Overview of data assimilation methods

Femke C. Vossepoel, based on the book of Geir Evensen, Femke C. Vossepoel and Peter Jan van Leeuwen

Book available from
`https://github.com/geirev/Data-Assimilation-Fundamentals.git`

Springer Textbooks in Earth Sciences, Geography and Environment

Geir Evensen · Femke C. Vossepoel · Peter Jan van Leeuwen

Data Assimilation Fundamentals

This open-access textbook's significant contribution is the unified derivation of data-assimilation techniques from a common fundamental and optimal starting point, namely Bayes' theorem. Unique for this book is the "top-down" derivation of the assimilation methods. It starts from Bayes theorem and gradually introduces the assumptions and approximations needed to arrive at today's popular data-assimilation methods. This strategy is the opposite of most textbooks and reviews on data assimilation that typically take a bottom-up approach to derive a particular assimilation method. E.g., the derivation of the Kalman Filter from control theory and the derivation of the ensemble Kalman Filter as a low-rank approximation of the standard Kalman Filter. The bottom-up approach derives the assimilation methods from different mathematical principles, making it difficult to compare them. Thus, it is unclear which assumptions are made to derive an assimilation method and sometimes even which problem it aspires to solve. The book's top-down approach allows categorizing data-assimilation methods based on the approximations used. This approach enables the user to choose the most suitable method for a particular problem or application. Have you ever wondered about the difference between the ensemble 4DVar and the "ensemble randomized likelihood" (EnRML) methods? Do you know the differences between the ensemble smoother and the ensemble-Kalman smoother? Would you like to understand how a particle flow is related to a particle filter? In this book, we will provide clear answers to several such questions. The book provides the basis for an advanced course in data assimilation. It focuses on the unified derivation of the methods and illustrates their properties on multiple examples. It is suitable for graduate students, post-docs, scientists, and practitioners working in data assimilation.

ISBN 978-3-030-96708-6

9 783030 967086

▶ springer.com

Evensen · Vossepoel · Leeuwen

Data Assimilation Fundamentals

TEXTBOOK

Geir Evensen · Femke C. Vossepoel
Peter Jan van Leeuwen

Data Assimilation Fundamentals

A Unified Formulation of the State and Parameter Estimation Problem

OPEN ACCESS

Springer

Data assimilation minimises a cost function to fit a model to observations

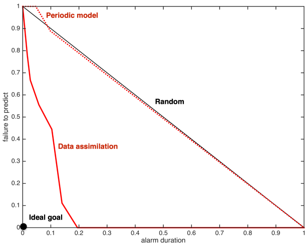
Current data assimilation methods are computationally too expensive

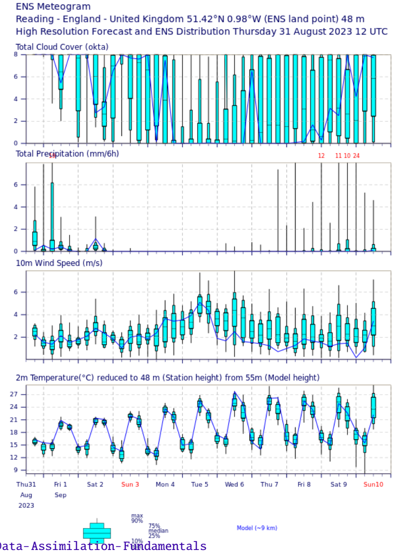
Data assimilation provides information about our models and their uncertainties

Who loves linear algebra?

Why data assimilation in earth systems?

Why data assimilation in earth systems?





- improving re-analysis
- improving forecasts
- scenario modelling ('what if?')

Application areas in Earth Systems (biased view)

Ocean reanalysis

Groundwater management

Climate (re)analysis

Flood forecasting

Weather forecasting

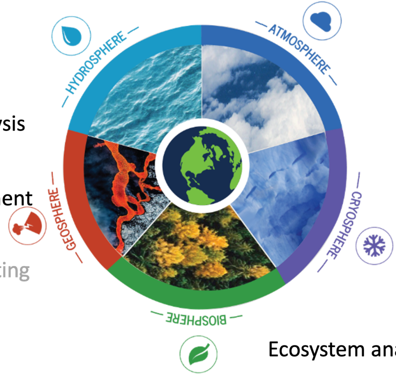
Climate forecasting

Geotechnical risk analysis

Mass balance analysis

Reservoir management

Earthquake forecasting



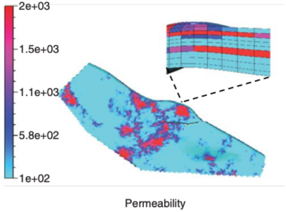
Ecosystem analysis

NASA

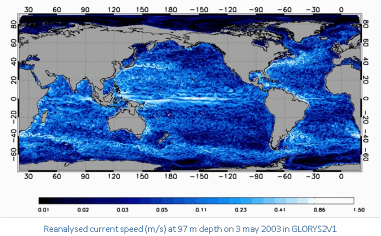
Different objectives of data assimilation

What to estimate

- state estimation (initial conditions, time evolution)
- parameter estimation



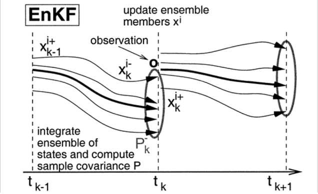
(Glegola et al., 2013)



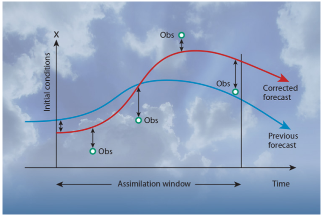
(MERCATOR)

Common classification (I)

- Ensemble methods
  - Ensemble Kalman Filter (EnKF)
  - Ensemble Smoothers: ES, ESMDA, IES
- Variational methods
  - 4D-Var, En4DVar
  - Randomized Maximum Likelihood (RML), EnRML
- Nonlinear methods
  - Particle Filter
  - Particle Flow Filter



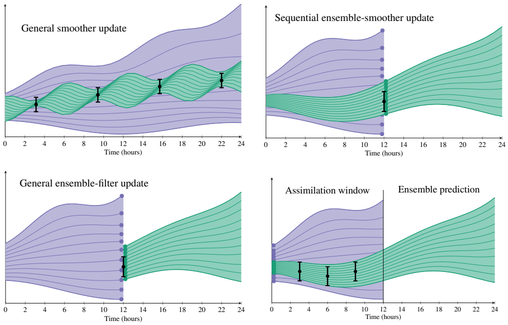
EnKF (Reichle, 2002)



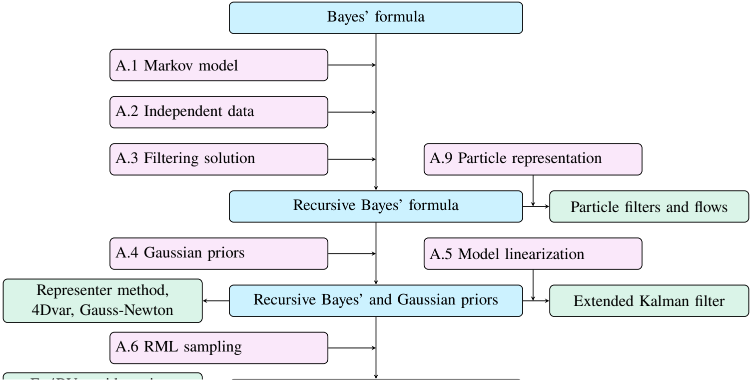
4DVar (ECMWF, 2017)

Common classification (II)

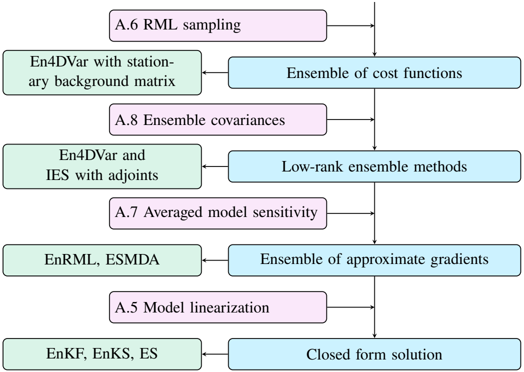
- Smoothers
  - (Ensemble) 4D-Var
  - (Ensemble) Randomized Maximum Likelihood (RML, EnRML)
  - Ensemble Smoothers: ES, ESMDA, IES
- Filters
  - Kalman Filter, Extended Kalman Filter
  - EnKF
  - Particle Filter, Particle Flow Filter



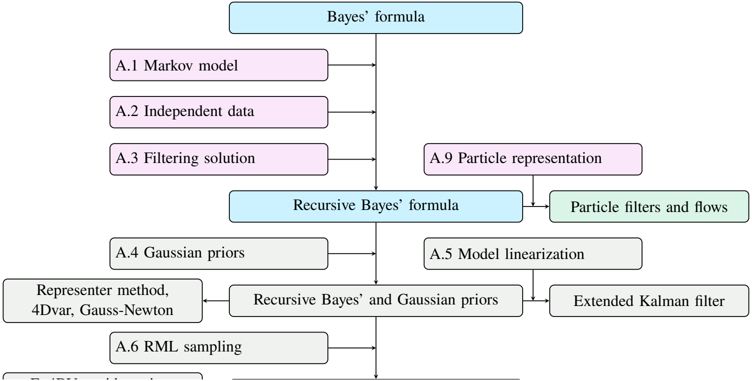
Overview of approximations and methods



Overview of approximations and methods



Overview of approximations and methods

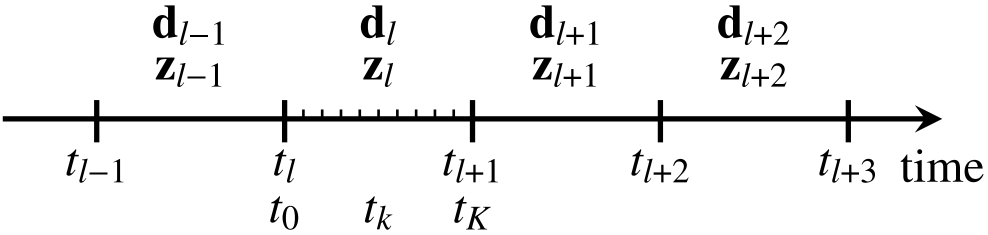


We start from Bayes' theorem

$$\boxed{f(\mathcal{Z}|\mathcal{D}) = \frac{f(\mathcal{D}|\mathcal{Z})\,f(\mathcal{Z})}{f(\mathcal{D})}}$$

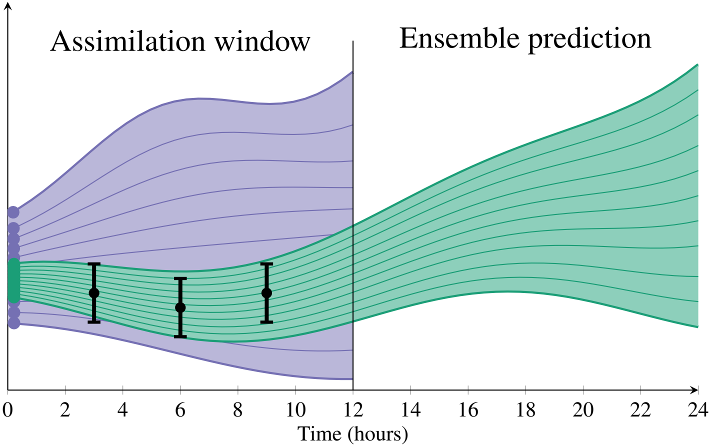
- $\mathcal{Z} = (\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_L)$ is the vector of state variables on all the assimilation windows.
- $\mathcal{D} = (\mathbf{d}_1, \ldots, \mathbf{d}_L)$ is the vector containing all the measurements.

Split time into data-assimilation windows



- We consider the DA problem for one single window.
- Errors propagate from one window to the next.

Weather forecasting approach



Model is Markov process

Approximation 1 (Model is 1st-order Markov process)

We assume the dynamical model is a 1st-order Markov process.

$$f(\mathbf{z}_l|\mathbf{z}_{l-1}, \mathbf{z}_{l-2}, \ldots, \mathbf{z}_0) = f(\mathbf{z}_l|\mathbf{z}_{l-1})$$

Independent measurements

Approximation 2 (Independent measurements)
We assume that measurements are independent between different assimilation windows.

Independent measurements have uncorrelated errors

$$f(\mathcal{D}|\mathcal{Z}) = \prod_{l=1}^{L} f(\mathbf{d}_l|\mathbf{z}_l) \tag{23}$$

Bayes becomes

$$f(\mathcal{Z}|\mathcal{D}) \propto \prod_{l=1}^{L} f(\mathbf{d}_l|\mathbf{z}_l) \prod_{l=1}^{L} f(\mathbf{z}_l|\mathbf{z}_{l-1}) f(\mathbf{z}_0) \tag{24}$$

Recursive form of Bayes

$$f(\mathbf{z}_1, \mathbf{z}_0 | \mathbf{d}_1) = \frac{f(\mathbf{d}_1|\mathbf{z}_1) f(\mathbf{z}_1|\mathbf{z}_0) f(\mathbf{z}_0)}{f(\mathbf{d}_1)}, \tag{25}$$

$$f(\mathbf{z}_2, \mathbf{z}_1, \mathbf{z}_0 | \mathbf{d}_1, \mathbf{d}_2) = \frac{f(\mathbf{d}_2|\mathbf{z}_2) f(\mathbf{z}_2|\mathbf{z}_1) f(\mathbf{z}_1, \mathbf{z}_0|\mathbf{d}_1)}{f(\mathbf{d}_2)}, \tag{26}$$

$$\vdots$$

$$f(\mathcal{Z}|\mathcal{D}) = \frac{f(\mathbf{d}_L|\mathbf{z}_L) f(\mathbf{z}_L|\mathbf{z}_{L-1}) f(\mathbf{z}_{L-1}, \ldots, \mathbf{z}_0|\mathbf{d}_{L-1}, \ldots, \mathbf{d}_1)}{f(\mathbf{d}_L)}. \tag{27}$$

Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Make use of Markovian property

$$f(\mathbf{z}_1|\mathbf{d}_1) = \frac{f(\mathbf{d}_1|\mathbf{z}_1) \int f(\mathbf{z}_1|\mathbf{z}_0) f(\mathbf{z}_0)\, d\mathbf{z}_0}{f(\mathbf{d}_1)} = \frac{f(\mathbf{d}_1|\mathbf{z}_1) f(\mathbf{z}_1)}{f(\mathbf{d}_1)}, \tag{28}$$

$$f(\mathbf{z}_2|\mathbf{d}_1, \mathbf{d}_2) = \frac{f(\mathbf{d}_2|\mathbf{z}_2) \int f(\mathbf{z}_2|\mathbf{z}_1) f(\mathbf{z}_1|\mathbf{d}_1)\, d\mathbf{z}_1}{f(\mathbf{d}_2)} = \frac{f(\mathbf{d}_2|\mathbf{z}_2) f(\mathbf{z}_2|\mathbf{d}_1)}{f(\mathbf{d}_2)}, \tag{29}$$

$$\vdots$$

Filtering assumption

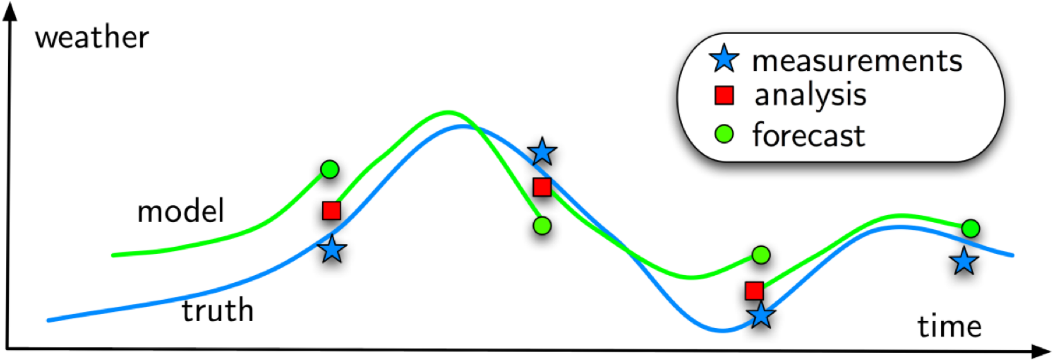
Approximation 3 (Filtering assumption)

We approximate the full smoother solution with a sequential data-assimilation solution. We only update the solution in the current assimilation window, and we do not project the measurement's information backward in time from one assimilation window to the previous ones.

Bayes' for the assimilation window

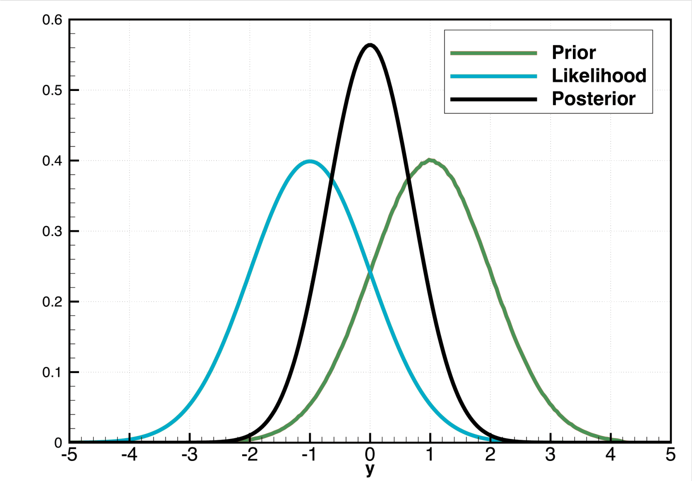
$$\boxed{f(\mathbf{z}|\mathbf{d}) = \frac{f(\mathbf{d}|\mathbf{z})f(\mathbf{z})}{f(\mathbf{d})} \qquad (30)}$$

Sequential data assimilation (EnKF, EKF, PF, PFF)

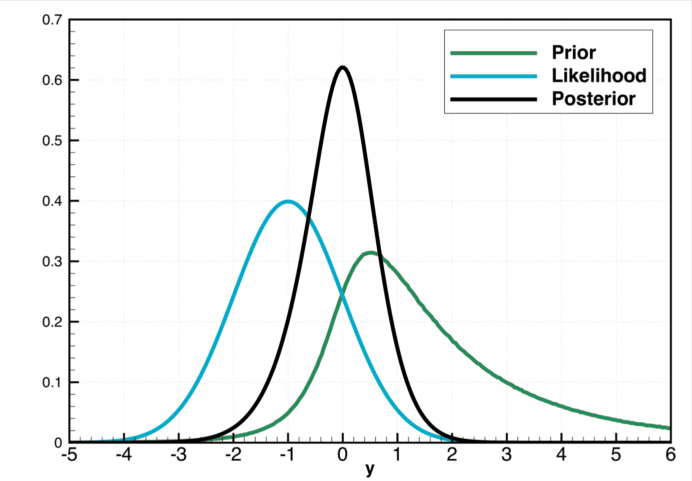


after Tandeo et al., 2018

Posterior is product of prior and likelihood

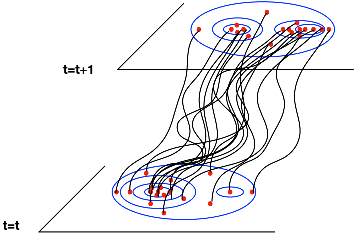


Posterior is product of prior and likelihood



Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Concept of particle filtering



Ensemble of N realisations (particles) to estimate pdf evolution

Posterior is proportional to prior times likelihood

Prior:

$$f(\mathbf{z}) = \Sigma_{j=1}^{N} \frac{1}{N} \delta(\mathbf{z} - \mathbf{z}_j)$$

Likelihood:

$f(\mathbf{d}|\mathbf{z}) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2}(\frac{\mathbf{d}-\mathbf{z}}{\sigma})^2}$

(Gaussian, can also be Lorentz function)

Posterior:

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

Particle filters for nonlinear data assimilation

Approximation 9 (Particle representation of the pdfs)

It is possible to approximate a probability density function by a finite ensemble of N model states (or particles) as

$$f(\mathbf{z}) \approx \sum_{j=1}^{N} \frac{1}{N} \delta(\mathbf{z} - \mathbf{z}_j), \tag{31}$$

where $\delta(\cdot)$ denotes the Dirac-delta function.

Importance sampling Monte Carlo

Use Monte Carlo samples to approximate the probability:

- generate N pseudo-random realisations z_j from $f(\mathbf{z}|\mathbf{d})$ with $j = 1, \ldots, N$.
- evaluate for each realisation the outcome of the forward model and compute the arithmetic average of the results.

Importance sampling Monte Carlo

Approximate the distribution $f(\mathbf{z}|\mathbf{d})$ by:

$$f(\mathbf{z}|\mathbf{d}) = \sum_{j=1}^{N} w_j \delta(\mathbf{z} - \mathbf{z}_j), \tag{32}$$

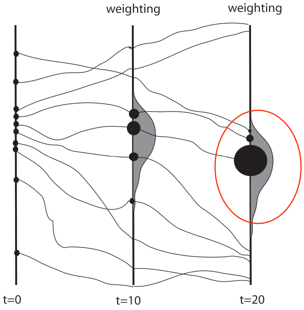
where $\delta_{\mathbf{z}_j}$ is a Dirac delta, and likelihood weights w_j are given by

$$w_j = \frac{f(\mathbf{d}|\mathbf{z}_j)}{f(\mathbf{d})} = \frac{f(\mathbf{d}|\mathbf{z}_j)}{\sum_{j=1}^{N} f(\mathbf{d}|\mathbf{z}_j))}. \tag{33}$$

denominator: normalization to ensure the weights add up to one,
$f(\mathbf{d}) = \int f(\mathbf{d}|\mathbf{z}) f(\mathbf{z})\, d\mathbf{z} \approx \sum_{j=1}^{N} f(\mathbf{d}|\mathbf{z}_j)$.

Degeneracy

Size of circles indicates weight

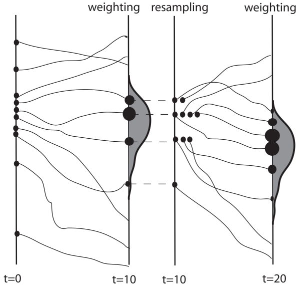


van Leeuwen, 2017, 10.5802/afst.1560

Samples that are closest to observations obtain largest weight. Some samples move very far from the observations and obtain a low weight. This means that effectively, there are only few samples left!

Importance Resampling

Size of circles indicates weight



van Leeuwen, 2017, 10.5802/afst.1560

Samples that are closest to observations obtain largest weight and are being duplicated. Low weight samples are removed from the ensemble.

Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Current challenges in data assimilation

- Estimating full distribution of uncertainty
- Dealing with nonlinearities
- Avoiding degeneracy and reducing computational costs
- Estimating model error
- Using data-assimilation outcomes to improve models
- Coupled data assimilation
- Support decision making

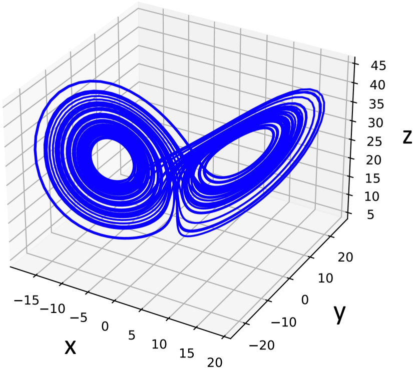
Examples (I)

EnKF with the Lorenz ’63 model

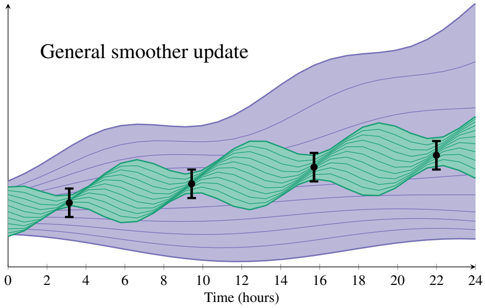
$$\frac{\partial x}{\partial t} = \sigma(y - x), \tag{34}$$

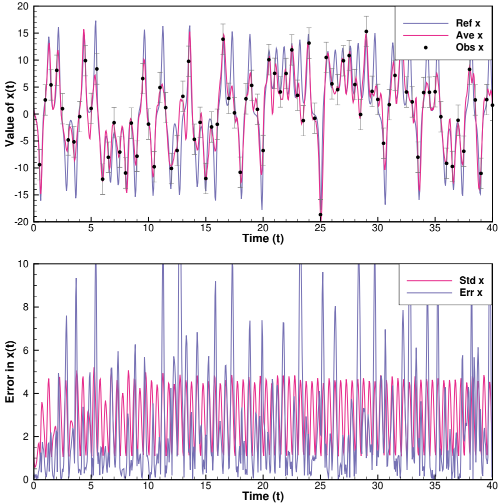
$$\frac{\partial y}{\partial t} = \rho x - y - xz, \tag{35}$$

$$\frac{\partial z}{\partial t} = xy - \beta z. \tag{36}$$



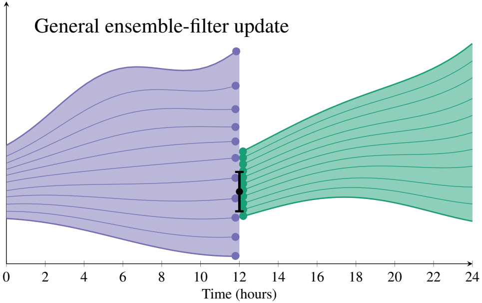
Ensemble Smoother

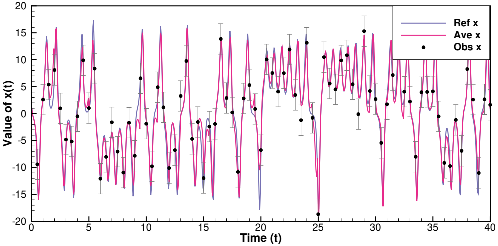


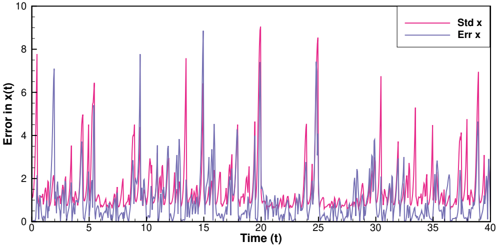


Available from: `https://github.com/geirev/Data-Assimilation-Fundamentals`

General Filter (EnKF)

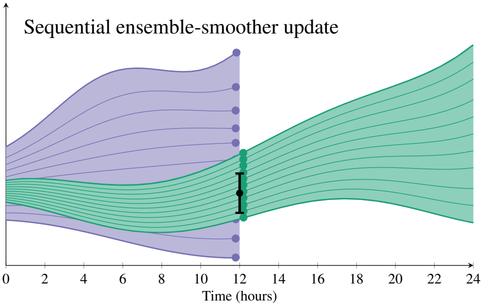


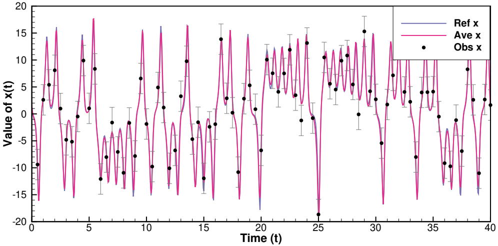


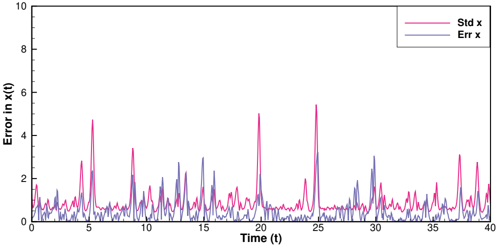


Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Recursive smoother, EnKS







Available from: `https://github.com/geirev/Data-Assimilation-Fundamentals`

Examples (II)

Particle filter, particle flow filter and EnKF for earthquake forecasting

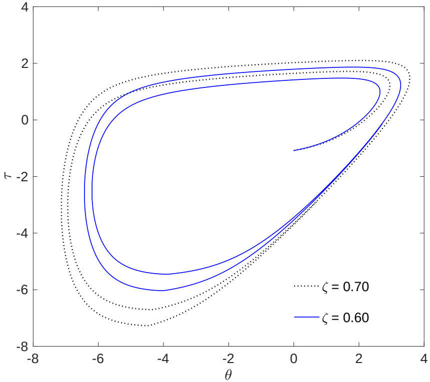
Available data:

1. Ground motion (GPS, seismometers)
2. Occassionaly, a subsurface measurement of strain

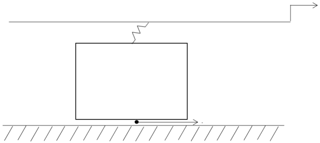
Challenges:

1. Very little data
2. Uncertainty in both model and observations

State- and parameter estimation with a 0D Burridge-Knopoff model



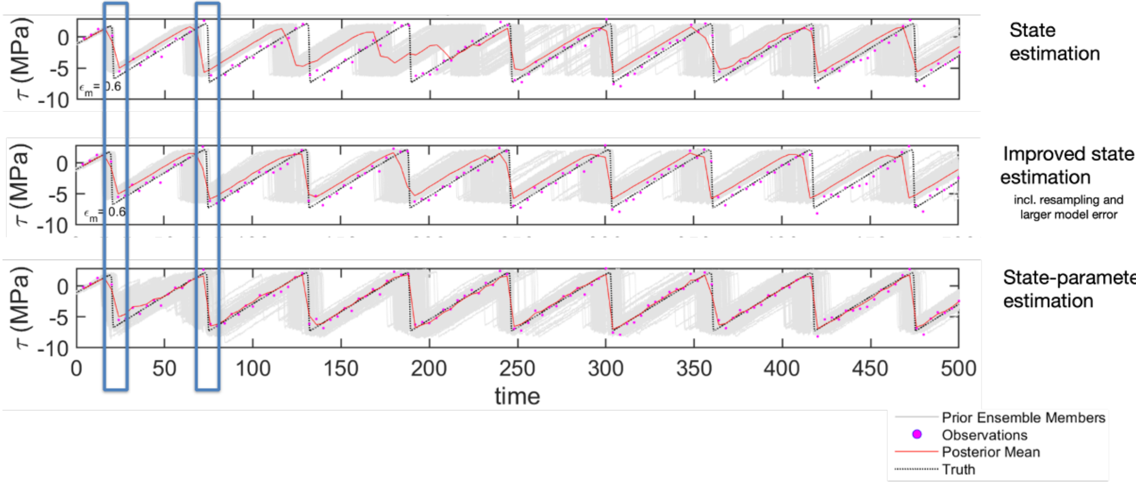
$$\frac{\partial \theta}{\partial t} = -v(\theta + (1 + \zeta)\ln v),$$
$$\frac{\partial u}{\partial t} = v - 1 + \eta,$$
$$\frac{\partial v}{\partial t} = -F^2\left(u + \frac{1}{\xi}(\theta + \ln v)\right),$$



Banerjee et al. (NPG, 2023):
https://npg.copernicus.org/articles/30/101/2023/

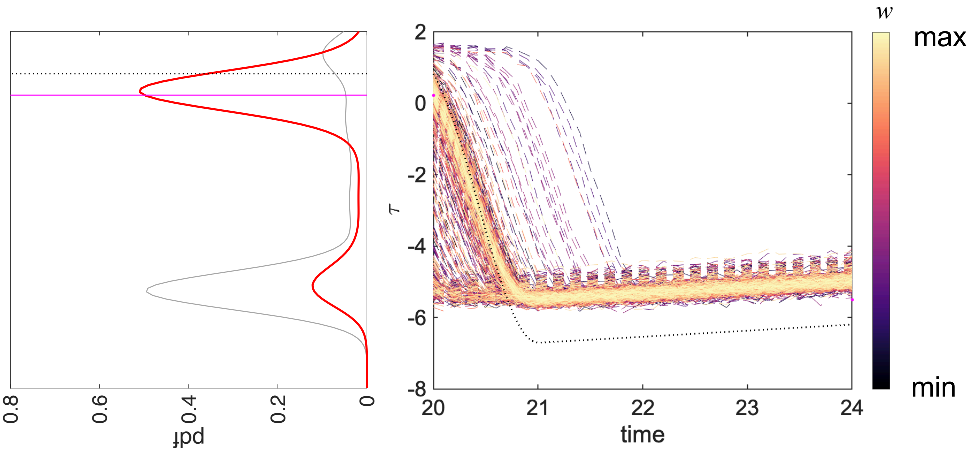
Parameter bias: Seismic cycle with 'true' (0.7) and 'biased' (0.6) ζ parameter

State vs State-Parameter estimation, using a particle filter



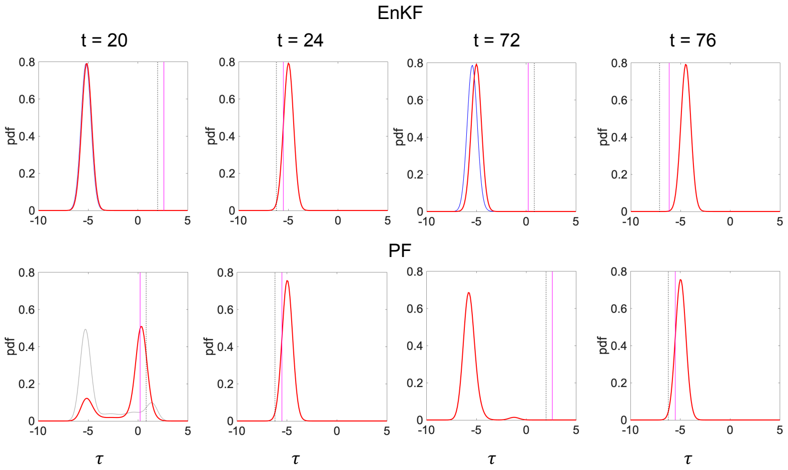
Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Shear-stress estimate with particle filter, for coseismic phase



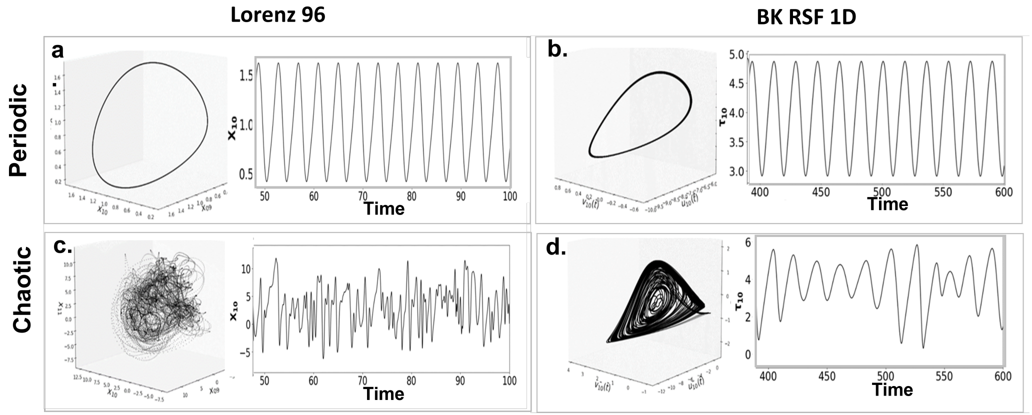
Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

PF tends to estimate shear stress better than EnKF



Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Ongoing: Burridge-Knopoff compared to Lorenz



Available from: https://github.com/geirev/Data-Assimilation-Fundamentals

Data assimilation methods

- can be derived based on Bayes' theorem
- provide information on model and data uncertainty
- need to deal with nonlinearities in the model
- can suffer from degeneracy and high computational costs

Current challenges:

- Estimating full distribution of uncertainty
- Dealing with nonlinearities
- Avoiding degeneracy and reducing computational costs
- Estimating model error
- Using data-assimilation outcomes to improve models
- Coupled data assimilation
- Support decision making

Thank you!

More details:

- Evensen, G., F.C. Vossepoel, P.J. van Leeuwen, Data Assimilation Fundamentals, open access, Springer, 2021
- Banerjee et al (NPG, 2023): https://npg.copernicus.org/articles/30/101/2023/
- Diab-Montero et al (in preparation)