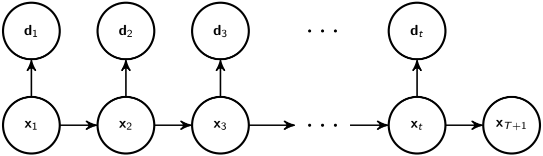
Multimodality in the Kalman Filter and Ensemble Kalman Filter

Maxime Conjard, Henning Omre

Department of Mathematical Sciences
NTNU

28/5/2018

Kalman Model



Process model's assumptions:

1. Gaussian initial distribution $f(\mathbf{x}_1)$
2. Single site dependence and conditional independence
3. Gauss-linear forward and likelihood model:

$$f(\mathbf{x}_{t+1}|\mathbf{x}_t) = \varphi_p(\mathbf{x}_{t+1}, \mathbf{B}\mathbf{x}_t, \mathbf{\Sigma}_{x|x})$$
$$f(\mathbf{d}_t|\mathbf{x}_t) = \varphi_p(\mathbf{d}_t, \mathbf{H}\mathbf{x}_t, \mathbf{\Sigma}_{d|x})$$

[Kalman(1960)],[Myrseth and Omre(2010)]

Kalman Model

Properties

Properties:

1. Analytically tractable, conjugate prior
2. Models linear unimodal processes

Selection Gaussian distribution

Let $\mathbf{A} \subset \mathbb{R}^q$,and

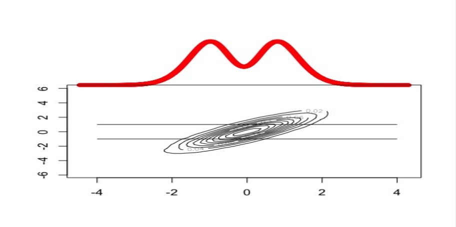
$$\begin{bmatrix} \mathbf{x}_0 \\ \boldsymbol{\nu} \end{bmatrix} \sim \varphi_{p+q}\left(\begin{bmatrix} \mathbf{x}_0 \\ \boldsymbol{\nu} \end{bmatrix}; \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_{x_0} \\ \boldsymbol{\mu}_{\nu} \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right)$$

then $\mathbf{x}_{0,A} = [\mathbf{x}_0 | \boldsymbol{\nu} \in A]$ is Selection Gauss.

Flexibility

1. Skewness
2. Multimodality
3. Conjugate prior to a Gauss-linear likelihood and forward model

Bimodality



[Azzalini and Valle(1996)],[Rimstad and Omre(2014)]

Selection Gaussian Kalman Model

Model's assumptions:

1. Selection Gaussian initial distribution $f(\mathbf{x}_1)$
2. Single site response and conditional independence
3. Gauss-linear forward and likelihood model:

$$
\begin{aligned}
f(\mathbf{x}_{t+1}|\mathbf{x}_t) =& \varphi_p(\mathbf{x}_{t+1}; \mathbf{B}\mathbf{x}_t, \mathbf{\Sigma}_{x|x}) \\
f(\mathbf{d}_t|\mathbf{x}_t) =& \varphi_p(\mathbf{d}_t; \mathbf{H}\mathbf{x}_t, \mathbf{\Sigma}_{d|x})
\end{aligned}
$$

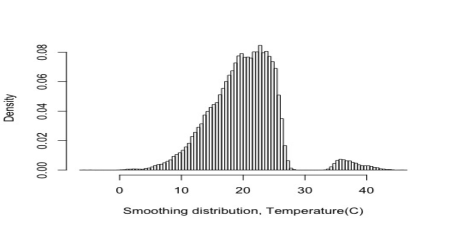
[Naveau et al.(2005)]

Selection Gaussian Kalman Model

Properties

1. Analytically tractable
2. Models multimodality
3. Easy to implement

Marginal smoothing distribution



Implementation

1. We start with $\begin{bmatrix} \mathbf{x}_1 \\ \boldsymbol{\nu} \end{bmatrix}$ that is Gaussian
2. We increment (update) to $\begin{bmatrix} \mathbf{x}_1 \\ \boldsymbol{\nu} \\ \mathbf{d}_1 \end{bmatrix}$ that is still Gaussian
3. We increment (forward) to $\begin{bmatrix} \mathbf{x}_2 \\ \mathbf{x}_1 \\ \boldsymbol{\nu} \\ \mathbf{d}_1 \end{bmatrix}$
4. etc ...

Implementation

1. Access to Kalman filtering $\mathbf{x}_t|\mathbf{d}_1, ..., \mathbf{d}_t$, smoothing $\mathbf{x}_s|\mathbf{d}_1, ..., \mathbf{d}_t, s \leq t$ and inversion $\mathbf{x}_1|\mathbf{d}_1, ..., \mathbf{d}_T$.
2. Fast computation
3. Conserve a Gaussian structure

Example: Backtracking the 2D Heat equation

The heat equation:

$$\frac{\partial T}{\partial t} - \nabla^2 T = 0$$
$$\nabla T . \mathbf{n} = 0$$

Modelled using finite differences on $[0, 1] \times [0, 1]$, it gives the following Gauss-linear forward model:

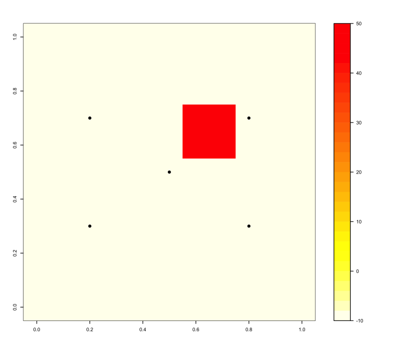
$$f(\mathbf{T}_{t+1}|\mathbf{T}_t) = \varphi_p(\mathbf{T}_{t+1}, \mathbf{B}\mathbf{T}_t, \mathbf{\Sigma}_{T|T}) \quad (1)$$

Data is collected at 5 different locations using the following Gauss-linear likelihood model:

$$f(\mathbf{d}_t|\mathbf{T}_t) = \varphi_p(\mathbf{d}_t, \mathbf{H}\mathbf{T}_t, \mathbf{\Sigma}_{d|T}) \quad (2)$$

Example: Backtracking the 2D Heat equation

Initial Heat map

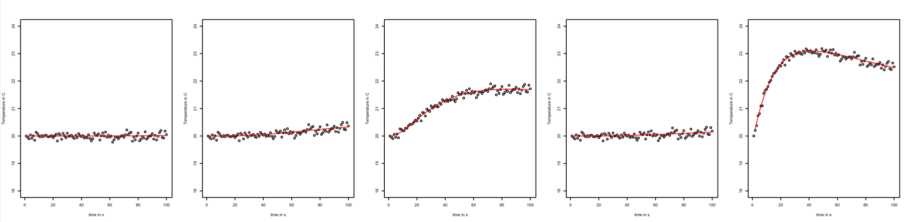


Facts

1. Discontinuous initial conditions
2. 5 data collection points

Example: Backtracking the 2D Heat equation

Data collected Vs True process

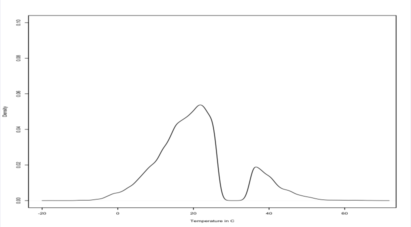


Parameters

1. $dt = 1s$
2. $\mathbf{\Sigma}_{d|x} = 0.01\mathbf{I}$.

Initial model: A reflection of our a priori knowledge

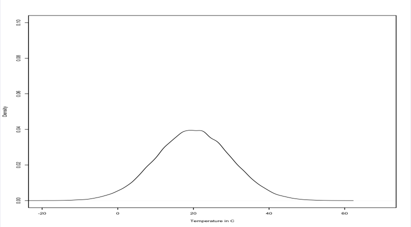
Scenario 1: Sel-Gauss initial model



Properties

1. Two lobes.

Scenario 2:Gaussian initial model



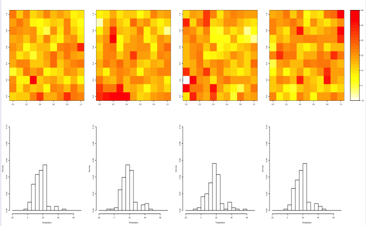
Properties

1. $E(\mathbf{x}_1) = 20.$
2. $\mathbf{Var}(\mathbf{x}_1) = 100.$

Initial model: A reflection of our a priori knowledge

Realizations from the initial distribution:

Sel-Gauss initial distribution



Gaussian initial distribution

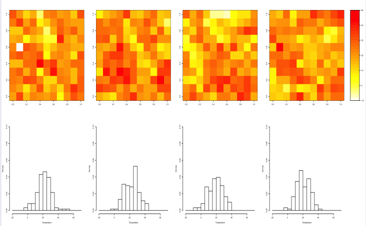
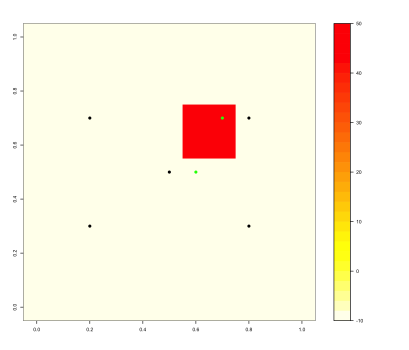


Exhibit $[x_{1,i}|d_1, \ldots, d_T]$ at 2 different locations

Initial Heat map

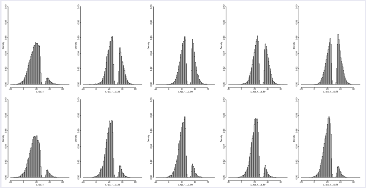


Facts

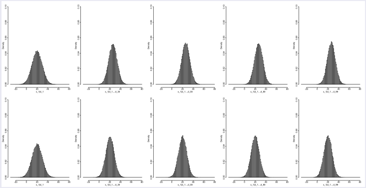
1. Compare the marginal distribution at two different point
2. One inside, one outside

Exhibit $[x_{1,i}|d_1, ..., d_T]$ at 2 different locations

Sel-Gauss initial model



Gaussian initial model

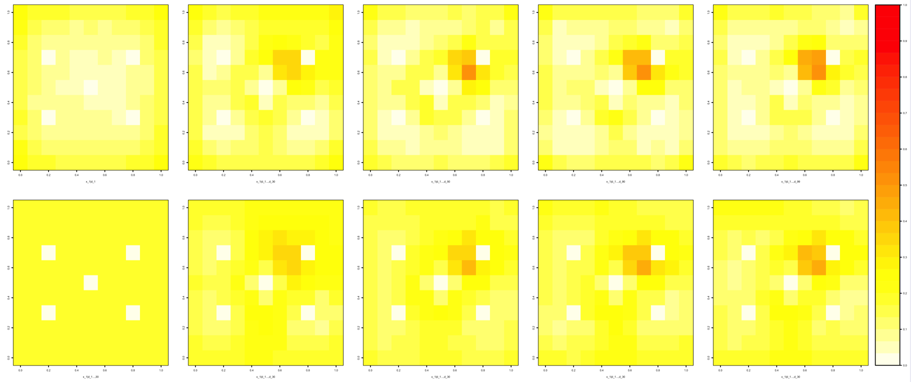


Global behavior

We define $LR(\mathbf{x})$ as:

$$LR_i(\mathbf{x}) = P(x_i > 28,75) \quad \forall i \in [1, p]$$

$LR(\mathbf{x}_1 | d_1, ..., d_t)$ for different values of t



Algorithm:EnKF for the Sel-Gauss (EnKF(SG))

- Initiate

$$n_e = \text{no. of ensemble members}$$
$$\begin{bmatrix} \mathbf{x}_0^{u(i)} \\ \boldsymbol{\nu}_0^{u(i)} \end{bmatrix}, \; i = 1, ..., n_e \text{ iid. } f(\mathbf{x}_0^u, \boldsymbol{\nu}_0^u) = N(\boldsymbol{\mu}_0^u, \boldsymbol{\Sigma}_0^u)$$
$$\mathbf{d}_0^{(i)} = \mathbf{H}\mathbf{x}_0^{u(i)} + \boldsymbol{\eta}_0^i, \; i = 1, ..., n_e \text{ with } \boldsymbol{\eta}_0 \sim N(0, \boldsymbol{\Sigma}_0^{d|x})$$

- Iterate $t = 0, ..., T$

$$\text{Estimate } \boldsymbol{\Sigma}_{x,\nu,d} \text{ from } \{(\mathbf{x}_t^{u(i)}, \boldsymbol{\nu}_t^{u(i)}, \mathbf{d}_t^i), i = 1, ..., n_e\}$$
$$\begin{bmatrix} \mathbf{x}_t^{c(i)} \\ \boldsymbol{\nu}_t^{c(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_t^{u(i)} \\ \boldsymbol{\nu}_t^{u(i)} \end{bmatrix} + \boldsymbol{\Gamma}_{x,\nu,d}\boldsymbol{\Sigma}_d^{-1}(\mathbf{d}_t - \mathbf{d}_t^i), \; i = 1, ..., n_e$$
$$\begin{bmatrix} \mathbf{x}_{t+1}^{u(i)} \\ \boldsymbol{\nu}_{t+1}^{u(i)} \end{bmatrix} = \begin{bmatrix} g(\mathbf{x}_t^{c(i)}) \\ \boldsymbol{\nu}_t^{c(i)} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\delta}_t \\ \mathbf{0} \end{bmatrix}, \; i = 1, ..., n_e \text{ with } \boldsymbol{\delta}_t \sim N(0, \boldsymbol{\Sigma}_t^{x|x})$$
$$d_{t+1}^{(i)} = \mathbf{H}\mathbf{x}_{t+1}^{u(i)} + \boldsymbol{\eta}_{t+1}^i, \; i = 1, ..., n_e \text{ with } \boldsymbol{\eta}_{t+1} \sim N(0, \boldsymbol{\Sigma}_{t+1}^{d|x})$$

- Estimate $\boldsymbol{\mu}_{T+1}^u, \boldsymbol{\Sigma}_{T+1}^u$ and assess $f(\mathbf{x}_{T+1}|\mathbf{d}_0, ..., \mathbf{d}_T, \nu \in A)$

Algorithm:EnKF for the Sel-Gauss (EnKF(SG))

1. Non gaussian output: Ensemble of $\mathbf{x}, \boldsymbol{\nu}$ rather than $\mathbf{x}|\boldsymbol{\nu} \in A$
2. Forward step made easy by :

$$g(\mathbf{x}_t|\nu \in A) = g(\mathbf{x}_t)|\boldsymbol{\nu} \in A$$

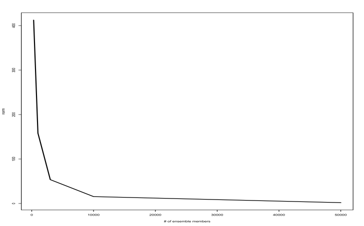
Test on a linear forward model: Previous example

Consider now:

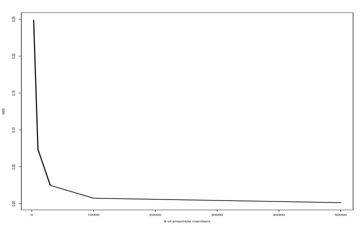
$$\begin{bmatrix} \mathbf{x}_{t+1}^{u(i)} \\ \boldsymbol{\nu}_{t+1}^{u(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{B} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{t}^{c(i)} \\ \boldsymbol{\nu}_{t}^{c(i)} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\delta}_t \\ \mathbf{0} \end{bmatrix} \quad i = 1, \dots, n_e$$

We "show" that the EnKF(SG) converges numerically to the Selection Gauss Kalman Filter as $n_e \to \infty$ when the forward model is linear.

Expected Value: Norm of the difference for $x_2|d_1, d_2$



Covariance matrix: Norm of the difference for $x_2|d_1, d_2$



Ongoing work: Use EnKF for parameter estimation

1. Idea: Put a Sel-Gauss prior on the parameter, one lobe per possible value for the parameter (diffusivity coefficient, but also porosity).
2. Use the EnKF to estimate the parameters.

A. Azzalini and A. Dalla Valle.
The multivariate skew-normal distribution.
Biometrika, 83(4):715–726, 1996.

E Kalman.
A new approach to linear filtering and prediction problems.
Transactions of the ASME–Journal of Basic Engineering, 82(Series D):35–45, 1960.

I. Myrseth and H. Omre.
The Ensemble Kalman Filter and Related Filters, pages 217–246.
John Wiley and Sons, Ltd, 2010.

P. Naveau, M. Genton, and X. Shen.
A skewed kalman filter.
Journal of Multivariate Analysis, 94(2):382 – 400, 2005.

K. Rimstad and H. Omre.
Skew-gaussian random fields.
Spatial Statistics, 10:43 – 62, 2014.