An ETKF approach for initial state and parameter estimation in ice sheet modelling

Bertrand Bonan

University of Reading

DARC Seminar, Reading – October 30, 2013

Context

Work supervised by Maëlle Nodet[1,2] and Catherine Ritz[3].

Joined work also with Vincent Peyaud[3].

- [1]: Inria, Laboratoire Jean Kuntzmann (Grenoble)
- [2]: Université Joseph Fourier – Grenoble 1
- [3]: CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement (Grenoble)

Part of project ADAGe funded by French "Agence Nationale de la Recherche".

http://www-lgge.obs.ujf-grenoble.fr/pdr/ADAGe/

Outline

1. Problem description
2. Large-scale ice sheet model
3. Data assimilation
   - Twin experiment
   - ETKF and LETKF
   - Initial ensemble
   - Results

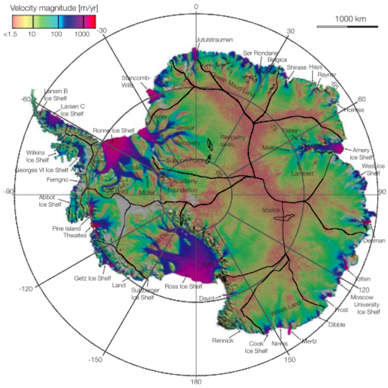
Outline

Sea level change: Antarctica & Greenland contribution

Motivations: Computation of the ice discharge of Antarctica and Greenland in the near future, thanks to simulations of polar ice sheet model.

Ice discharge:

- narrow outlets,
- closely linked to ice velocities,
- highly sensitive to basal friction parameters,
- highly sensitive to bedrock topography.



Surface ice velocities [Rignot *et al.* 2011]

Basal parameters

Typically 2 types of basal parameters are mandatory for simulation:

- **bedrock topography** $B_{soc}(x)$
- **basal friction parameters** in basal shear stress relationship

$$\boldsymbol{\tau}_b(x,t) = \begin{cases} -C(x,t)|\mathbf{u}_b(x,t)|^{m(x,t)-1}\mathbf{u}_b(x,t) & \text{if } T_b(x,t) = T_{melt} \\ 0 & \text{if } T_b(x,t) < T_{melt} \end{cases}$$

with $\mathbf{u}_b$ basal velocity, T_b basal ice temperature, $C > 0$ and m basal friction parameters.

Simplified relationship

$$\boldsymbol{\tau}_b(x,t) = \begin{cases} -\beta(x)\mathbf{u}_b(x,t) & \text{if } T_b(x,t) = T_{melt} \\ 0 & \text{if } T_b(x,t) < T_{melt} \end{cases}$$

with $\beta > 0$.

What is behind β

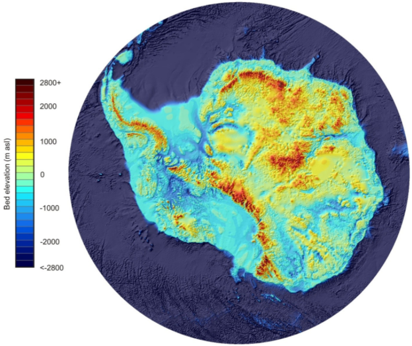
- sliding of ice over the bed
- deformation of the bed

due to

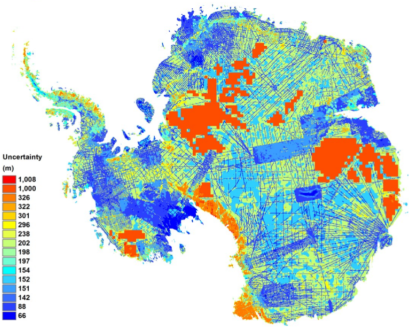
- effective pressure = ice pressure - water pressure
  (ex: thin water layer between ice and bedrock, water in cavities, ...)
- rock debris
- sediments
- ...

Poorly known basal parameters ...

Illustration for bedrock topography with Bedmap2 [Fretwell *et al.* 2013]



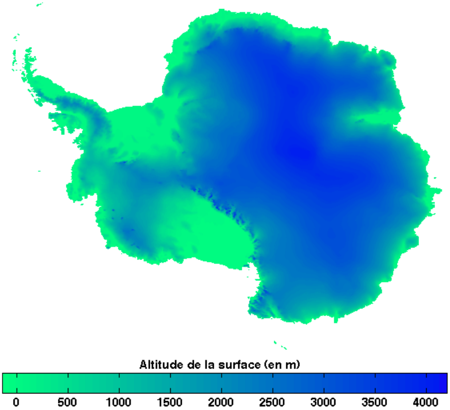
Bedmap2 estimation of bed topography



Estimated uncertainty of bed topography

... Fortunately surface observations

- Ice velocity
- Surface elevation (ex.: [Bamber et al, 2009])



Overview

What we want:

- Build estimation of actual basal parameters:
  - **bedrock topography** B_{soc}
  - **sliding coefficient** $\beta = 10^{\alpha}$
- Give an estimation of the uncertainty of our estimation.

What we have:

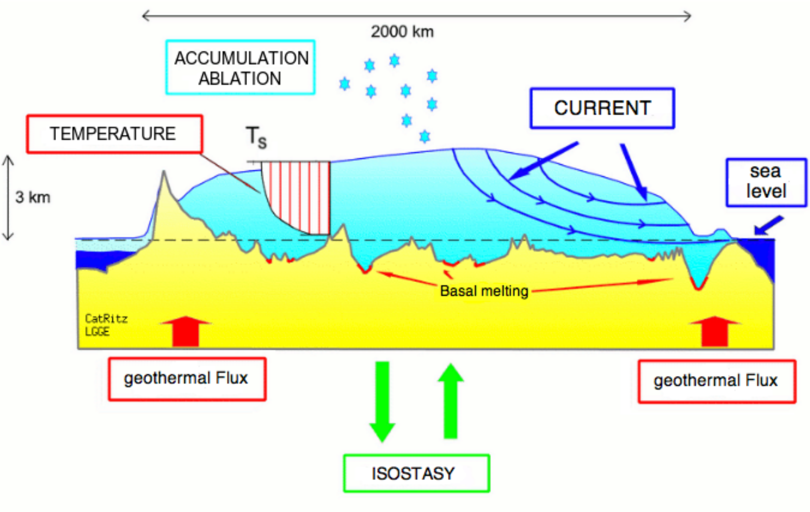
- 20 years of observations of surface elevation and surface velocities
- Sparse observations of bedrock topography
- Time dependent ice sheet model

What we need: efficient data assimilation system

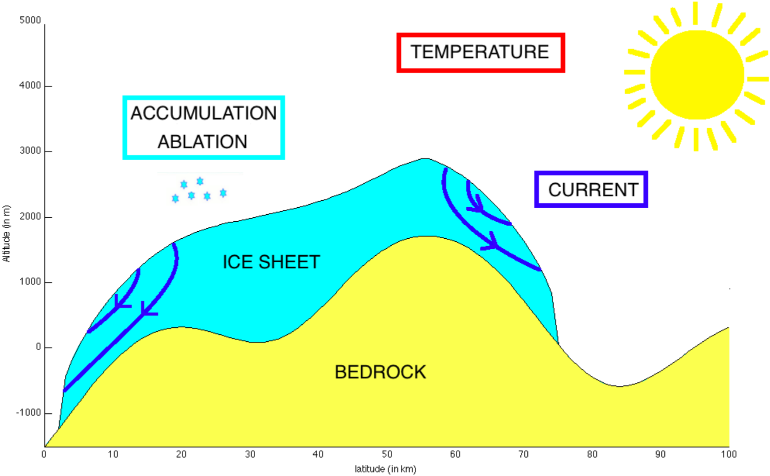
To develop the method: **Synthetic experiments along a flowline**

Outline

Ice dynamics processes



Simplified physics



Model equations: mass balance

Flowline SIA model (1D + time) with a sliding law

Mass balance equation:

$$\frac{\partial H}{\partial t} = \dot{b}_m - \frac{\partial(\overline{U}H)}{\partial x} \qquad\qquad H_{|t=0} = H_0$$

with

- x latitude, t time
- $H(x,t)$ ice thickness, $H_0(x)$ initial ice thickness
- $\overline{U}(x,t)$ ice velocity averaged over ice thickness
- $\dot{b}_m(x,t)$ surface mass balance rate

Model equations: dynamics (1)

Vertically averaged ice velocity is a diagnostic variable = no partial derivative in time involved, computed from geometry at each time step

$$\begin{aligned} \overline{U} &= u_{def} + u_{slid} \\ S &= B_{soc} + H \\ H &\geq 0 \end{aligned}$$

with

- $S(x,t)$ surface elevation
- $B_{soc}(x,t)$ bedrock topography
- u_{def} ice velocity component due to ice deformation
- u_{slid} ice velocity component due to sliding

Model equations: dynamics (2)

u_{def} is the averaged velocity due to ice deformation. Under SIA approximation

$$u_{def} = -a_1 \frac{\partial S}{\partial x} \frac{H^2}{3} - a_2 \left(\frac{\partial S}{\partial x} \right)^3 \frac{H^4}{5}$$

with a_1, a_2 given coefficients

u_{slid} is the velocity component due to basal sliding

$$u_{slid} = -\frac{1}{\beta} \rho_i \, g H \frac{\partial S}{\partial x}$$

with $\beta > 0$ basal friction coefficient.

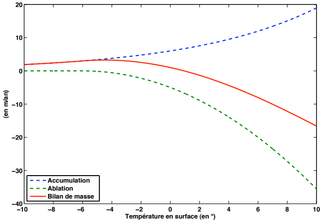
Model equations: mass balance rate

$$\dot{b}_m(x,t) = Acc(x,t) + Abl(x,t)$$

- $Acc(x,t) \geq 0$ ice accumulation rate (mainly snow precipitation)
- $Abl(x,t) \leq 0$ ice ablation rate (snow melting)

Both are function of surface temperature T_s

$$T_s(x,t) = F_{clim}(t) + \lambda x + \gamma S(x,t)$$



Outline

Outline

Twin experiment

What we control: at each grid point

- Ice thickness H
- Bedrock topography B_{soc}
- Basal parameter β

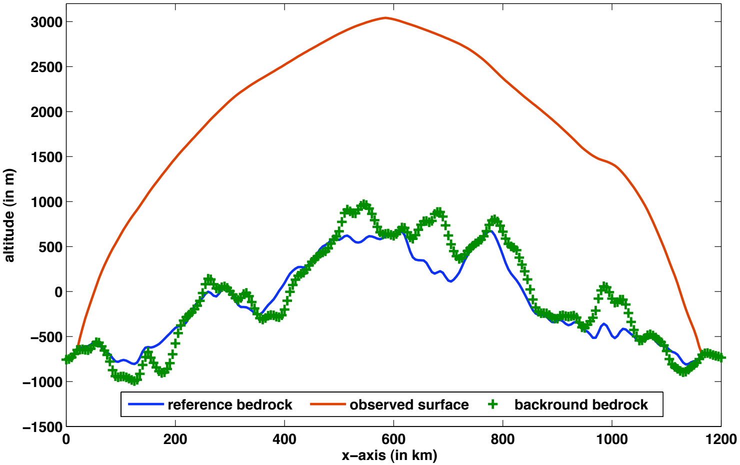
What we observe: each year during 20 years (20 observation times)

- Surface elevation S at each grid point ($\sigma_S = 2\ m$)
- Surface ice velocity u_s at each grid point ($\sigma_u = 3\ m/yr$)
- Bedrock topography B_{soc} at few grid points ($\sigma_B = 20\ m$)

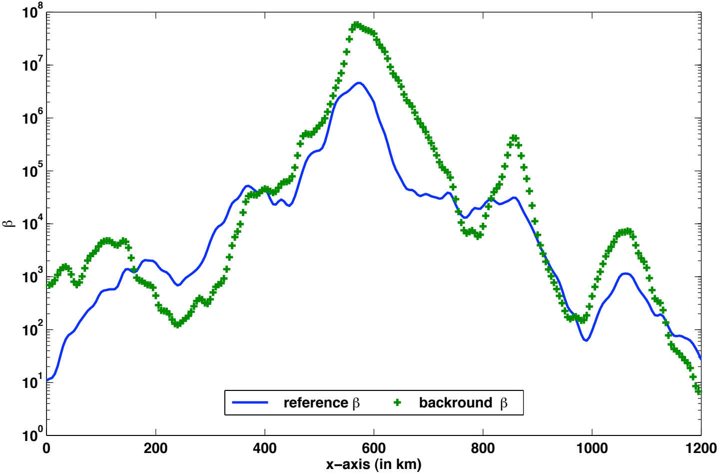
DA system:

- ETKF [Wang et al., 2004], [Ott et al. 2004]
- LETKF [Hunt et al., 2007]

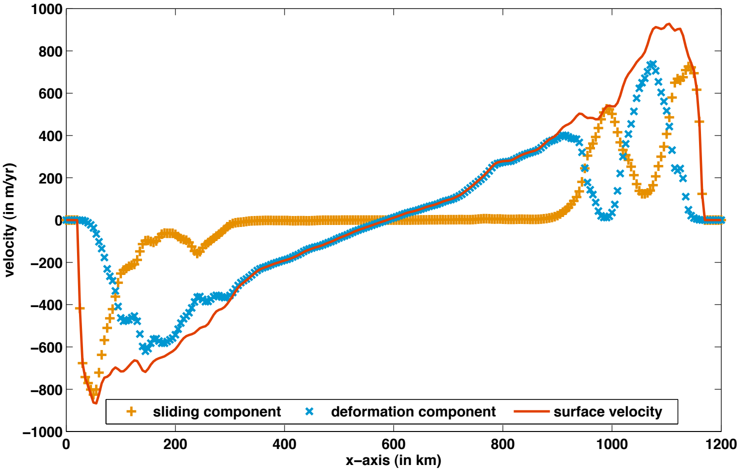
Background and reference states (1)



Background and reference states (2)



Background and reference states (3)



Outline

Basis of Ensemble Kalman Filter (EnKF)

Ensemble formulation:

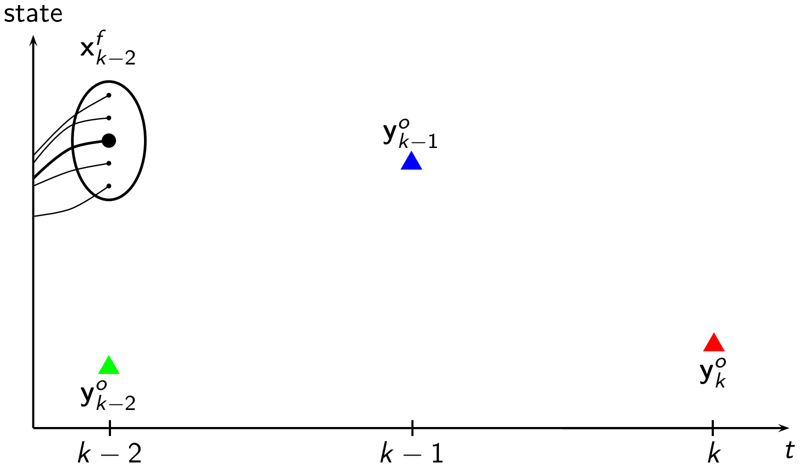
Let $\mathbf{x}_k$ ensemble of N_{ens} state estimations of model at time t_k
(in our case, ice thickness, sliding coefficient and bedrock topography)

- $\mathbf{x}_k^{(i)}$ ith member of ensemble, $i = 1, \ldots, N_{ens}$
- $\overline{\mathbf{x}}_k$ ensemble mean
- $\mathbf{P}_{e,k}$ ensemble covariance matrix

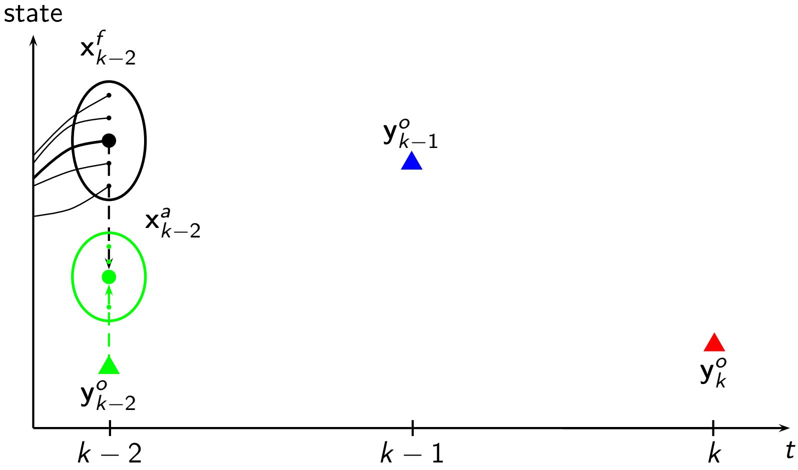
2 step algorithm:

- **Forecast**: Predict $\mathbf{x}_k^f$ from $\mathbf{x}_{k-1}^a$ using the model.
- **Analysis**: Correct $\mathbf{x}_k^f$ with observations $\mathbf{y}_k^o$ (in our case, surface velocities, surface elevation and sparse bedrock topography) to produce $\mathbf{x}_k^a$.

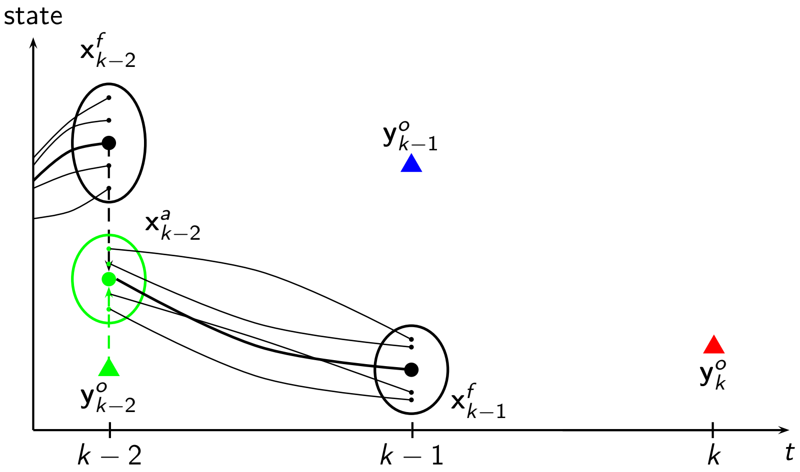
Sequential data assimilation: EnKF sequence



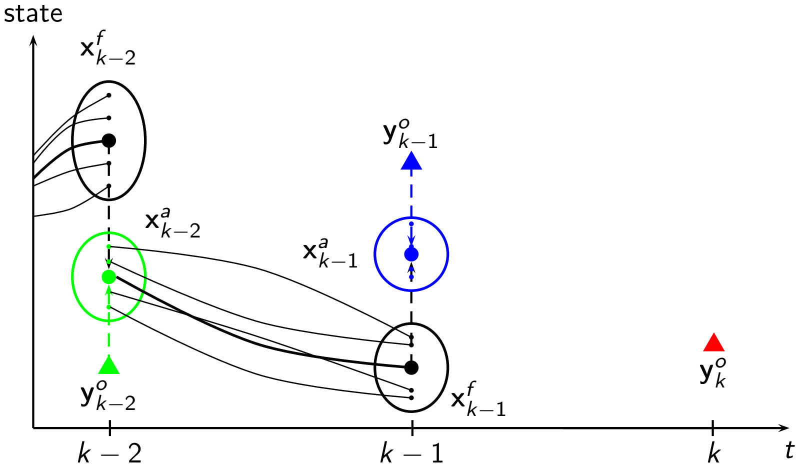
Sequential data assimilation: EnKF sequence



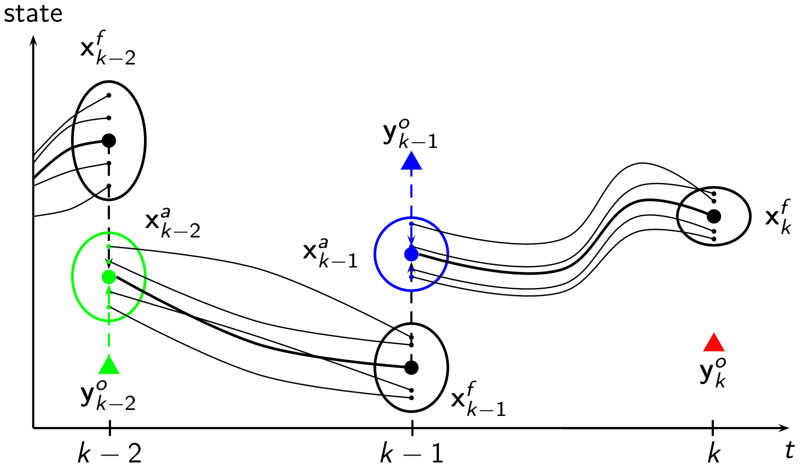
Sequential data assimilation: EnKF sequence



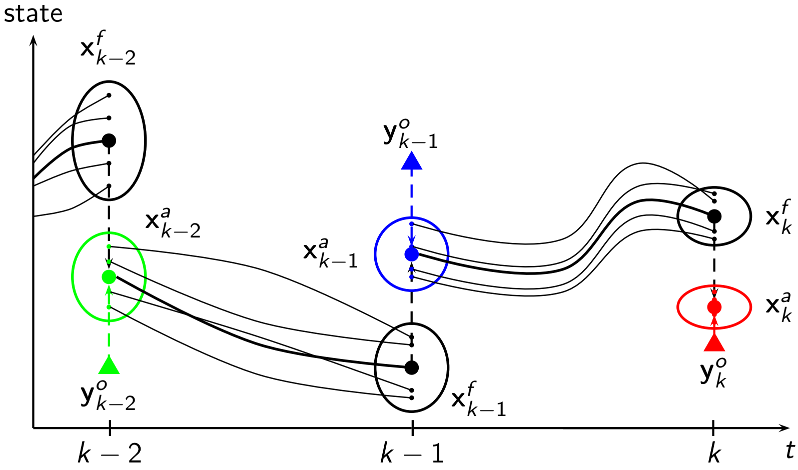
Sequential data assimilation: EnKF sequence



Sequential data assimilation: EnKF sequence



Sequential data assimilation: EnKF sequence



Ensemble Transform Kalman Filter (ETKF)

Detailled in [Hunt et al. 2007], [Harlim and Hunt 2007].

Analysis step

Build an ensemble $\left\{ \mathbf{x}^{a(i)}, i = 1 \dots N_{ens} \right\}$ such as $\overline{\mathbf{x}}^a$ minimum of

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x} - \overline{\mathbf{x}}^f \right)^T \mathbf{P}_e^{f\,-1} \left(\mathbf{x} - \overline{\mathbf{x}}^f \right) + \frac{1}{2} \left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}) \right)^T \mathbf{R}^{-1} \left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}) \right)$$

and $\mathbf{P}_e^a \approx \mathbf{P}^a$.

Attention: $\mathrm{rank}(\mathbf{P}_e^f) \leq N_{ens} - 1$ by construction so P_e^f not invertible.

First hypothesis

Transform Filter

Assume $\mathbf{x} = \overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}$ with

- $\mathbf{w} \in \mathbb{R}^{N_{ens}}$
- $\mathbf{X}^f$ matrix whose ith column is $\mathbf{x}^{f(i)} - \overline{\mathbf{x}}^f$

and minimise $\mathcal{J}(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w})$

Why?

- $\mathbf{w} \in \mathbb{R}^{N_{ens}}$, $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$ and $N_{ens} < n_\mathbf{x}$ (generally).
- $\left(\mathbf{x} - \overline{\mathbf{x}}^f\right)^T {\mathbf{P}_e^f}^{-1} \left(\mathbf{x} - \overline{\mathbf{x}}^f\right) = \mathbf{w}^T {\mathbf{X}^f}^T {\mathbf{P}_e^f}^{-1} \mathbf{X}^f \mathbf{w}$ (background term)

  if $\mathcal{V} = \text{Vect}\left(\left\{\mathbf{x}^{f(i)} - \overline{\mathbf{x}}^f, i = 1 \ldots N_{ens}\right\}\right)$,

  $\mathbf{P}_e^f = \frac{1}{N-1} \mathbf{X}^f {\mathbf{X}^f}^T$ invertible as a restriction to $\mathcal{V}$.

Where are we?

$$\mathcal{J}(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}) = \frac{N-1}{2} \mathbf{w}^T {\mathbf{X}^f}^T \left(\mathbf{X}^f {\mathbf{X}^f}^T \right)^{-1} \mathbf{X}^f \mathbf{w} + \frac{1}{2} \left(\mathbf{y}^o - \mathcal{H}(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}) \right)^T \mathbf{R}^{-1} \left(\mathbf{y}^o - \mathcal{H}(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}) \right)$$

Attention: $\mathcal{J}$ invariant on $\mathrm{Ker}(\mathbf{X}^f)$ and $\dim(\mathrm{Ker}(\mathbf{X}^f)) \geq 1$.
$\Longrightarrow$ non unicity of $\mathbf{w}$ for minimal $\mathcal{J}$.

Gauge-fixing term

Minimise under constraints

To avoid this invariance we can minimise $\mathcal{J}$ over the $\mathbf{w}$ that have a null orthogonal projection on the kernel of $\mathbf{X}^f$

$$\left(\mathbf{I}_{N_{ens}} - \mathbf{X}^{f\,T}\left(\mathbf{X}^f\mathbf{X}^{f\,T}\right)^{-1}\mathbf{X}^f\right)\mathbf{w} = 0$$

It is equivalent to add a gauge-fixing term $\mathcal{G}$ into $\mathcal{J}$

$$\mathcal{G}(\mathbf{w}) = \frac{N-1}{2}\mathbf{w}^T\left(\mathbf{I}_{N_{ens}} - \mathbf{X}^{f\,T}\left(\mathbf{X}^f\mathbf{X}^{f\,T}\right)^{-1}\mathbf{X}^f\right)\mathbf{w}$$

and minimise the new cost function $\widetilde{\mathcal{J}}$.

Where are we?

$$\widetilde{\mathcal{J}}(\mathbf{w}) = \frac{N-1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\left(\mathbf{y}^o - \mathcal{H}(\overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w})\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \mathcal{H}(\overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w})\right)$$

Second hypothesis

Linear approximation for observation operator

$$\mathcal{H}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w}\right) \approx \overline{\mathbf{y}}^f + \mathbf{Y}^f\mathbf{w}$$

with

- $\mathbf{y}^{f(i)} = \mathcal{H}\left(\mathbf{x}^{f(i)}\right)$
- $\overline{\mathbf{y}}^f = \frac{1}{N}\sum_{i=1}^{N_{ens}} \mathbf{y}^{f(i)}$
- matrix $\mathbf{Y}^f$ whose ith column is $\mathbf{y}^{f(i)} - \overline{\mathbf{y}}^f$

Finally, we search to minimise a quadratic cost function

$$\widetilde{\mathcal{J}}^*(\mathbf{w}) = \frac{N-1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right)$$

ETKF analysis step

Direct computation of:

- $\overline{\mathbf{w}}^a = \widetilde{\mathbf{P}}^a \, {\mathbf{Y}^f}^T \mathbf{R}^{-1} \left(\mathbf{y}^o - \overline{\mathbf{y}^f} \right)$ (minimum of $\widetilde{\mathcal{J}}^*$)
- $\widetilde{\mathbf{P}}^a = \left((N_{ens} - 1) \, \mathbf{I}_{N_{ens}} + {\mathbf{Y}^f}^T \mathbf{R}^{-1} \mathbf{Y}^f \right)^{-1}$ $\left(= Hess \left(\widetilde{\mathcal{J}}^*(\overline{\mathbf{w}}^a) \right)^{-1} \right)$
- $\mathbf{W}^a = \left((N_{ens} - 1) \widetilde{\mathbf{P}}^a \right)^{1/2}$ (symetric square root matrix)

Analysis ensemble $\left\{ \mathrm{x}^{a(i)}, i = 1 \ldots N_{ens} \right\}$ is defined as

- $\overline{\mathrm{x}}^a = \overline{\mathrm{x}}^f + \mathbf{X}^f \overline{\mathbf{w}}^a$ (mean)
- $\mathbf{X}^a = \mathbf{X}^f \mathbf{W}^a$ (anomalies matrix)
- $\mathrm{x}^{a(i)} = \overline{\mathrm{x}}^a + \mathbf{X}^a_i$ (with $\mathbf{X}^a_i$ ith column of $\mathbf{X}^a$)

Outline

1. Problem description
2. Large-scale ice sheet model
3. Data assimilation
   - Twin experiment
   - ETKF and LETKF
   - Initial ensemble
   - Results

Initial ensemble: Cooking recipe (1)

- Generate an ensemble of $\alpha = \log(\beta)$ following a Gaussian law $\mathcal{N}(\alpha^b, \mathbf{B}_\alpha)$ with

$$[\mathbf{B}_\alpha]_{ij} = \sigma_\alpha^2 \exp\left(-\frac{(x_i - x_j)^2}{L}\right)$$

- Generate an ensemble of B_{soc} following a Gaussian law $\mathcal{N}(B_{soc}^b, \mathbf{B}_b)$
- Generate an ensemble of S following a Gaussian law $\mathcal{N}(S^{obs}, \sigma_S^2 \mathbf{I})$
- Compute $H = S - B_{soc}$ and correct to avoid negative ice thicknesses.
- For each member, run the model 1 year in order to obtain more physically balanced ice sheets.
- Rescale the produced ensemble so that its mean remains equal to the background.

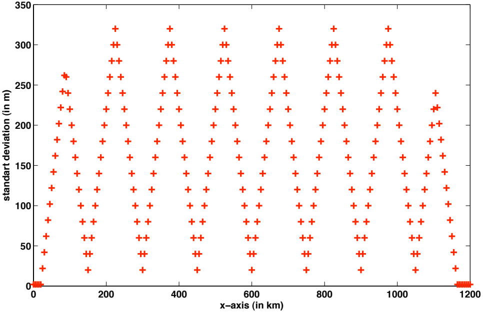
Warning: we use observations to produce initial ensemble!

Initial ensemble: Cooking recipe (2)

$$B_b = \mathbf{\Sigma} \mathbf{C} \mathbf{\Sigma}$$

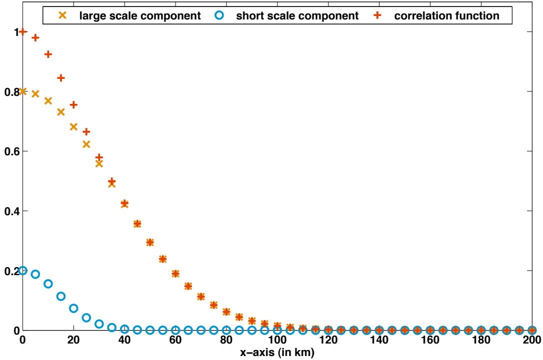
with:

- Σ square root of diagonal matrix of variances
- $\mathbf{C}$ correlation matrix

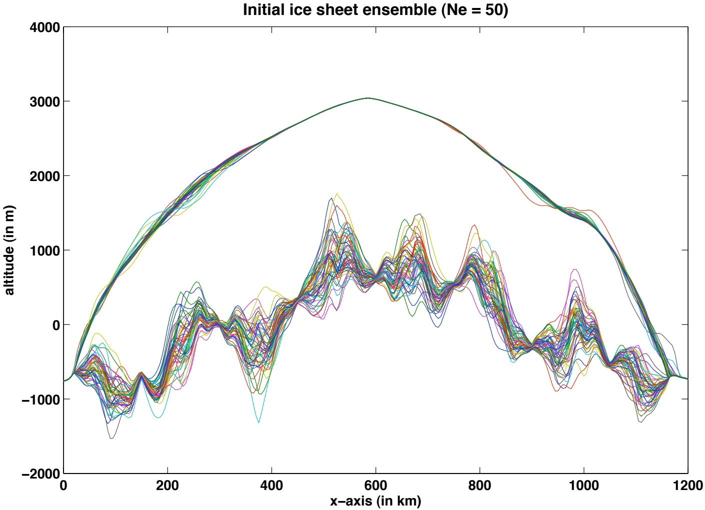


Initial ensemble: Cooking recipe (3)

$$[\mathbf{C}]_{ij} = c_1 \exp\left(-\frac{(x_i - x_j)^2}{L_1}\right) + c_2 \exp\left(-\frac{(x_i - x_j)^2}{L_2}\right)$$

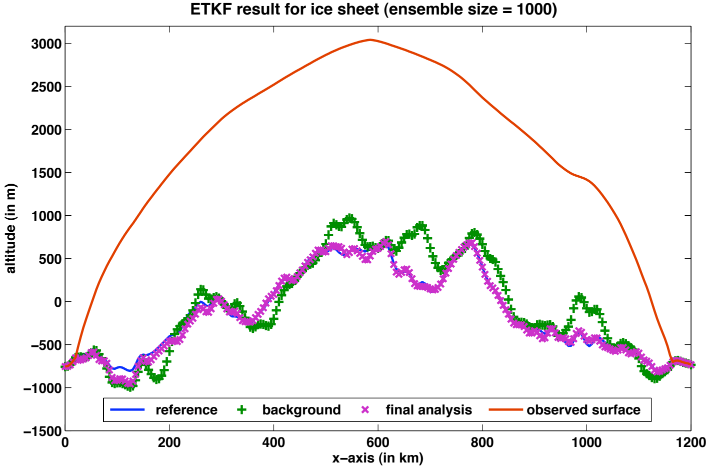


Initial ensemble

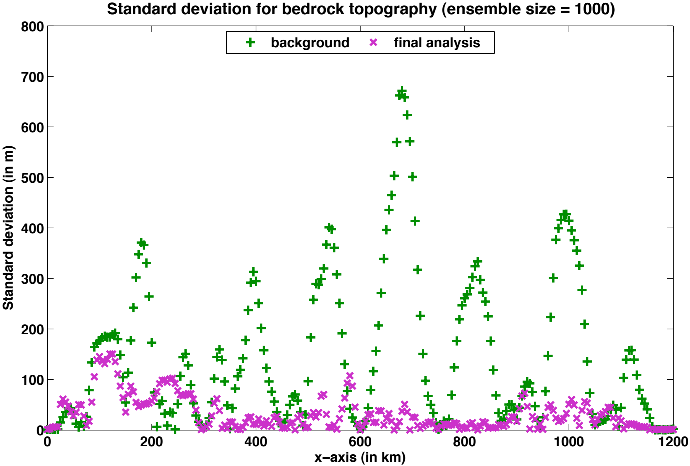


Outline

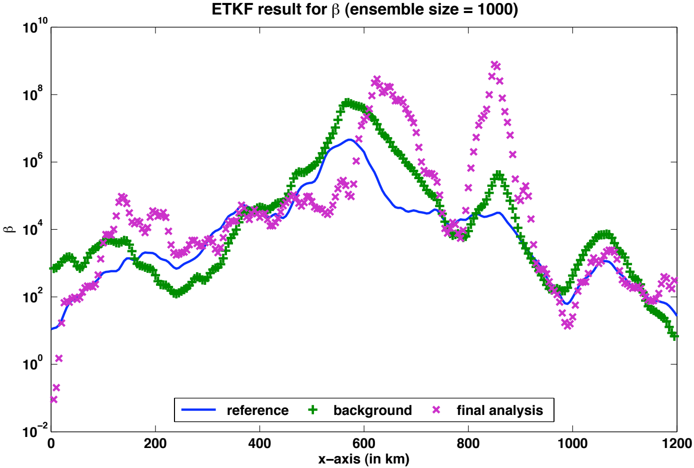
ETKF ($N_{ens} = 1000$)



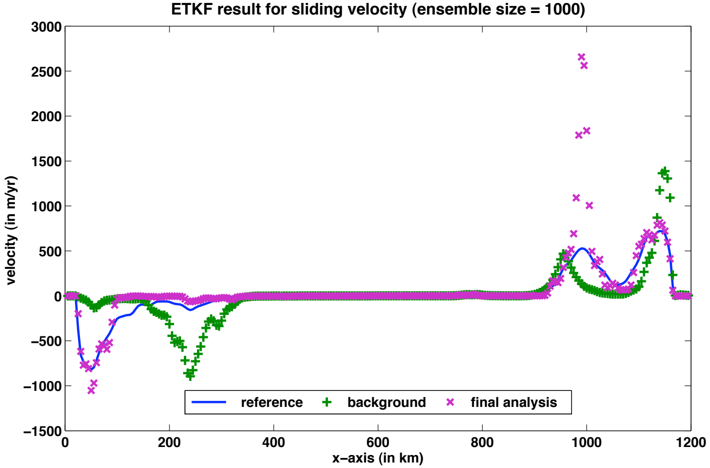
ETKF ($N_{ens} = 1000$)



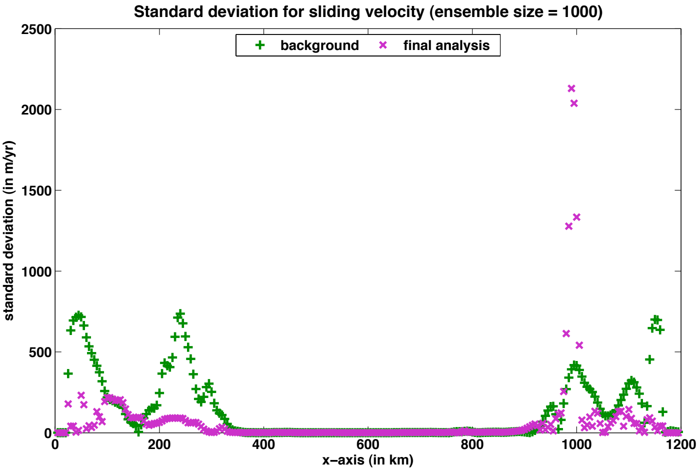
ETKF ($N_{ens} = 1000$)



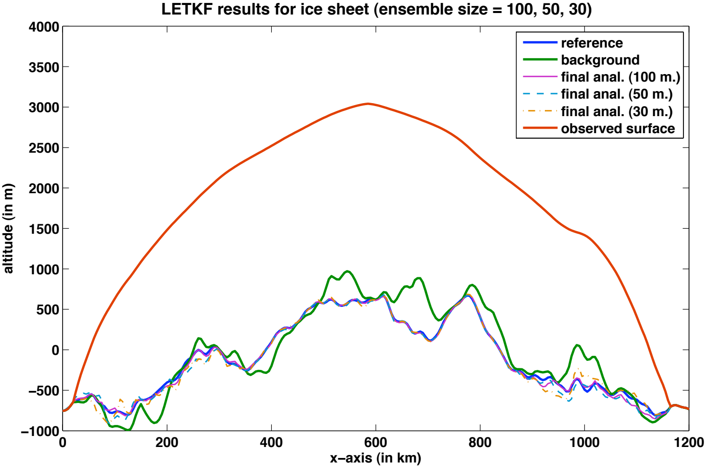
ETKF ($N_{ens} = 1000$)



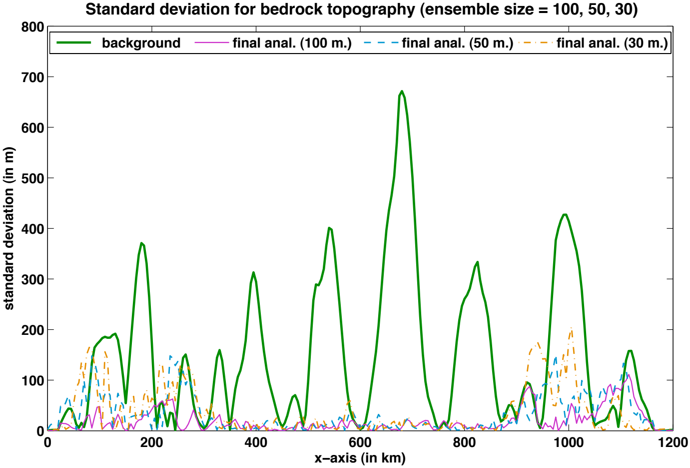
ETKF ($N_{ens} = 1000$)



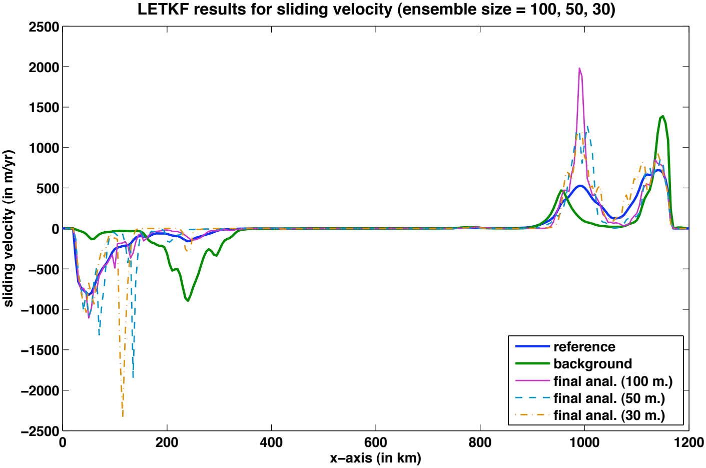
LETKF (N_{ens} = 30, 50, 100)



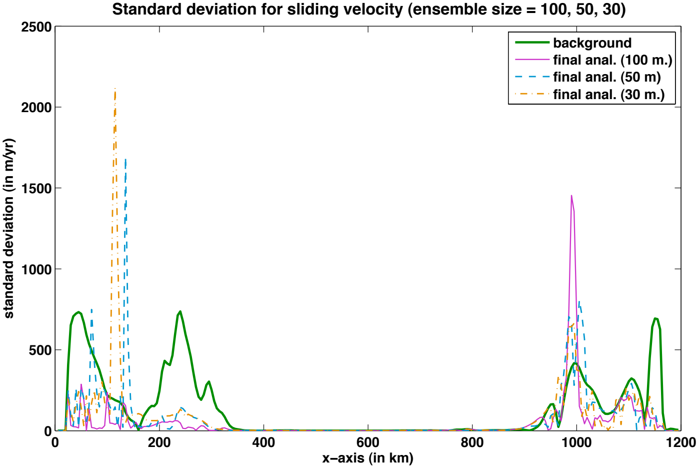
LETKF ($N_{ens} = 30, 50, 100$)



LETKF (N_{ens} = 30, 50, 100)



LETKF (N_{ens} = 30, 50, 100)



Ontgoing and future works

- Compare with variational approaches used in glaciology
- Add more physics: ice shelves, criterion on basal temperature for example.
- Perform experiments on a real flow line
- Implement this method into a 2D ice sheet model.

Thank you for your attention!

Polar bears finally migrate to Antarctica