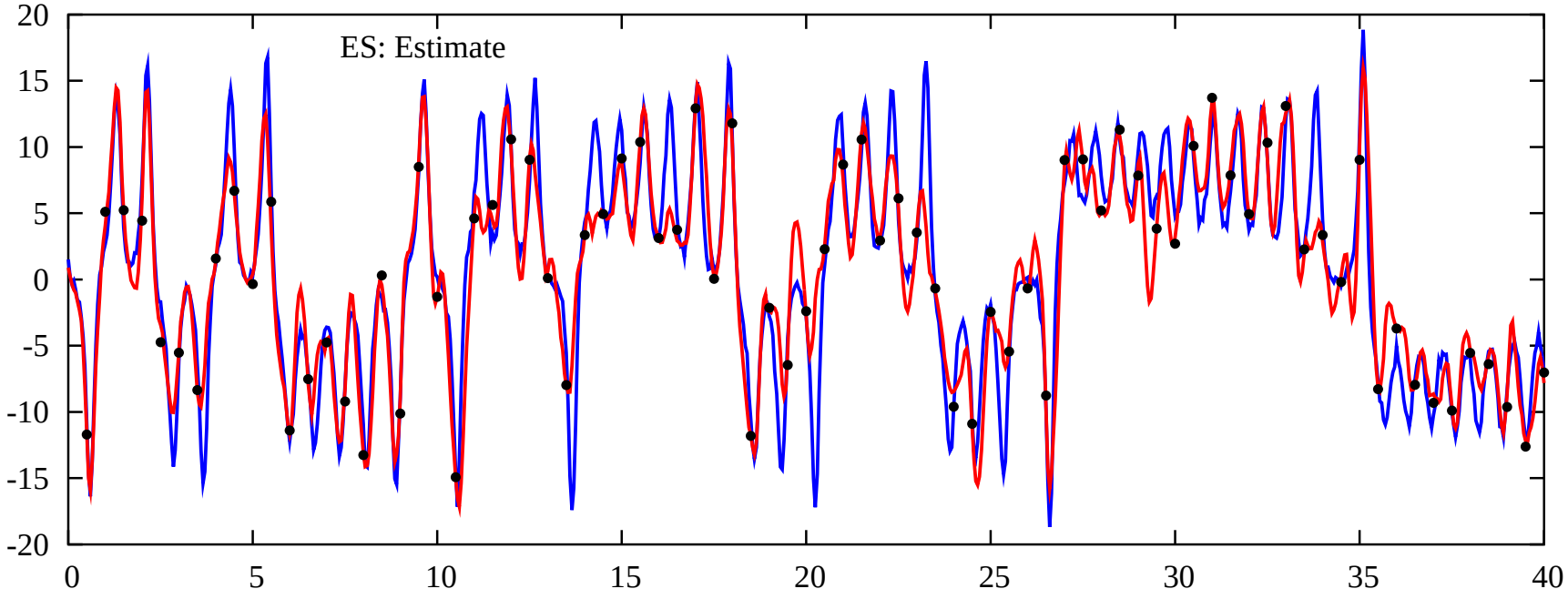
Computes the Bayesian update using linear ensemble equation

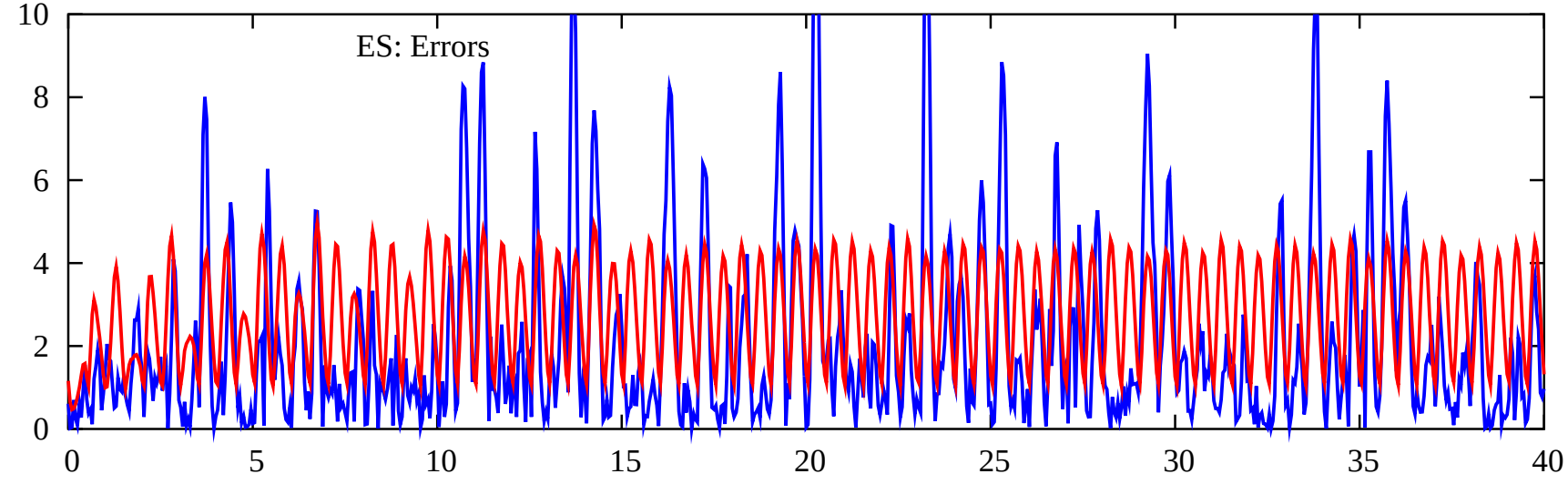
$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}) =$$

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b) \prod_{j=1}^{m} f(\boldsymbol{d}_j | \boldsymbol{\psi}_{i(j)}, \boldsymbol{\alpha}),$$



ES: Example with Lorenz equations





ES: summary

Gauss–Markov interpolation in space and time.

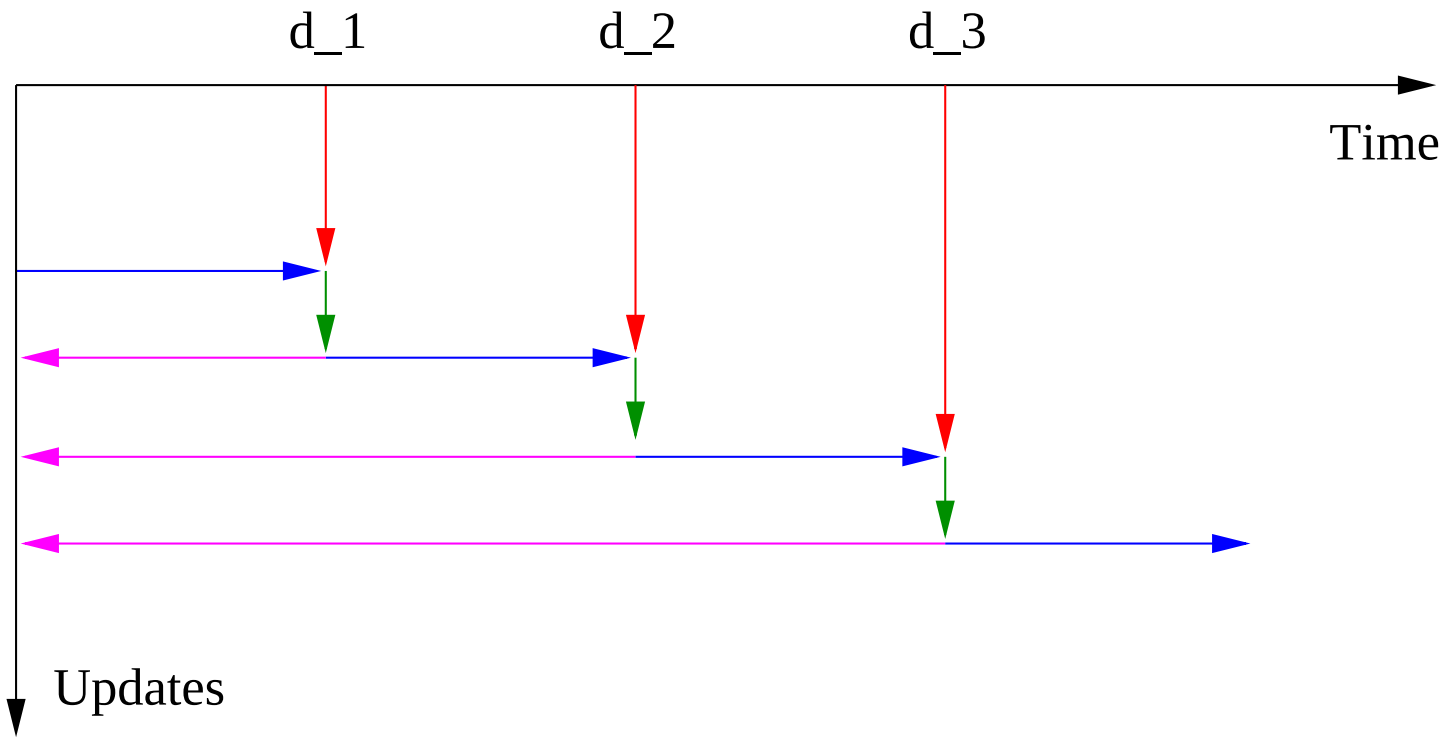
- Creates an ensemble for the model evolution.
- Assumes Gaussian pdf for model evolution.
- Computes variance minimizing ensemble analysis.
- Exact solution for linear problems.

EnKS: The ensemble Kalman smoother

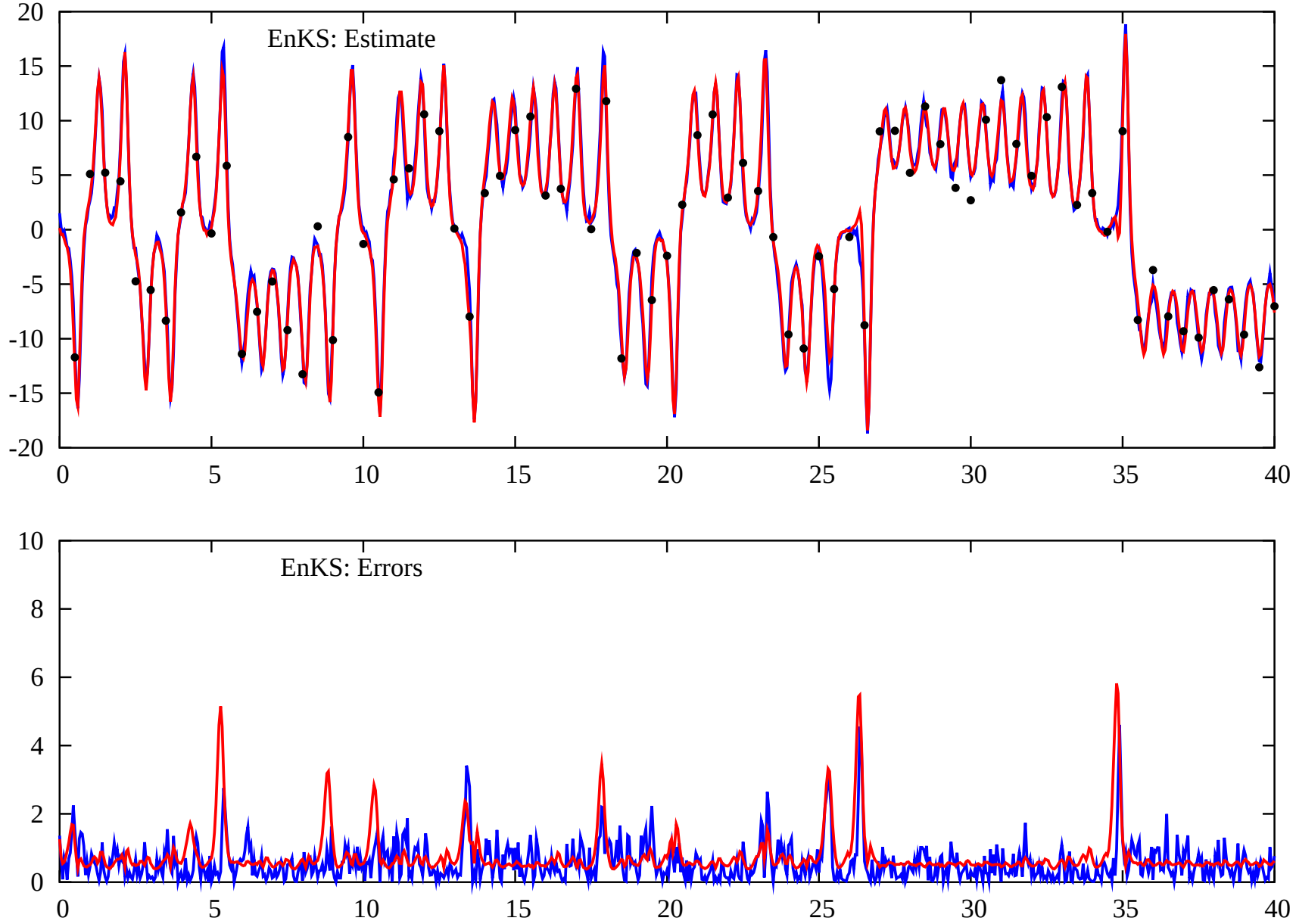
- Bayes with sequential processing of data: First update is

$$f(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}, \boldsymbol{\psi}_0, \boldsymbol{\psi}_b | \boldsymbol{d}_1) \propto$$

$$f(\boldsymbol{\alpha}) f(\boldsymbol{\psi}_0) f(\boldsymbol{\psi}_b) \prod_{i=1}^{i(1)} f(\boldsymbol{\psi}_i | \boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) f(\boldsymbol{d}_1 | \boldsymbol{\psi}_{i(1)}, \boldsymbol{\alpha}),$$



EnKS solution

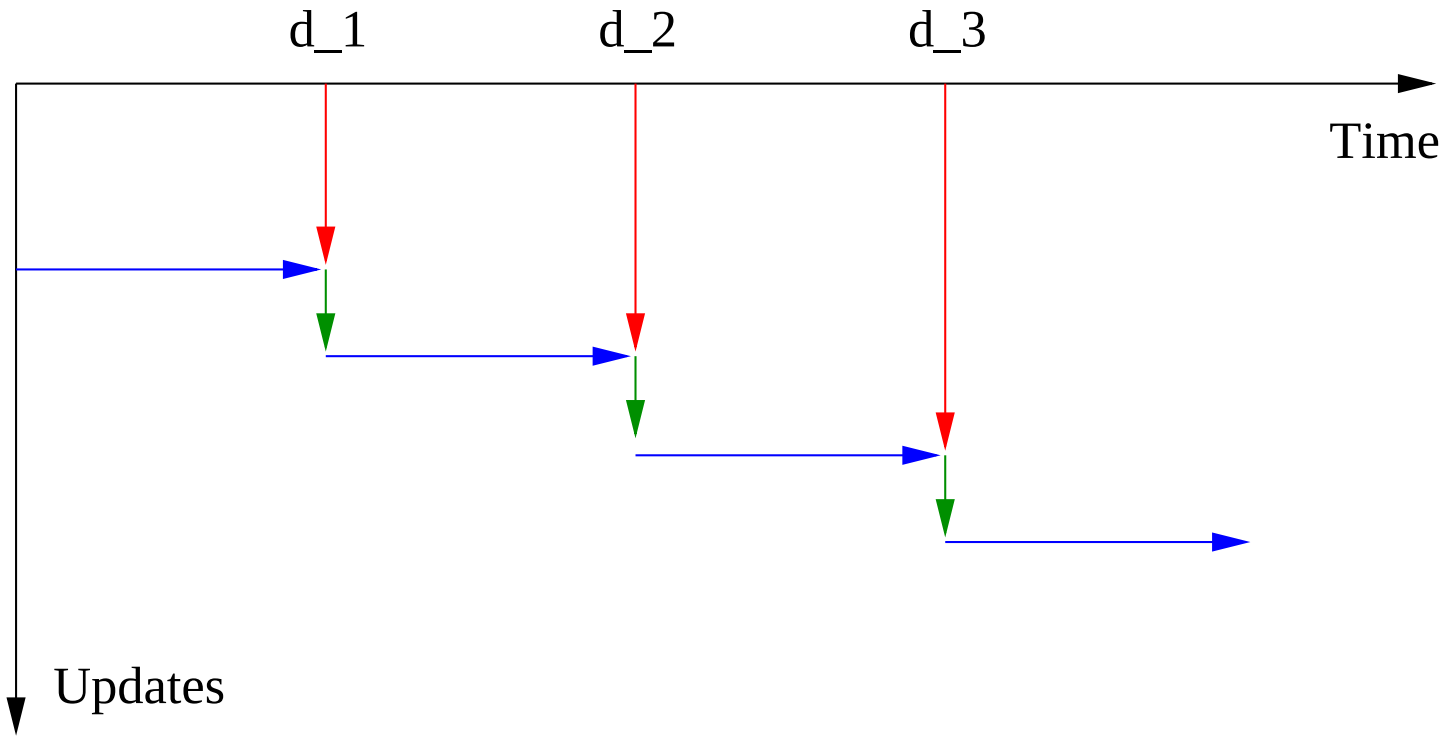


EnKS summary

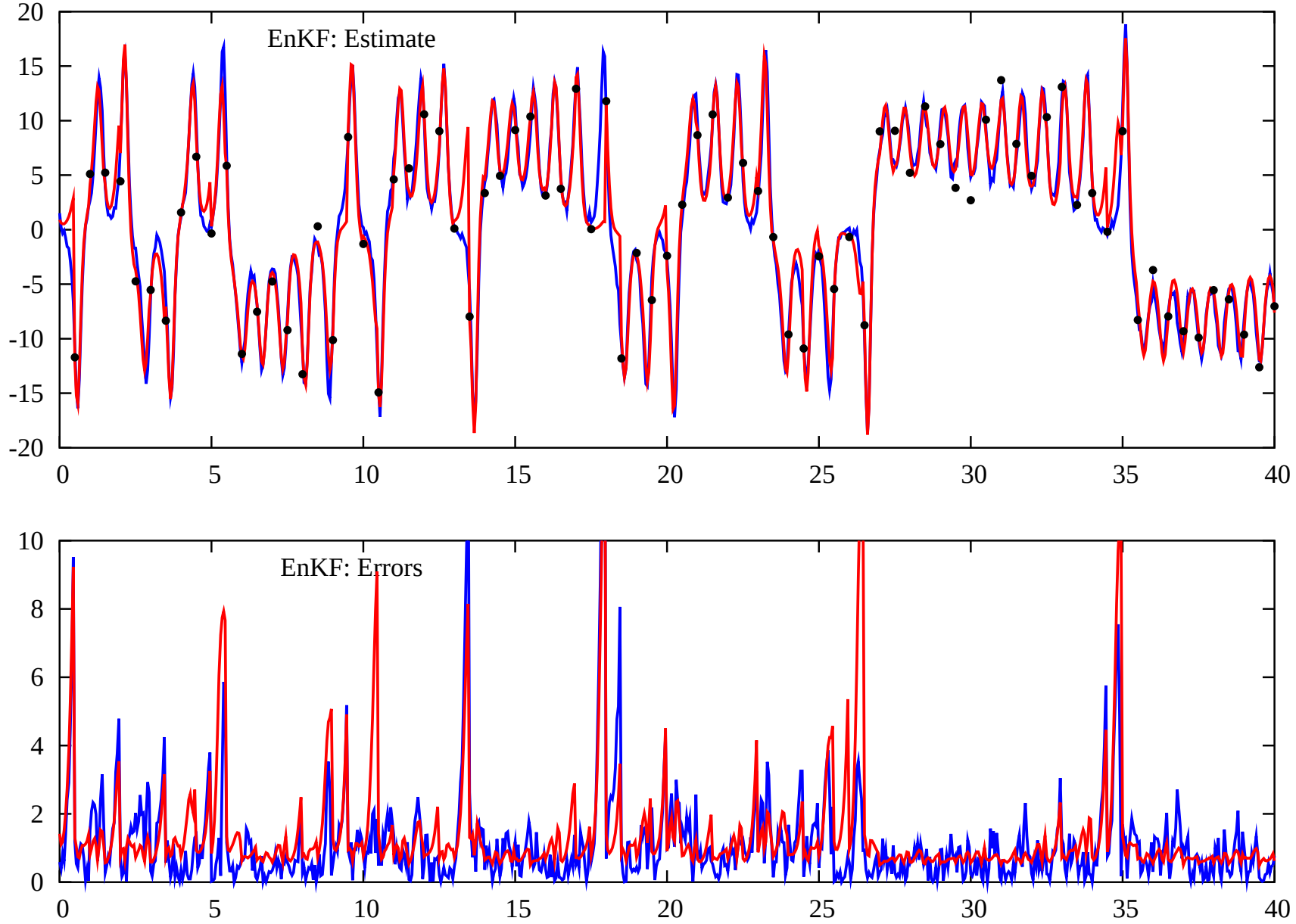
- ES and EnKS give identical results for linear models.
- EnKS is superior to the ES with nonlinear models.
  - Sequential processing of measurements introduces “Gaussianity”.
  - Ensemble is kept close to the true state.

EnKF: Ensemble Kalman Filter

- Special case of EnKS.
- Only update state at data times
- EnKF forms a prior for the EnKS.
- A prediction from EnKF and EnKS will be identical.



EnKF solution



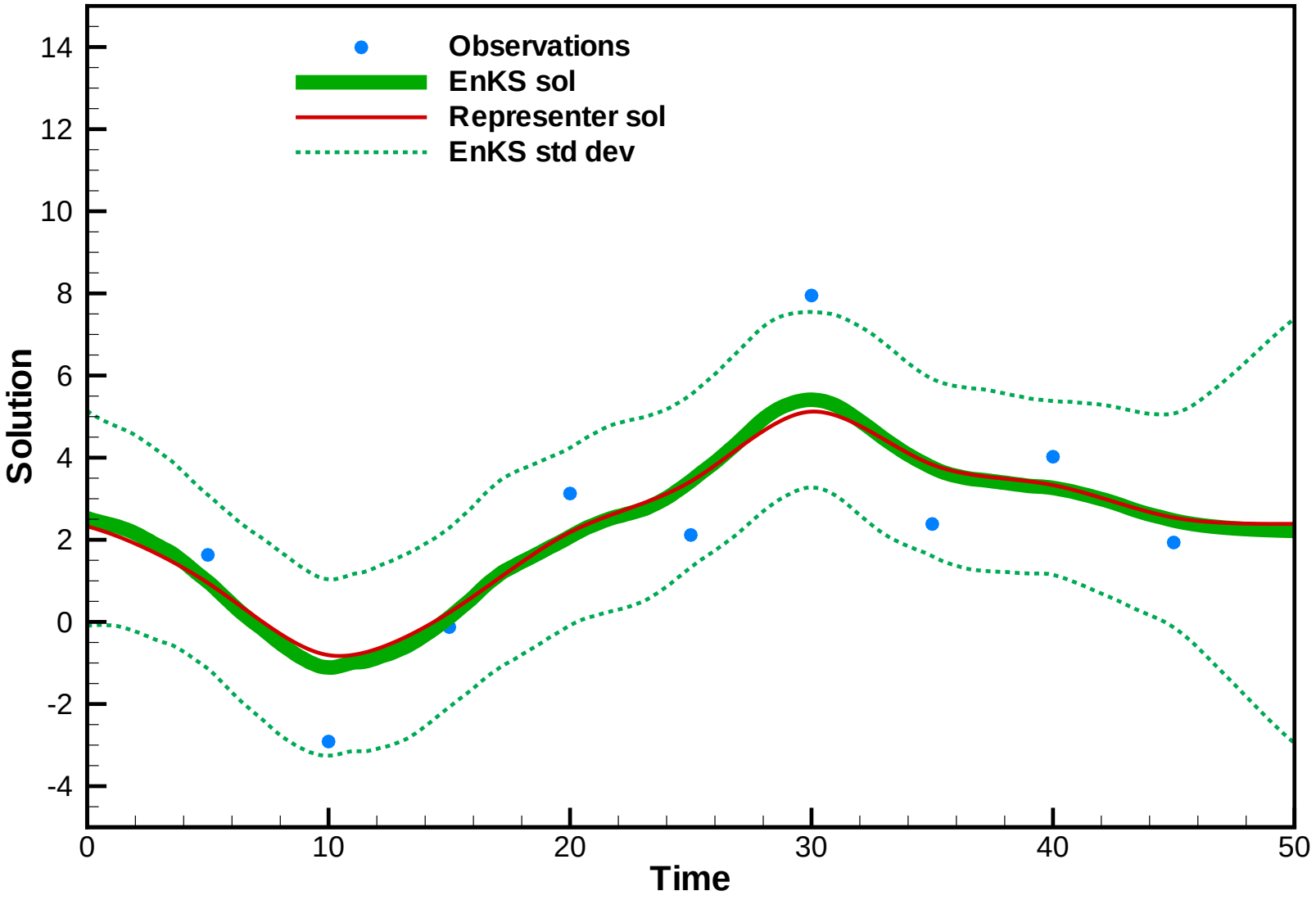
Example

- Scalar model

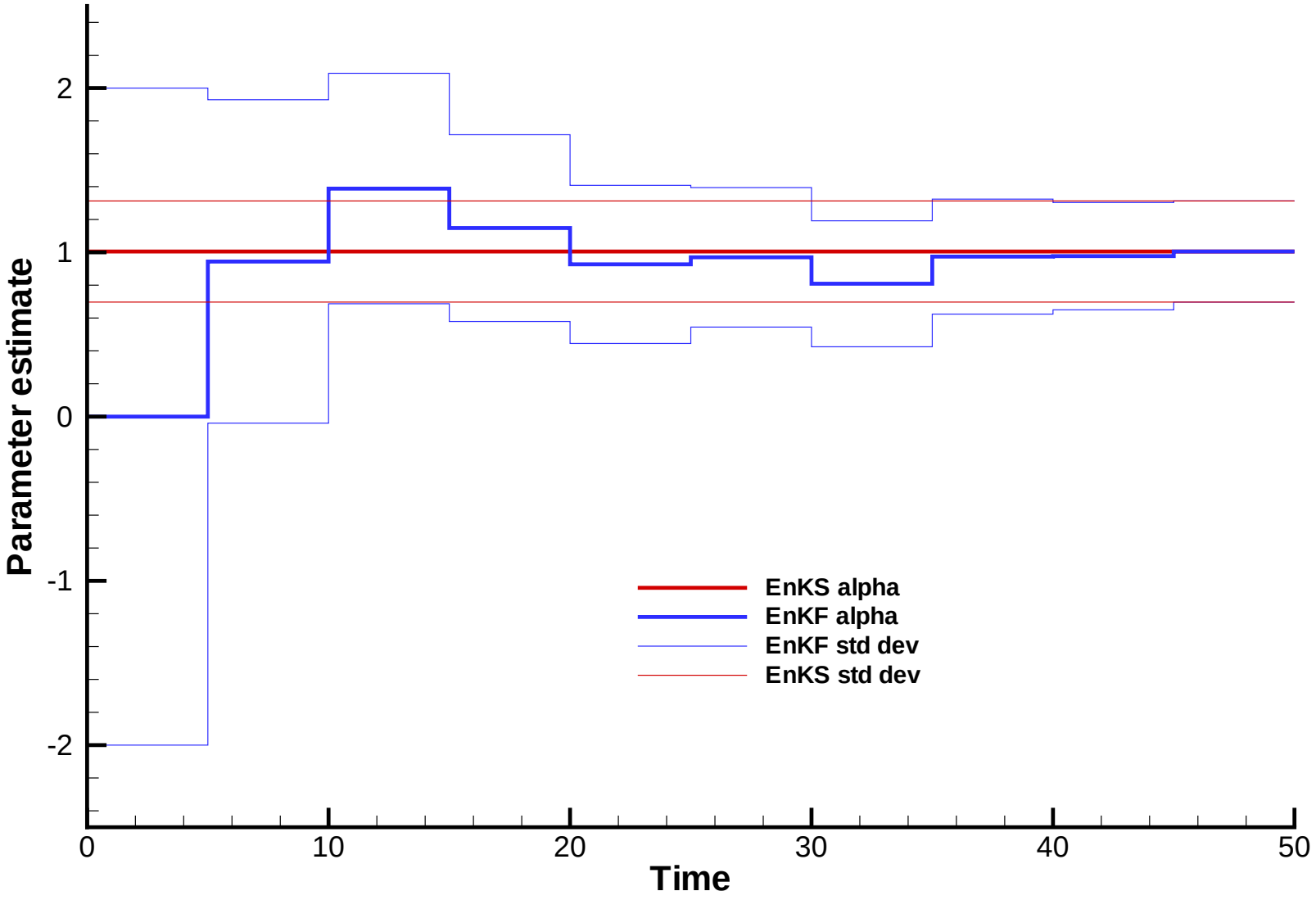
$$\frac{\partial \psi}{\partial t} = 1 - \alpha + q,$$
$$\psi(t=0) = 3 + a,$$
$$\alpha = 0 + \alpha',$$
$$\mathcal{M}(\psi) = \boldsymbol{d} + \boldsymbol{\epsilon}.$$

- True parameter value is $\alpha = 1$.
- Model operator is linear and independent of ψ.
- Solved using EnKF, EnKS and Representer methods.
- Exponential time correlation for model errors.

State and parameter estimation

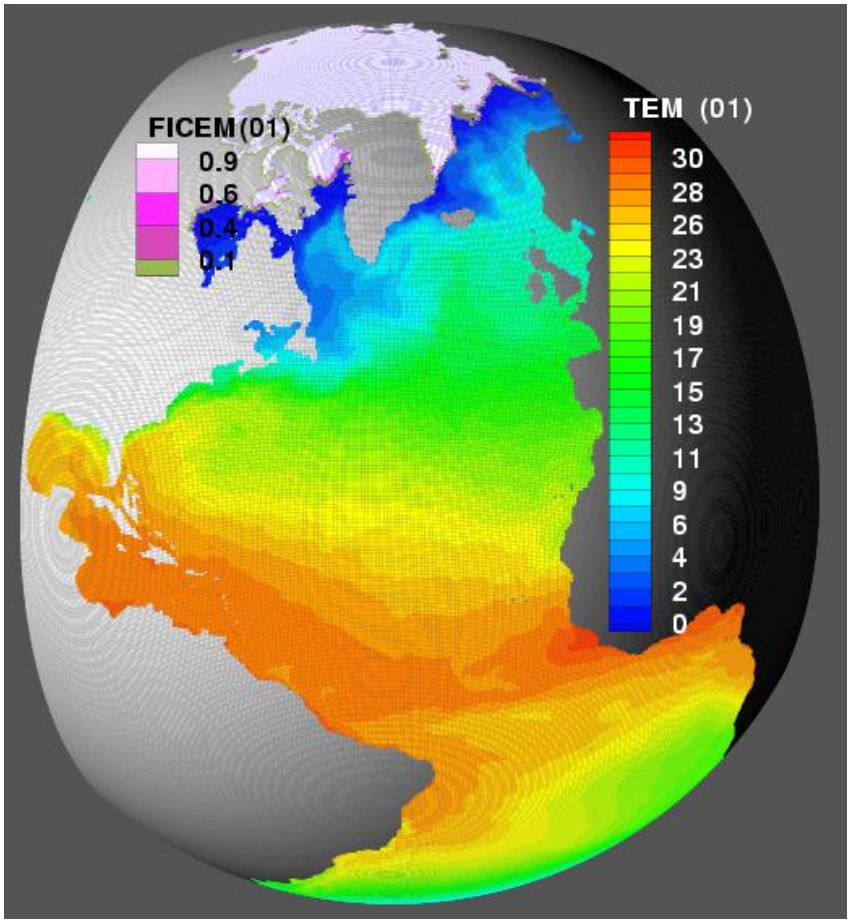


Estimate of parameter



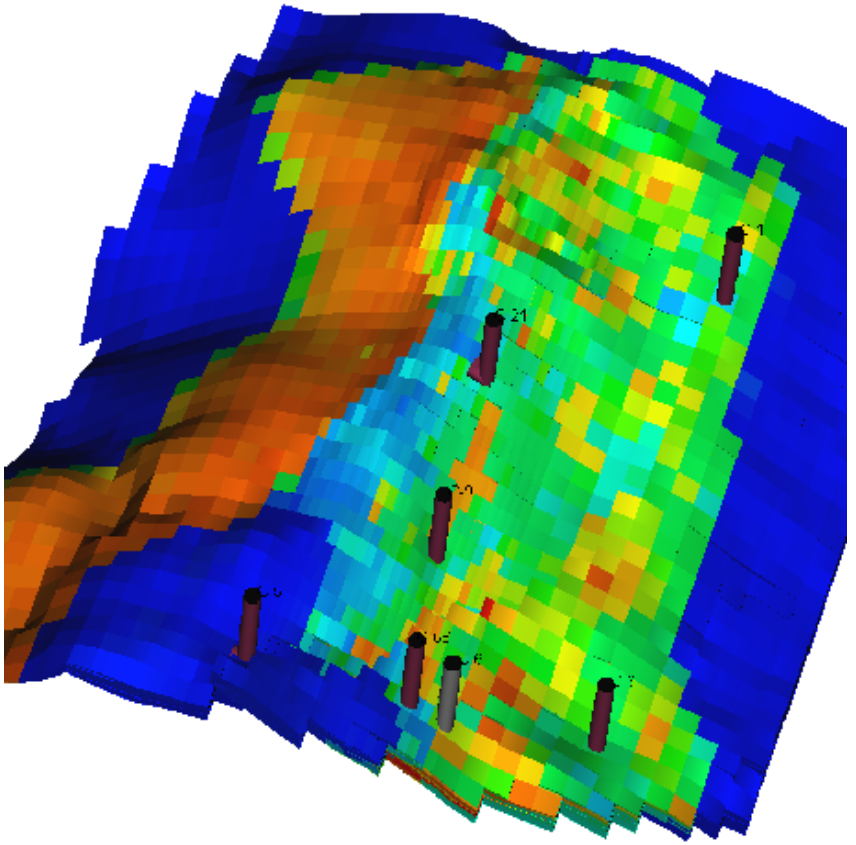
TOPAZ ocean forecasting system

- State vector has 26 400 000 unknowns.
- Around 10 000 measurements are assimilated.
- Assimilation method is EnKF.
- Operated for several years "topaz.nersc.no"



Reservoir Simulation Models

- EnKF for history matching in ECLIPSE.
- Estimates porosity, permeability and other parameters.
- Assimilates production and 4D-seismic data.



Summary

www.nersc.no/∼geir/EnKF

- Joint state and parameter estimation problem!
- Bayesian formulation:
  - Variational methods (multiple minima; MLH).
    - Hard to compute error statistics!
    - Do not allow for sequential processing of measurements!
  - Ensemble methods (Gaussianity assumption; mean).
    - Provide estimate with error statistics.
    - Allow for sequential processing of measuremets.
      - Introduces Gaussianity!
      - Advantage with nonlinear dynamics.