Ensemble methods using a selection-Gaussian initial distributions

Maxime Conjard and Henning Omre

Department of Mathematical Sciences
NTNU

04/06/2019

Selection Gaussian distribution

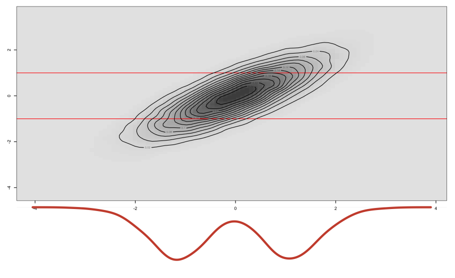
Let $A \subset \mathbb{R}^q$, and

$$\begin{bmatrix} \mathbf{r} \\ \boldsymbol{\nu} \end{bmatrix} \sim N_{p+q}\left(\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_r \\ \boldsymbol{\mu}_\nu \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_r & \boldsymbol{\Gamma}_{r,\nu} \\ \boldsymbol{\Gamma}_{\nu,r} & \boldsymbol{\Sigma}_\nu \end{bmatrix}\right)$$

A selection Gaussian random variable is defined as :

$$\mathbf{r}_A = (\mathbf{r} | \boldsymbol{\nu} \in A)$$

Example 1D



Selection Gaussian distribution

1. Let $\varphi_p(\mathbf{r}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote the pdf of a Gaussian distribution with parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
2. Let $\Phi_q(A; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = P(Y \in A)$.
3. Then, $\mathbf{r}_A = (\mathbf{r} | \boldsymbol{\nu} \in A)$ has the following pdf:

$$f_{p+q}(\mathbf{r}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, A) = \varphi_p(\mathbf{r}; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r) \frac{\Phi_q(A; \boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*)}{\Phi_q(A; \boldsymbol{\mu}_\nu, \boldsymbol{\Sigma}_\nu)}$$

   with:

$$\begin{aligned} \boldsymbol{\mu}^* =& \boldsymbol{\mu}_\nu + \boldsymbol{\Gamma}_{\nu,r} \boldsymbol{\Sigma}_r^{-1} (\mathbf{r} - \boldsymbol{\mu}_r) \\ \boldsymbol{\Sigma}^* =& \boldsymbol{\Sigma}_\nu - \boldsymbol{\Gamma}_{\nu,r} \boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Gamma}_{r,\nu} \end{aligned}$$

Note: Gaussian pdfs are a subset of the selection-Gaussian pdfs (consider $\boldsymbol{\Gamma}_{\nu,r} = 0$).

Selection Gaussian distribution

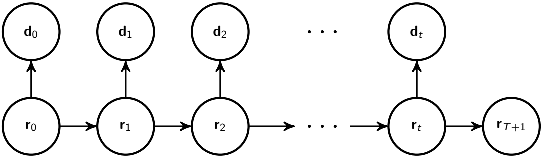
Properties

1. Can represent skewness, multimodalilty, and heavy-tailedness.
2. Conjugate prior to a Gauss-linear likelihood and forward model: analytically tractable Bayesian recursive algorithm (Selection Kalman Filter)

Bayesian inversion

$$\underbrace{f(\mathbf{r}_A|\mathbf{d})}_{\downarrow \atop selection-Gaussian} \propto \underbrace{f(\mathbf{d}|\mathbf{r}_A)}_{\downarrow \atop Gauss-linear} \underbrace{f(\mathbf{r}_A)}_{\downarrow \atop selection-Gaussian}$$

Hidden Markov Model



1. Selection-Gaussian initial distribution $f(\mathbf{r}_0)$
2. Forward and likelihood model:

$$[\mathbf{r}_{t+1}|\mathbf{r}_t] = \omega_t(\mathbf{r}_t, \epsilon_t^r) \sim f(\mathbf{r}_{t+1}|\mathbf{r}_t)$$
$$[\mathbf{d}_t|\mathbf{r}_t] = \gamma_t(\mathbf{r}_t, \epsilon_t^d) \sim f(\mathbf{d}_t|\mathbf{r}_t)$$

Augmented state space

Recall:

$$\begin{bmatrix} \mathbf{r} \\ \boldsymbol{\nu} \end{bmatrix} \sim N_{p+q}\left(\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_r \\ \boldsymbol{\mu}_\nu \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_r & \boldsymbol{\Gamma}_{r,\nu} \\ \boldsymbol{\Gamma}_{\nu,r} & \boldsymbol{\Sigma}_\nu \end{bmatrix}\right)$$

with $\mathbf{r}_A = (\mathbf{r}|\boldsymbol{\nu} \in A)$. Since:

$$[g(\mathbf{r}_t|\nu \in A)] = [g(\mathbf{r}_t)|\boldsymbol{\nu} \in A]$$

Augmented forward model:

$$\begin{bmatrix} \mathbf{r}_{t+1} \\ \boldsymbol{\nu}_{t+1} \end{bmatrix} = \begin{bmatrix} \omega_t(\mathbf{r}_t, \epsilon_t^r) \\ \boldsymbol{\nu}_t \end{bmatrix}.$$

Conditioning done updating the augmented state vector.

Selection-EnKF (SEnKF)

Time series of ensembles defined as
$\mathbf{e}_t = \{(\mathbf{r}_t^{u(i)}, \boldsymbol{\nu}_t^{u(i)}, \mathbf{d}_t^i), i = 1, ..., n_e\}, \forall t = 0, \cdots, T+1$ and has the following covariance matrix:

$$\boldsymbol{\Sigma}_{r\nu d} = \begin{bmatrix} \boldsymbol{\Sigma}_{r\nu} & \boldsymbol{\Gamma}_{r\nu,d} \\ \boldsymbol{\Gamma}_{d,r\nu} & \boldsymbol{\Sigma}_d \end{bmatrix}$$

Simple rewrite of the EnKF with an augmented state space.

Selection-EnKF (SEnKF)

- Initiate

$$
\begin{aligned}
&n_e = \text{no. of ensemble members}\\
&\begin{bmatrix} \mathbf{r}_0^{u(i)} \\ \boldsymbol{\nu}_0^{u(i)} \end{bmatrix} \sim N(\boldsymbol{\mu}_0^u, \boldsymbol{\Sigma}_0^u),\ i = 1, \dots, n_e\\
&\mathbf{d}_0^{(i)} = \gamma_0(\mathbf{r}_0^{u(i)}, \epsilon_0^{d(i)}),\ \epsilon_0^{d(i)} \sim \mathcal{U}_n[0,1] \text{ iid } , i = 1, \dots, n_e
\end{aligned}
$$

- Iterate $t = 0, \dots, T$

$$
\begin{aligned}
&\text{Estimate } \boldsymbol{\Sigma}_{r\nu d} \text{ from } \mathbf{e}_t \longrightarrow \hat{\boldsymbol{\Sigma}}_{r\nu d}\\
&\begin{bmatrix} \mathbf{r}_t^{c(i)} \\ \boldsymbol{\nu}_t^{c(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_t^{u(i)} \\ \boldsymbol{\nu}_t^{u(i)} \end{bmatrix} + \hat{\boldsymbol{\Gamma}}_{r\nu,d} \hat{\boldsymbol{\Sigma}}_d^{-1} (\mathbf{d}_t - \mathbf{d}_t^i), \quad i = 1, \dots, n_e\\
&\begin{bmatrix} \mathbf{r}_{t+1}^{u(i)} \\ \boldsymbol{\nu}_{t+1}^{u(i)} \end{bmatrix} = \begin{bmatrix} \omega_t(\mathbf{r}_t^{c(i)}, \epsilon_t^{r(i)}) \\ \boldsymbol{\nu}_t^{c(i)} \end{bmatrix},\ \epsilon_t^{r(i)} \sim \mathcal{U}_n[0,1] \text{ iid},\ i = 1, \dots, n_e\\
&\mathbf{d}_{t+1}^{(i)} = \gamma_{t+1}(\mathbf{r}_{t+1}^{u(i)}, \epsilon_{t+1}^{d(i)}),\ \epsilon_{t+1}^{d(i)} \sim \mathcal{U}_n[0,1] \text{ iid},\ i = 1, \dots, n_e
\end{aligned}
$$

- End iterate
- Estimate $\boldsymbol{\mu}_{T+1}^u, \boldsymbol{\Sigma}_{T+1}^u$ from $\mathbf{e}_{T+1}$

Selection-EnKF (SEnKF)

Assume that:

$$\begin{matrix}\mathbf{r_{T+1}} \\ \boldsymbol{\nu}\end{matrix} | \mathbf{d} \sim N_{p+q}\left(\tilde{\boldsymbol{\mu}}^u_{T+1}, \tilde{\boldsymbol{\Sigma}}^u_{T+1}\right)$$

Block sampling MH algorithm to estimate:

$$\mathbf{r}_A = [\mathbf{r}|\boldsymbol{\nu} \in A, \mathbf{d}] \rightarrow Selection - Gaussian$$

Synthetic case studies

Diffusion equation

$$\begin{aligned}\frac{\partial r_t(\mathbf{x})}{\partial t} - \nabla \cdot (\lambda(\mathbf{x}) \nabla r_t(\mathbf{x})) =& q \\ \nabla r_t(\mathbf{x}) \cdot \mathbf{n} =& 0\end{aligned}$$

Consider $\mathbf{r}_0, \dots, \mathbf{r}_{T+1}$ and $\boldsymbol{\lambda}$ to be random variables. Finite difference $\rightarrow$ perfect forward model:

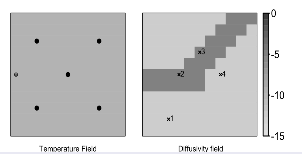
$$\omega([\mathbf{r}_t, \boldsymbol{\lambda}], \mathbf{0}) = \begin{bmatrix} \omega^*(\mathbf{r}_t, \boldsymbol{\lambda}) \\ \boldsymbol{\lambda} \end{bmatrix}$$

Observations collected at 5 locations with a Gauss-linear likelihood model:

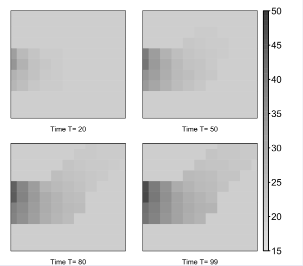
$$[\mathbf{d}_t | \mathbf{r}_t] = \gamma_t(\mathbf{r}_t, \epsilon_d) = \mathbf{H}\mathbf{r}_t + \epsilon_d, \quad \epsilon_d \sim N(0, \sigma_d^2 \mathbb{I})$$

Test I: Assessing the static diffusivity field

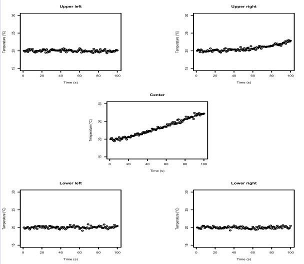
Initial Heat map



Temperature field



Data collection, $\sigma_d = 0.1$

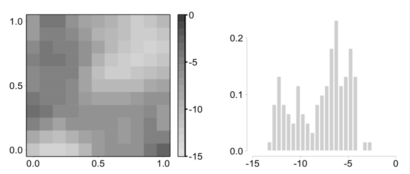


Prior model

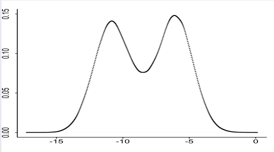
Prior knowledge

1. Equiprobable lobes the true diffusivities marginally
2. Prior model spatially stationary bar border effects.
3. Spatial smoothness

Realization and spatial histogram

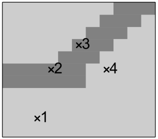


Marginal distribution



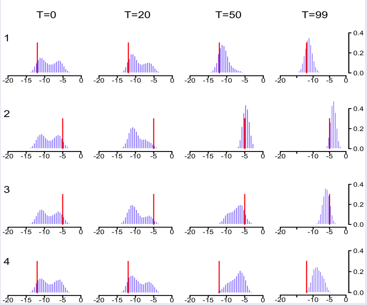
Marginal distributions at the monitoring locations

Posterior distribution $[log(\boldsymbol{\lambda})_i | d_{0:T}]$:

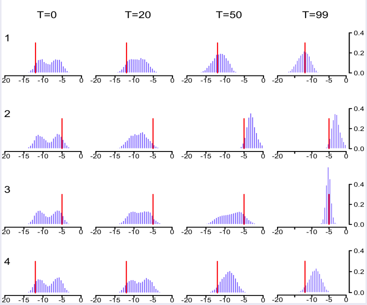


Diffusivity field

SEnKF



EnKF



MMAP and RMSE

MMAP

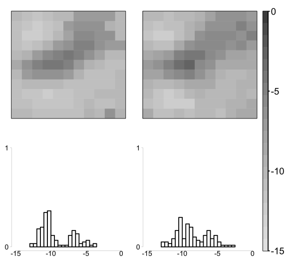
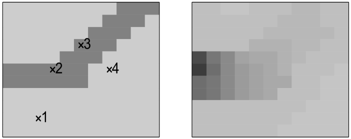


Figure: SENKF (left) and ENKF (right)



Diffusivity field

Temperature field at T=99

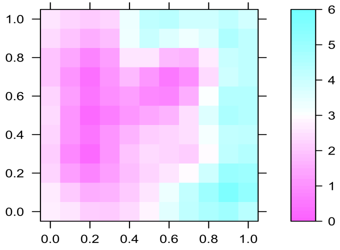
RMSE

	SEnKF	ENKF
$RMSE_{T=99}$	2.39	2.83

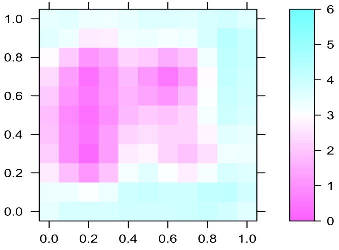
Table: RMSE of the MMAP prediction at time $T = 99$.

Prediction variance

SEnKF

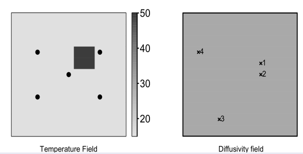


EnKF

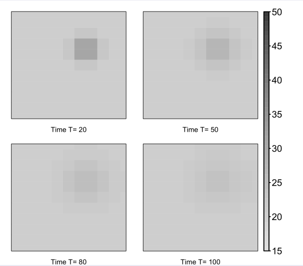


Test II: Assessing the initial dynamic field

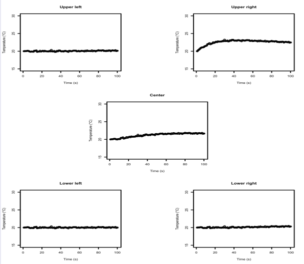
Initial Heat map



Temperature field



Data collection, $\sigma_d = 0.1$

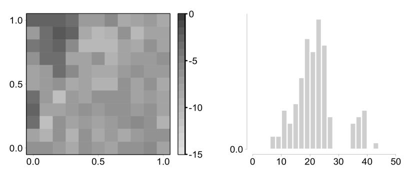


Prior model

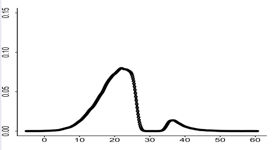
Prior knowledge

1. Two lobes about the true initial temperatures
2. Prior model spatially stationary bar border effects.
3. Spatial smoothness

Realization and spatial histogram

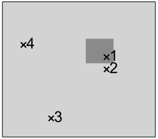


Marginal distribution



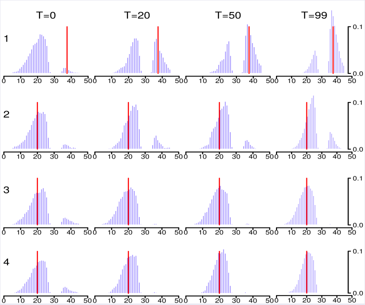
Marginal distributions at the monitoring locations

Posterior distribution $[(r_0)_i|d_{0:T}]$:

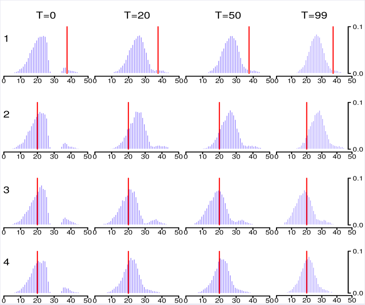


Diffusivity field

SEnKS



EnKS



MMAP and RMSE

MMAP

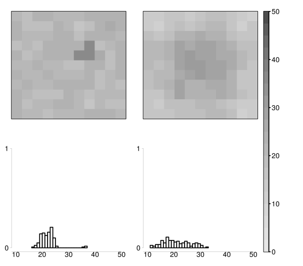
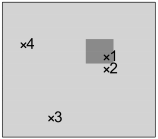


Figure: SENKS (left) and ENKS (right)



Diffusivity field

RMSE

	SEnKS	ENKS
$RMSE_{T=99}$	3.22	5.25

Table: RMSE of the MMAP prediction o at time $T = 99$.

Challenges

Sampling of posterior distribution sensitive to covariance matrix $\rightarrow$ large $n_e \simeq 10000$

1. **SVD**: comparable for $n_e \downarrow$
2. Reference covariance matrix $n_e \simeq 10000$ - The distance to the reference matrix decrease as $n_e \uparrow$
3. **Localization**: seems to solve the problem ($n_e = 500$) at the expense of contrast.

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

Geir Evensen.
Data assimilation. The ensemble Kalman filter, volume 307.
01 2006.
ISBN 9783642037108.

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.

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

Henning Omre and Kjartan Rimstad.
Bayesian spatial inversion and conjugate selection gaussian prior models, 2018.

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