Ensemble data assimilation using stochastic homogenization in a slow-fast system with tipping points

Lewis Mitchell, Georg Gottwald*

University of Vermont, *University of Sydney

International conference on ensemble methods in geophysical sciences, Toulouse, 12 November 2012

Challenges for EnKF: Small ensemble sizes

Small ensemble sizes $\longrightarrow$ underestimation of $\mathbf{P}_f$

Challenges for EnKF: Small ensemble sizes

Small ensemble sizes $\longrightarrow$ underestimation of $\mathbf{P}_f$

- $\nearrow$ larger ensembles
- $\searrow$ covariance inflation

Challenges for EnKF: Small ensemble sizes

Small ensemble sizes $\longrightarrow$ underestimation of $\mathbf{P}_f$ $\nearrow$ larger ensembles, $\searrow$ covariance inflation

Covariance inflation

$$\mathbf{P}_f \to \delta \mathbf{P}_f$$

where $\quad \delta > 1$

How to choose δ?

Challenges for EnKF: Subgrid-scale phenomena

Large forecast models computationally unwieldy $\rightarrow$ How to parametrize subgrid-scale phenomena?

> Generally cannot observe fast processes $\Rightarrow$ We want to use a reduced model in place of full deterministic model in EnKF

Climate models for data assimilation

Stochastic homogenization (*Khasminsky '66, Kurtz '73, Papanicolaou '76*) has been recently taken up in the context of climate models (works by *Crommelin, Franzke, Majda, Harlim, Timofejev, Vanden-Eijnden*).

IDEA: Consider $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$

$$dx = \frac{1}{\varepsilon} f_0(x,y)\,dt + f_1(x,y)\,dt$$
$$dy = \frac{1}{\varepsilon^2} g_0(x,y)\,dt + \frac{1}{\varepsilon}\sigma(x,y)\,dW_t$$

(For purely deterministic dynamics see *Melbourne and Stuart, Nonlinearity 2011*)

Assume the fast y-process is ergodic, and the average of f_0 over this measure is zero; then the statistics of the slow x-dynamics can be approximated in the limit $\varepsilon \to 0$ by

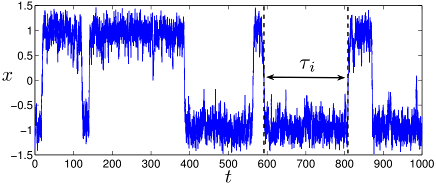
$$dX = F(X)\,dt + \Sigma(X)\,dB_t$$

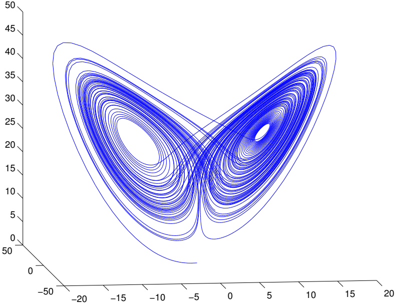
Toy model

4-D slow-fast system (Givon *et al.*, Nonlinearity **17** (2004))

$$\frac{dx}{dt} = x - x^3 + \frac{4}{90\varepsilon} y_2$$

$$\begin{aligned}
\frac{dy_1}{dt} &= \frac{10}{\varepsilon^2}(y_2 - y_1) \\
\frac{dy_2}{dt} &= \frac{1}{\varepsilon^2}(28y_1 - y_2 - y_1 y_3) \\
\frac{dy_3}{dt} &= \frac{1}{\varepsilon^2}\left(y_1 y_2 - \frac{8}{3} y_3\right)
\end{aligned}$$





Toy model - Stochastic model reduction

The time scale separation and ergodicity of the y-dynamics imply the existence of a reduced model, using *homogenization*:

$$\frac{dX}{dt} = X - X^3 + \sigma \frac{dW}{dt}$$

where σ is given by the integrated autocorrelation function of y_2:

$$\frac{\sigma^2}{2} = -\left(\frac{4}{90}\right)^2 \int_0^\infty y_2(t) \lim_{T\to\infty} \frac{1}{T} \int_0^T y_2(t+s) ds dt.$$

Numerically, $\sigma^2 \approx 0.113$.

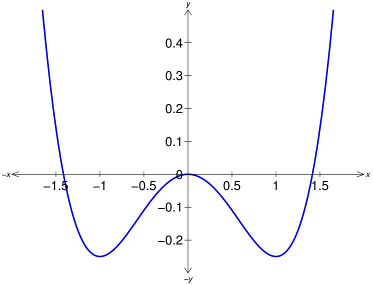
Toy model

Can be modelled as a particle in potential well:

$$V(x) = \frac{x^4}{4} - \frac{x^2}{2}$$

Then

$$\frac{dX}{dt} = -V'(x) + \sigma \frac{dW}{dt}$$



Time scales

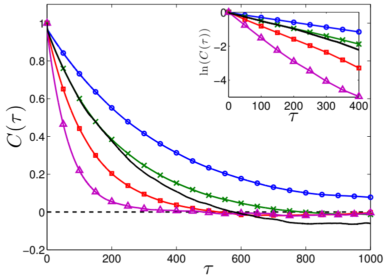
Transit time τ_t (Kramers, 1940)

Time taken to travel between metastable states $x^* = \pm 1$ is $\tau_t = 5.90$.

Exit time τ_e (Kramers, 1940)

Average residence time in one well $\tau_e = 75.6769$ (for $\sigma^2 = 0.126$).

Decorrelation time τ_{corr}



$\sigma^2 = 0.1$

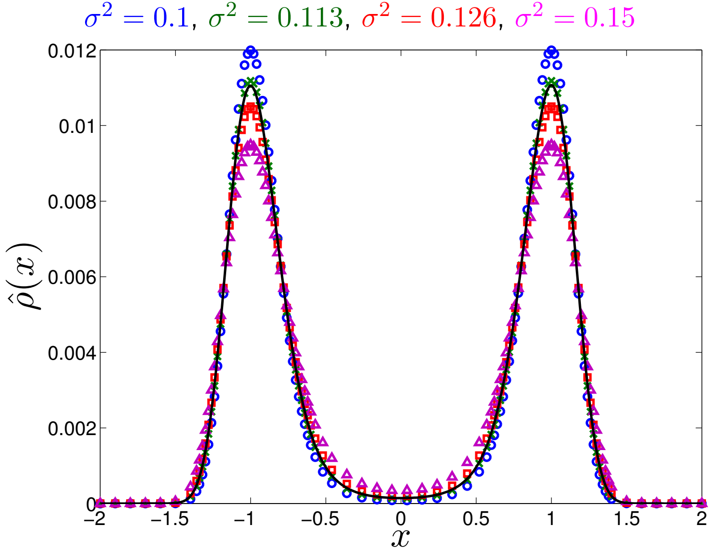
$\sigma^2 = 0.113$

$\sigma^2 = 0.126$

$\sigma^2 = 0.15$

Decorrelation time for full deterministic model is $\tau_{\text{corr}} = 233.0$

Sensitivity to σ^2



So $\sigma^2 = 0.113$ produces the best “climate”.

Data assimilation experiment

Twin experiment

1. Generate "truth" from time series of full deterministic model
2. Create "observations" of slow x-variable by adding Gaussian noise
3. Attempt to recover truth using
   - Full deterministic 4D forecast model
   - Reduced stochastic 1D model

Data assimilation experiment

Twin experiment

1. Generate "truth" from time series of full deterministic model
2. Create "observations" of slow x-variable by adding Gaussian noise
3. Attempt to recover truth using
   - Full deterministic 4D forecast model
   - Reduced stochastic 1D model

"BIG" QUESTION: Can reduced stochastic climate models be beneficial for forecasting and prediction?

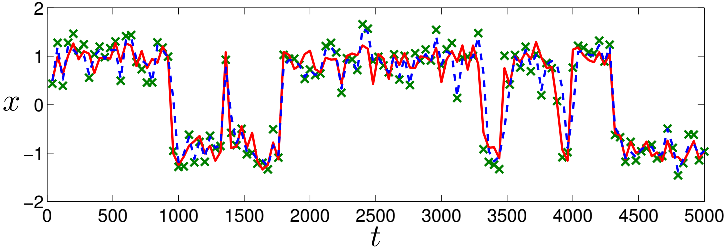
Under what circumstances and why can stochastic reduced models be beneficial as forecast models in an ensemble Kalman filter setting?

Can we achieve

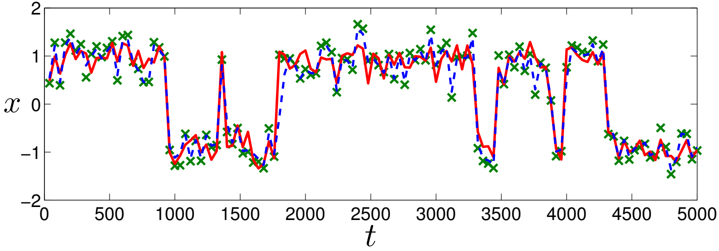
- computational gain?
- better analysis skill?

Numerical results

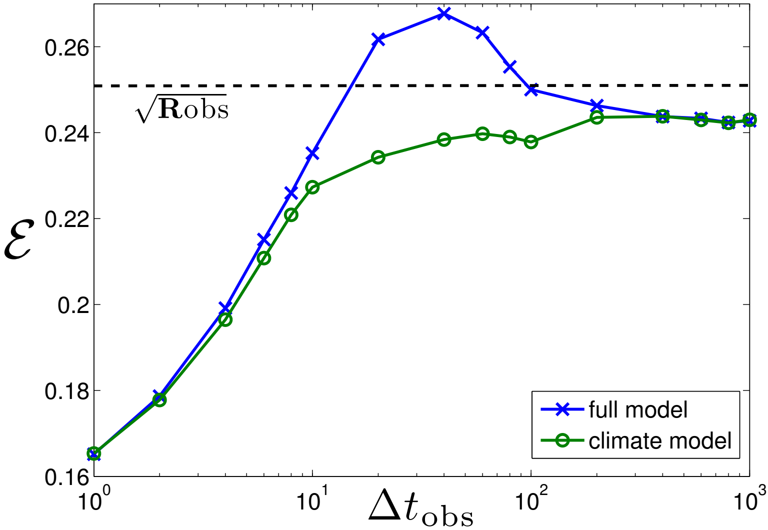
Full deterministic model:



Reduced climate model:



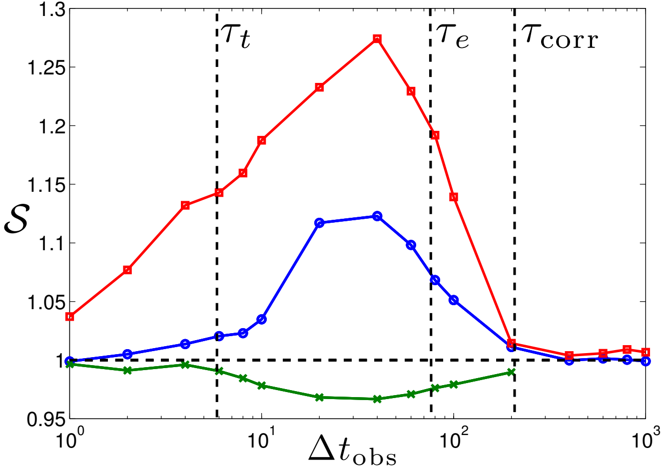
Numerical results



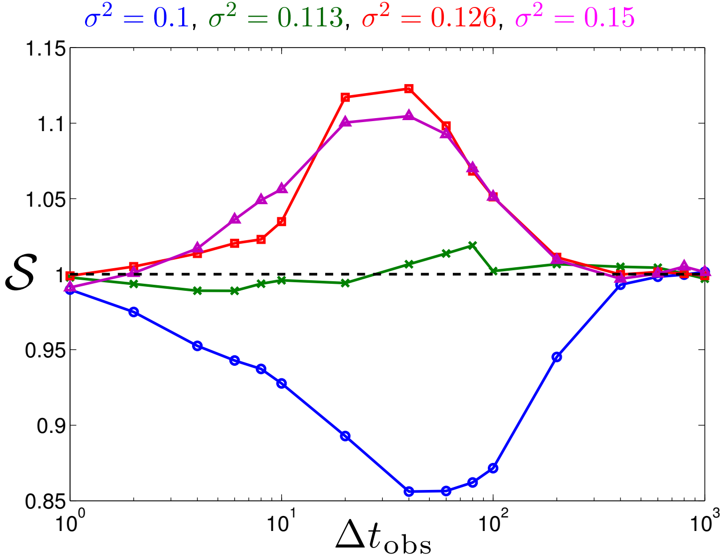
Numerical results

We define the *skill* $\mathcal{S} = \dfrac{\mathcal{E}_{\text{full}}}{\mathcal{E}_{\text{climate}}}$, $\mathcal{S} > 1$ is good!

Blue - all analyses, Green - metastable states, Red - transitions



Numerical results



Numerical results

The numerical results suggest that stochastic climate models are

- beneficial for observation time intervals $\Delta t_{\mathrm{obs}} \in (\tau_t, \tau_{\mathrm{corr}})$
- good at capturing the transitions between slow metastable states
- perform better than full system for diffusion larger than the "correct" value: $\sigma^2 > 0.113$

Numerical results

The numerical results suggest that stochastic climate models are

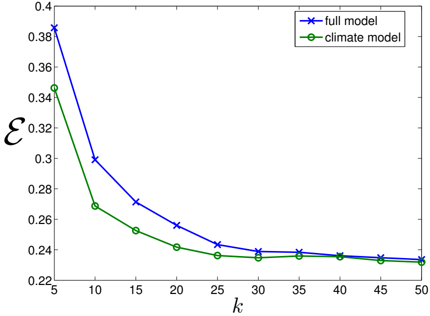
- beneficial for observation time intervals $\Delta t_{\mathrm{obs}} \in (\tau_t, \tau_{\mathrm{corr}})$
- good at capturing the transitions between slow metastable states
- perform better than full system for diffusion larger than the "correct" value: $\sigma^2 > 0.113$

So, if the climate model fails to accurately reproduce the statistics of the full model, why does it perform better here?

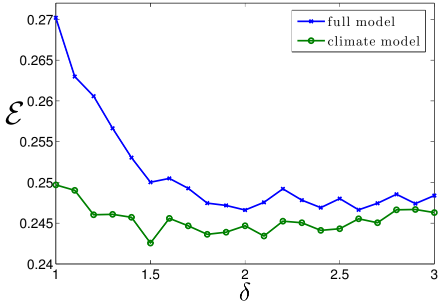
Numerical results

Small ensemble sizes $\longrightarrow$ underestimation of $\mathbf{P}_f$ $\nearrow$ larger ensembles $\searrow$ covariance inflation

Increasing ensemble size k:



Increasing covariance inflation δ:



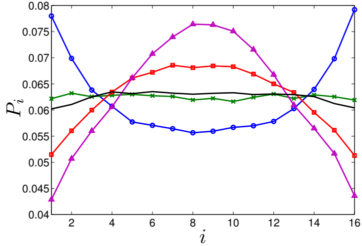
Ranked probability histograms

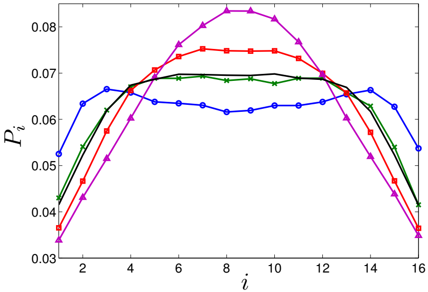
- sort the forecast ensemble $\mathbf{X}_f = [x_{f,1}, x_{f,2}, ..., x_{f,k}]$ and create bins $(-\infty, x_{f,1}]$, $(x_{f,1}, x_{f,2}]$, ... , $(x_{f,k}, \infty)$ at each forecast step
- increment whichever bin the actual truth falls into at each forecast step

Convex histogram: insufficient ensemble spread
Concave histogram: too much ensemble spread
Flat histogram: reliable ensemble for which each ensemble member has equal probability of being nearest to the truth

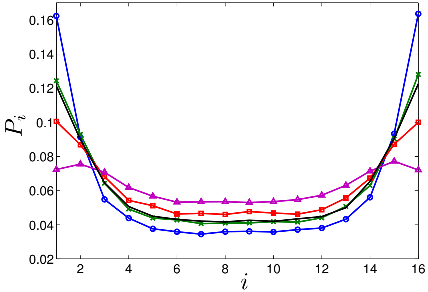
Ensemble spread for different forecast models

All variables:





Metastable states



Transitions

Conclusions and further work

Take-home point:

In the context of DA, models which produce best “climate” need not produce best analyses (and vice versa).

Conclusions and further work

Take-home point:

In the context of DA, models which produce best "climate" need not produce best analyses (and vice versa).

We can incorporate climatological information into the forecast step through stochastic parametrization of fast dynamics in the forecast model. This can produce more skilful analyses, because it

- creates ensembles with greater spread
- simulates larger ensembles (or covariance inflation)
- increases ensemble reliability

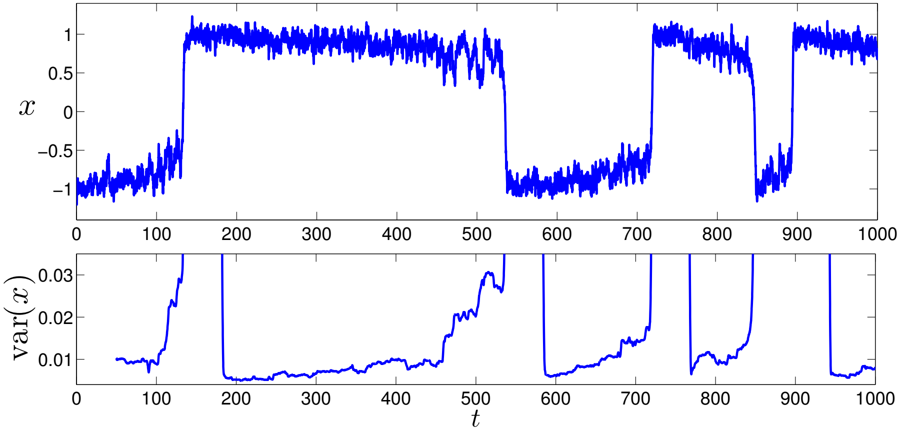
We will:

- Attack this problem of producing analyses that are consistent with climate statistics more directly
- Extend to the case with bifurcation-induced tipping

Bifurcation-induced tipping

Induce tipping by time-varying bifurcation parameter:

$$\frac{dx}{dt} = x - x^3 + \mu + \frac{\kappa}{\varepsilon} y_2$$



Homogenization

Consider the set of ODEs

$$\frac{dx}{dt} = \frac{1}{\epsilon} f_0(x, y) + f_1(x, y)$$
$$\frac{dy}{dt} = \frac{1}{\epsilon^2} g(x, y).$$

The backward Kolmogorov equation for $v(z, t) = \mathbb{E}(\phi(z(t))|z(0) = z_0)$, $z = (x, y)$ is

$$\frac{\partial v}{\partial t} = \frac{1}{\epsilon^2} \mathcal{L}_0 v + \frac{1}{\epsilon} \mathcal{L}_1 v + \mathcal{L}_2 v$$

where

$$\mathcal{L}_0 = g \cdot \nabla_y, \quad \mathcal{L}_1 = f_0 \cdot \nabla_x, \quad \mathcal{L}_2 = f_1 \cdot \nabla_x$$

Homogenization

Substituting an ansatz (Papanicolou (1976)) $v = v_0 + \epsilon v_1 + \epsilon^2 v_2 + \ldots$ and equating orders gives

$$O\left(\frac{1}{\epsilon^2}\right) : \mathcal{L}_0 v_0 = 0$$

$$O\left(\frac{1}{\epsilon}\right) : \mathcal{L}_0 v_1 = -\mathcal{L}_1 v_0$$

$$O(1) : \mathcal{L}_0 v_2 = \frac{\partial v_0}{\partial t} - \mathcal{L}_1 v_1 - \mathcal{L}_2 v_0$$

Homogenization

Substituting an ansatz (Papanicolou (1976)) $v = v_0 + \epsilon v_1 + \epsilon^2 v_2 + ...$ and equating orders gives

$$O\left(\frac{1}{\epsilon^2}\right) : \mathcal{L}_0 v_0 = 0 \qquad \Rightarrow v_0 = v_0(x,t)$$

$$O\left(\frac{1}{\epsilon}\right) : \mathcal{L}_0 v_1 = -\mathcal{L}_1 v_0$$

$$O(1) : \mathcal{L}_0 v_2 = \frac{\partial v_0}{\partial t} - \mathcal{L}_1 v_1 - \mathcal{L}_2 v_0$$

Homogenization

Substituting an ansatz (Papanicolou (1976)) $v = v_0 + \epsilon v_1 + \epsilon^2 v_2 + \ldots$ and equating orders gives

$$O\left(\frac{1}{\epsilon^2}\right) : \mathcal{L}_0 v_0 = 0 \qquad \Rightarrow v_0 = v_0(x,t)$$

$$O\left(\frac{1}{\epsilon}\right) : \mathcal{L}_0 v_1 = -\mathcal{L}_1 v_0 \qquad \Rightarrow \langle \mathcal{L}_1 v_0 \rangle_{\rho_\infty} = 0$$

$$\Rightarrow v_1 = -\mathcal{L}_0^{-1} \mathcal{L}_1 v_0$$

$$O(1) : \mathcal{L}_0 v_2 = \frac{\partial v_0}{\partial t} - \mathcal{L}_1 v_1 - \mathcal{L}_2 v_0$$

Homogenization

Substituting an ansatz (Papanicolou (1976)) $v = v_0 + \epsilon v_1 + \epsilon^2 v_2 + \ldots$ and equating orders gives

$$O\left(\frac{1}{\epsilon^2}\right) : \mathcal{L}_0 v_0 = 0 \qquad \Rightarrow v_0 = v_0(x,t)$$

$$O\left(\frac{1}{\epsilon}\right) : \mathcal{L}_0 v_1 = -\mathcal{L}_1 v_0 \qquad \Rightarrow \langle \mathcal{L}_1 v_0 \rangle_{\rho_\infty} = 0$$

$$\Rightarrow v_1 = -\mathcal{L}_0^{-1} \mathcal{L}_1 v_0$$

$$O(1) : \mathcal{L}_0 v_2 = \frac{\partial v_0}{\partial t} - \mathcal{L}_1 v_1 - \mathcal{L}_2 v_0$$

and the solvability condition for the $O(1)$ expression is

$$\begin{aligned} \frac{\partial v_0}{\partial t} &= \int \mathcal{L}_2 v_0 \rho_\infty(y) dy - \int \mathcal{L}_1 \mathcal{L}_0^{-1} \mathcal{L}_1 v_0 \rho_\infty(y) dy \\ &= x(1-x^2) \frac{\partial}{\partial x} v_0 - \left(\frac{4}{90}\right)^2 \left\langle y_2 \mathcal{L}_0^{-1} y_2 \right\rangle_{\rho_\infty} \frac{\partial^2}{\partial x^2} v_0 \end{aligned}$$

Homogenization

Define, for every function h with $\langle h \rangle_{\rho_\infty} = 0$,

$$H(x,y) = -\int_0^\infty e^{\mathcal{L}_0 t} h \, dt \; .$$

As $\mathcal{L}_0$ corresponds to an ergodic process, $\mathcal{L}_0 H = h$. Hence

$$\begin{aligned}
\langle y_2 \mathcal{L}_0^{-1} y_2 \rangle_{\rho_\infty} &= -\int_0^\infty \langle y_2 e^{\mathcal{L}_0 t} y_2 \rangle_{\rho_\infty} \, dt \\
&= -\int_0^\infty \{ \lim_{T\to\infty} \frac{1}{T} \int_0^T y_2(s) y_2(t+s) ds \} \, dt = \frac{\sigma^2}{2}.
\end{aligned}$$

Backward Kolmogorov equation is

$$\frac{\partial}{\partial t} v_0 = x(1-x^2) \frac{\partial}{\partial x} v_0 + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} v_0.$$

Exit time

Let D be the "well" domain for the particle. Then we wish to find the sojourn time $\tau_D^x = \inf\{t \geq 0 : X \notin D\}$. Then the *exit time* τ_e is the average sojourn time

$$\tau_e = \mathbb{E}(\tau_D^x).$$

Theorem

The exit time τ_e is the solution of the boundary value problem

$$\mathcal{L}\tau = -1 \qquad x \in D$$
$$\tau = 0 \qquad x \in \partial D$$

where $\mathcal{L} = -V'(x)\frac{d}{dx} + \frac{1}{2}\sigma^2 \frac{d^2}{dx^2}$ is the generator of the reduced model.

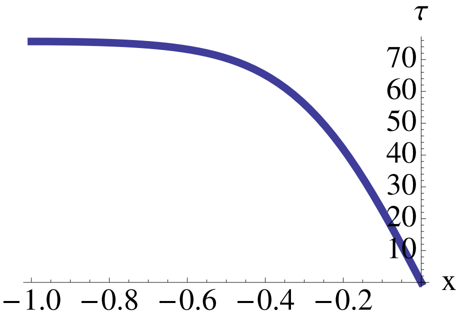
(Gardiner, 2003)

Exit time

We can find the average **exit time** for the particle from a well

$$\tau_e(x) = \frac{2}{\sigma^2}\int_x^0 \exp\left(\frac{2V(y)}{\sigma^2}\right)\int_{-1}^{y} \exp\left(\frac{-2V(z)}{\sigma^2}\right) dz dy$$

Gives exit time $\tau_e = 75.6769$
(for $\sigma^2 = 0.126$)



The Ensemble Kalman Filter (EnKF)

Step 3: Update of the ensemble

The ensemble needs to be consistent with

$$\mathbf{P}_a = \frac{1}{k-1}\mathbf{Z}'_a \left[\mathbf{Z}'_a\right]^T$$

Method of ensemble square root filters:

- Ensemble transform Kalman filter (ETKF) (*Tippett et al 2003*): $\mathbf{Z}'_a = \mathbf{Z}'_f\mathbf{S}$ with $\mathbf{S} \in \mathbb{R}^{k\times k}$
- Ensemble adjustment Kalman filter (EAKF) (*Anderson 2001*): $\mathbf{Z}'_a = \mathbf{A}\mathbf{Z}'_f$ with $\mathbf{A} \in \mathbb{R}^{N\times N}$

The Ensemble Kalman Filter (EnKF)

Step 3: Update of the ensemble

The ensemble needs to be consistent with

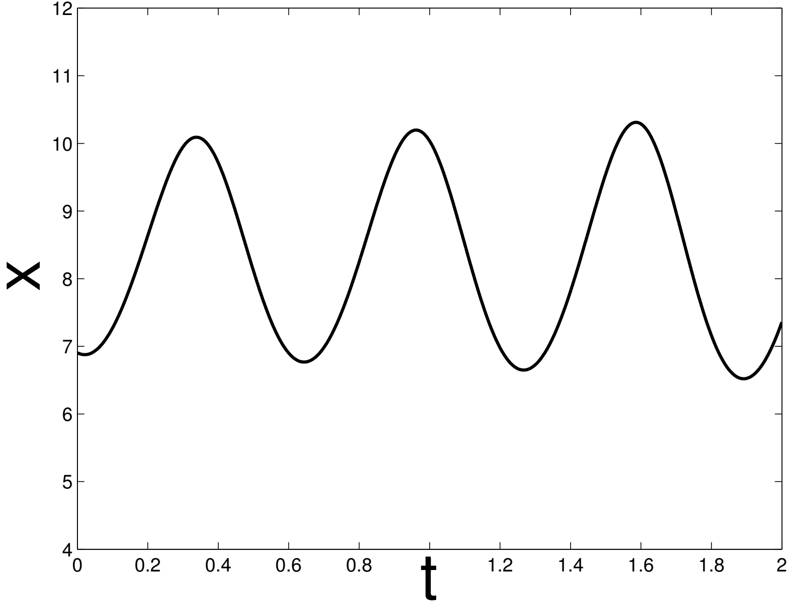
$$\mathbf{P}_a = \frac{1}{k-1}\mathbf{Z}'_a\left[\mathbf{Z}'_a\right]^T$$

Method of ensemble square root filters:

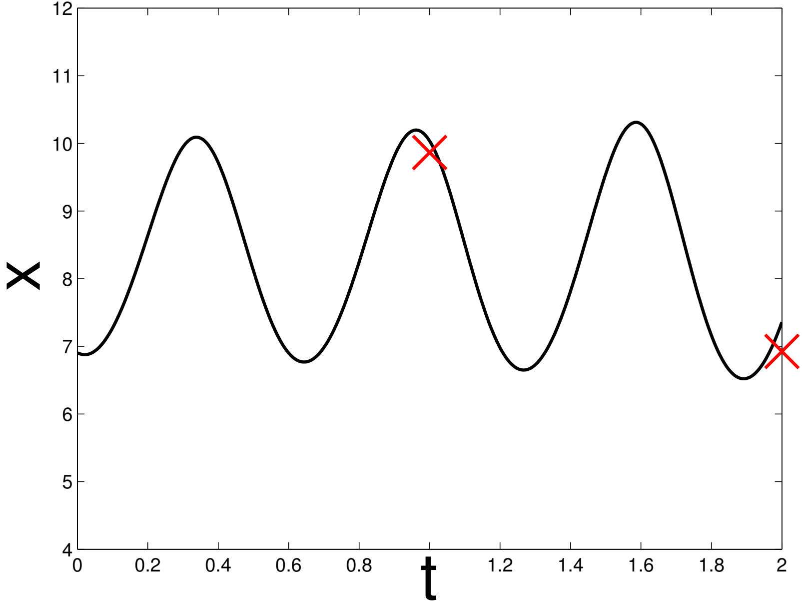
- Ensemble transform Kalman filter (ETKF) (*Tippett et al 2003*):
  $\mathbf{Z}'_a = \mathbf{Z}'_f\mathbf{S}$ with $\mathbf{S} \in \mathbb{R}^{k\times k}$
- Ensemble adjustment Kalman filter (EAKF) (*Anderson 2001*):
  $\mathbf{Z}'_a = \mathbf{A}\mathbf{Z}'_f$ with $\mathbf{A} \in \mathbb{R}^{N\times N}$

Step 4: Update of the forecast

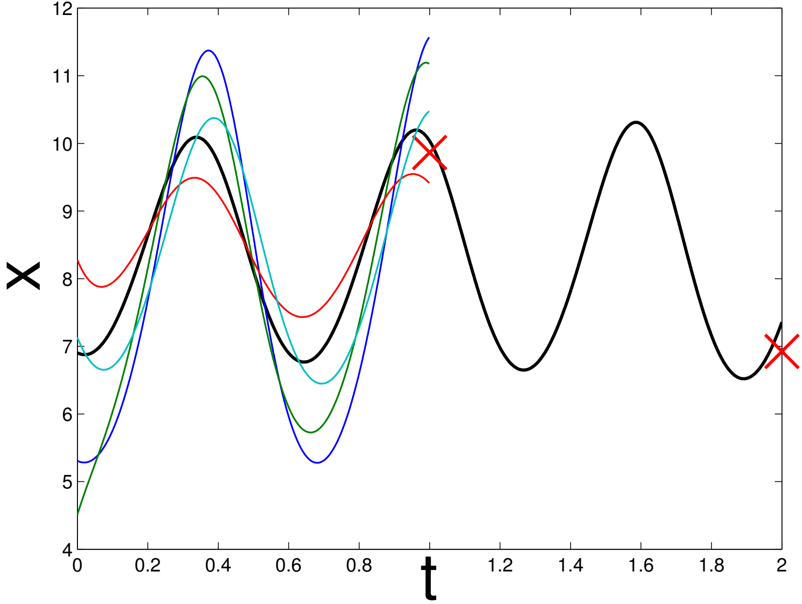
Set $\mathbf{Z}_b = \mathbf{Z}_a$ to propagate the ensemble forward again with the full dynamics to the next observation time



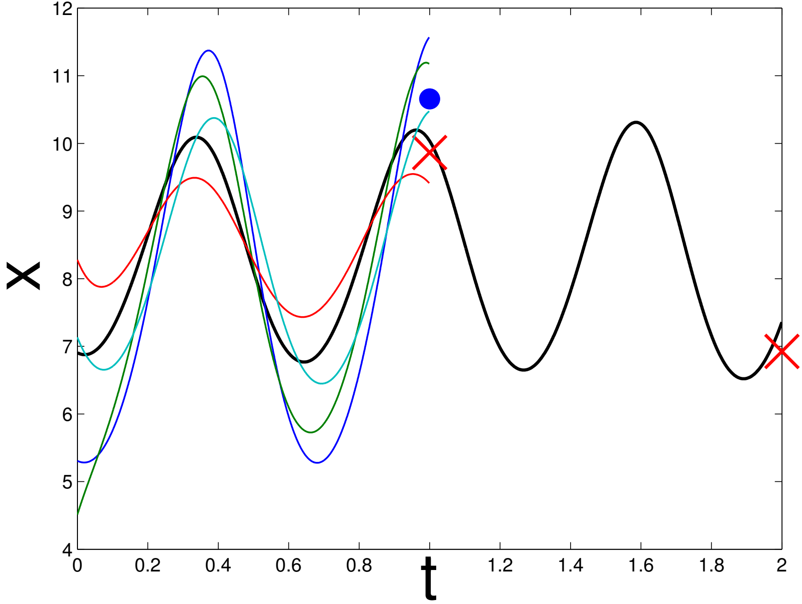
12
11
10
9
8
7
6
5
4
x
0
0.2
0.4
0.6
0.8
t
1.2
1.4
1.6
1.8
2



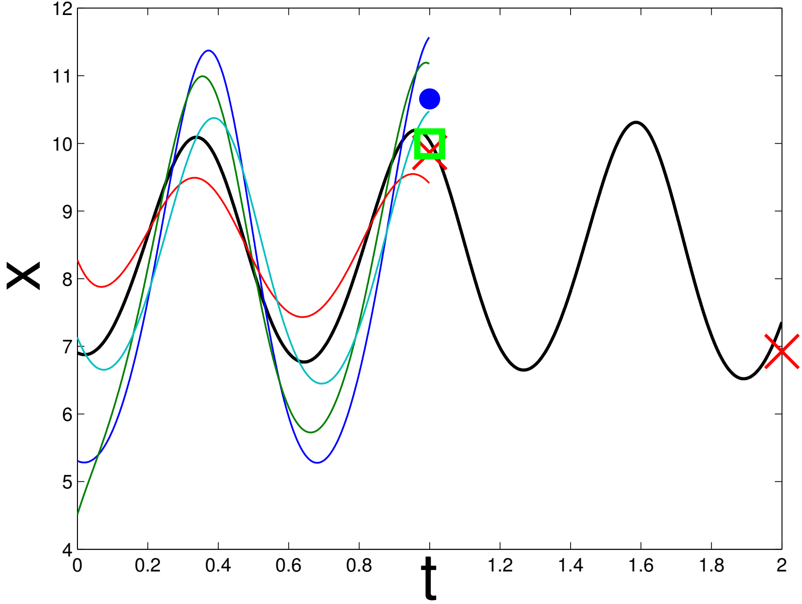
12
11
10
9
8
7
6
5
4
x
0
0.2
0.4
0.6
0.8
t
1.2
1.4
1.6
1.8
2



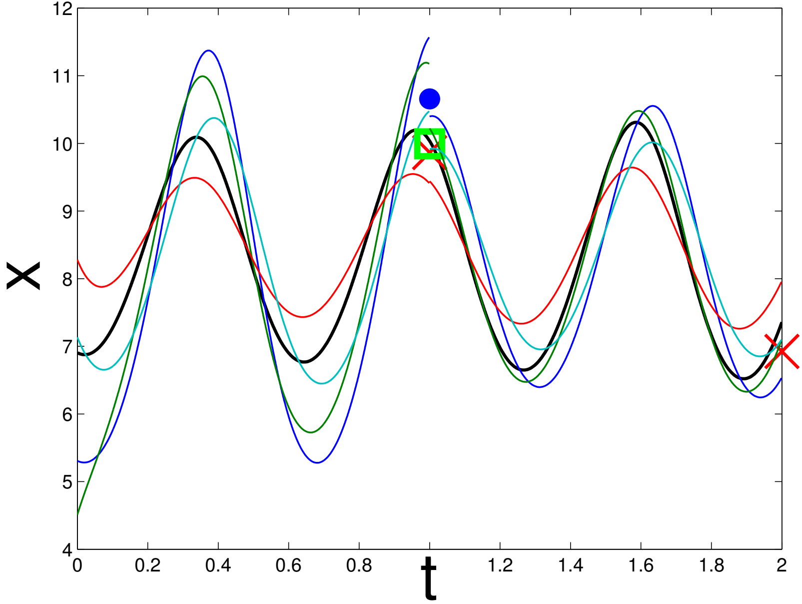
12
11
10
9
8
7
6
5
4
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
x
t



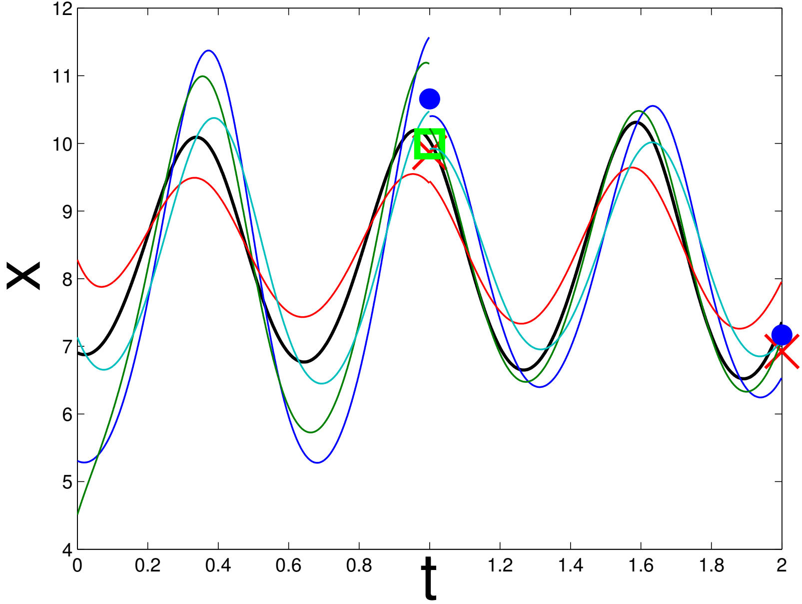
12
11
10
9
8
7
6
5
4
x
0
0.2
0.4
0.6
0.8
1.2
1.4
1.6
1.8
2
t



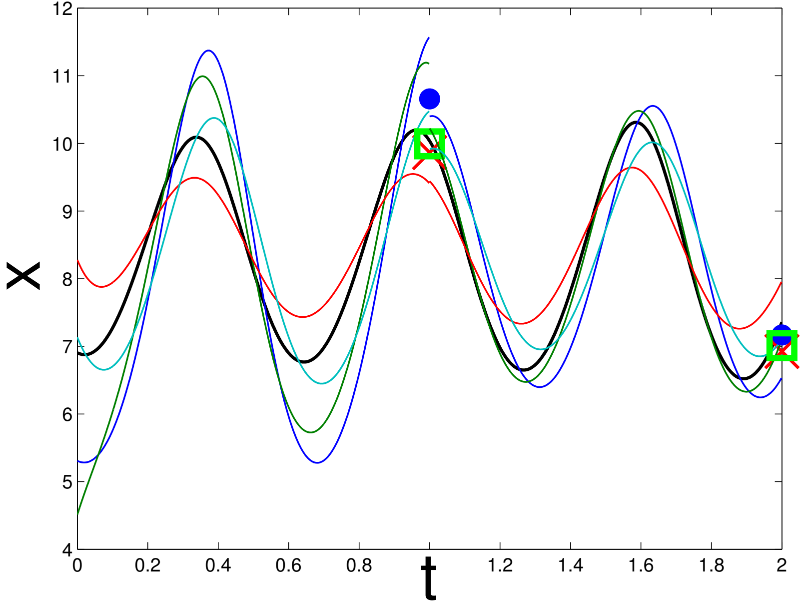
12
11
10
9
8
7
6
5
4
x
0
0.2
0.4
0.6
0.8
1.2
1.4
1.6
1.8
2
t



12
11
10
9
8
7
6
5
4
x
0
0.2
0.4
0.6
0.8
1.2
1.4
1.6
1.8
2
t



x
t
0
0.2
0.4
0.6
0.8
1.2
1.4
1.6
1.8
2
4
5
6
7
8
9
10
11
12

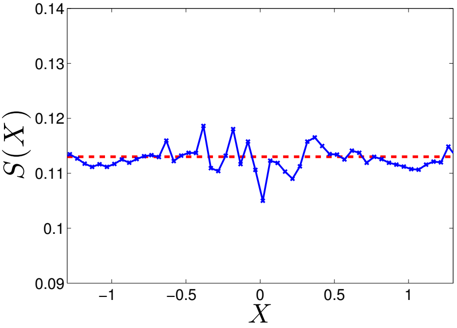


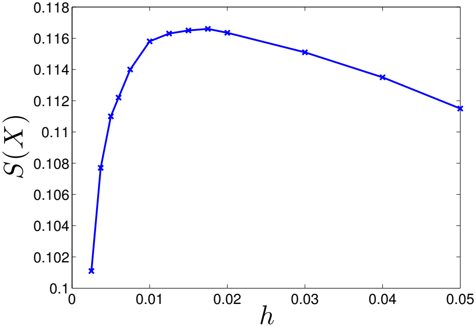
12
11
10
9
8
7
6
5
4
x
0
0.2
0.4
0.6
0.8
t
1.2
1.4
1.6
1.8
2

Parameter estimation

Estimate σ^2 from a long trajectory partitioned into bins $[X, X + \Delta X]$, sampled at coarse sampling time $h \gg dt$

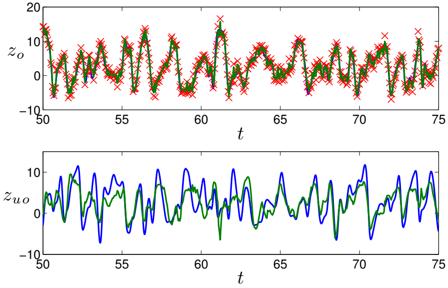
$$\sigma^2 \approx S(X) = \frac{1}{h} \langle (x^{n+1} - x^n)^2 \rangle \Big|_{x^n \in (X, X+\Delta X)}.$$



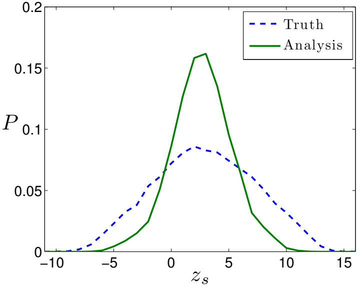


Data assimilation: Weather vs. Climate

Weather:



Climate:



A challenge for data assimilation in climate problems

- Weather models capture short-term dynamics well, but not long-term.
- Climate modelers interested in distribution of states averaged over long time scales (capturing the pdf)
- Uncertainty in how to find balance between short-term (window of predictability) and long-term (climate, pdf) models.

A challenge for data assimilation in climate problems

- Weather models capture short-term dynamics well, but not long-term.
- Climate modelers interested in distribution of states averaged over long time scales (capturing the pdf)
- Uncertainty in how to find balance between short-term (window of predictability) and long-term (climate, pdf) models.

We will study a situation where it is preferable to use a “worse” model

Summary of Ensemble Kalman Filtering (EnKF)

Step 1: Forecast step

$$\mathbf{Z}_f = F(\mathbf{Z}_b)$$

$$\mathbf{P}_f = \frac{1}{k-1}\mathbf{Z}'_f(t)[\mathbf{Z}'_f(t)]^T$$

Step 2: Analysis step

$$\mathbf{z}_a = \mathbf{z}_f + \mathbf{K}_{\text{obs}}(\mathbf{y} - \mathbf{H}\mathbf{z}_f)$$

$$\mathbf{K}_{\text{obs}} = \mathbf{P}_a\mathbf{H}^T\mathbf{R}_{\text{obs}}^{-1}\,, \quad \mathbf{P}_a = \left(\mathbf{P}_f^{-1} + \mathbf{H}^T\mathbf{R}_{\text{obs}}^{-1}\mathbf{H}\right)^{-1}$$

Step 3: Updating the ensemble

The ensemble needs to be consistent with

$$\mathbf{P}_a = \frac{1}{k-1}\mathbf{Z}'_a\left[\mathbf{Z}'_a\right]^T$$

by means of a transformation $\mathbf{Z}'_a = \mathbf{Z}'_f\mathbf{S}$ (we will use the ETKF)

Step 4: Update of the forecast

Set $\mathbf{Z}_b = \mathbf{Z}_a$, propagate the ensemble forward again with the forecast model.