Filtering Turbulent Systems with Stochastic Parameterized "Extended" Kalman Filter

John Harlim

Department of Mathematics
North Carolina State University

May 27, 2010

References:

a) Test Models for Improving Filtering with Model Errors through Stochastic Parameter Estimation (with B. Gershgorin and A.J. Majda), J. Comput. Phys., 229(1):1-31, 2010.

b) Improving Filtering and Prediction of Spatially Extended Turbulent Systems with Model Errors through Stochastic Parameter Estimation (with B. Gershgorin and A.J. Majda), J. Comput. Phys., 229(1):32-57, 2010.

c) Filtering Turbulent Sparsely Observed Geophysical Flows (with A.J. Majda), Monthly Weather Review, 138(4): 1050-1083, 2010.

Review Article: Mathematical Strategies for Filtering Turbulent Dynamical Systems (with B. Gershgorin and A.J. Majda), DCDS-A: 27(2), 441-486, 2010.

Ch 13 of: Systematic Strategies for Real Time Filtering of Turbulent Signals in Complex Systems (with A.J. Majda), Cambridge University Press (in preparation), 2010.

Online Model Error Estimation Strategy

A simple strategy to cope with model errors for filtering with an imperfect model nonlinear dynamical system depending on parameters, b,

$$\frac{du}{dt} = F(u, b)$$

is to augment the state variable u, by the parameters λ, and adjoin an approximate dynamical equation for the parameters

$$\frac{db}{dt} = g(b).$$

Online Model Error Estimation Strategy

A simple strategy to cope with model errors for filtering with an imperfect model nonlinear dynamical system depending on parameters, b,

$$\frac{du}{dt} = F(u, b)$$

is to augment the state variable u, by the parameters λ, and adjoin an approximate dynamical equation for the parameters

$$\frac{db}{dt} = g(b).$$

In hierarchical notation, filtering this augmented system is one way of estimating

$$[u, b|v] = [v|u, b]\frac{[u|b][b]}{[v]}$$

The classical separate bias Kalman Filter

Friedland (1969, 1982) considered the following linear filtering problem

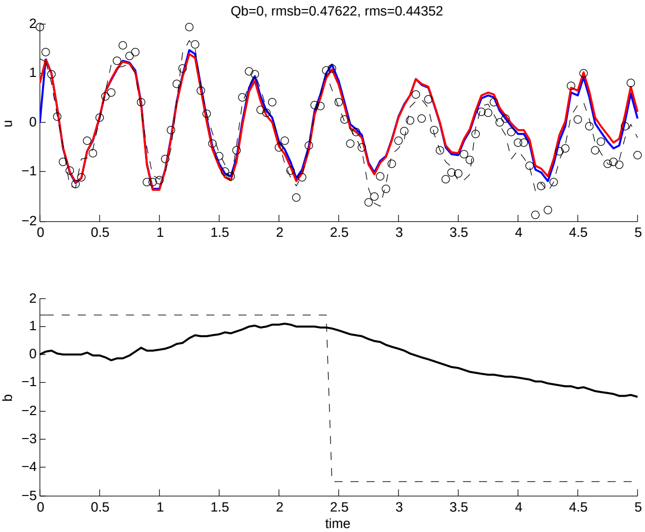
$$
\begin{aligned}
u_{m+1} &= Fu_m + B_{m+1} b_{m+1} + \sigma_{m+1} \\
b_{m+1} &= b_m + \sigma^b_{m+1}
\end{aligned}
$$

with a biased observation model

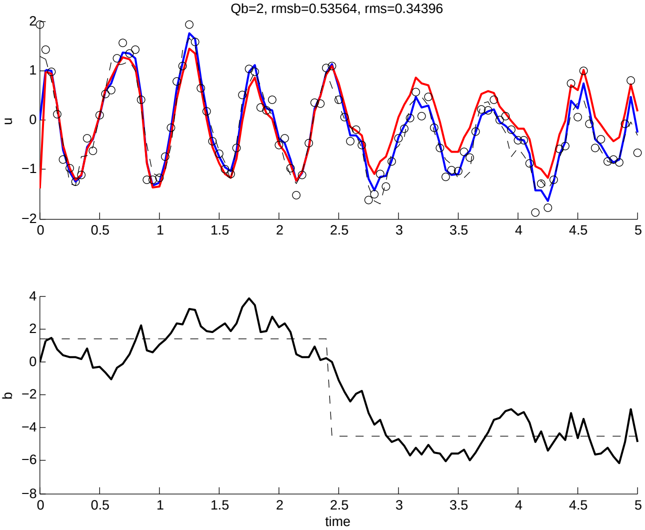
$$
v_{m+1} = Gu_{m+1} + C_{m+1} b_{m+1} + \sigma^o_{m+1}
$$

In this talk, we only consider unbiased observation, $C_m = 0$.

Constant bias model: $b_{m+1} = b_m$



White noise bias model: $b_{m+1} = b_m + \sigma^b_{m+1}$



Test model for true signal

Consider the following SDE

$$\frac{du(t)}{dt} = -\gamma(t)u(t) + \mathrm{i}\omega u(t) + \sigma \dot{W}(t) + f(t)$$

as a test model for filtering with model error.

Test model for true signal

Consider the following SDE

$$\frac{du(t)}{dt} = -\gamma(t)u(t) + \mathrm{i}\omega u(t) + \sigma \dot{W}(t) + f(t)$$

as a test model for filtering with model error.

To generate significant model errors as well as to mimic intermittent chaotic instability as often occurs in nature, we allow $\gamma(t)$ to switch between stable ($\gamma > 0$) and unstable ($\gamma < 0$) regimes according to a two-state Markov jump process.

Test model for true signal

Consider the following SDE

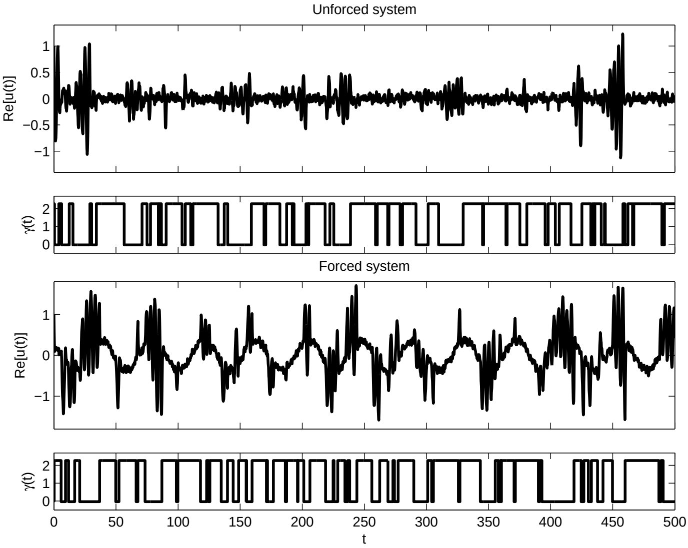
$$\frac{du(t)}{dt} = -\gamma(t)u(t) + \mathrm{i}\omega u(t) + \sigma\dot{W}(t) + f(t)$$

as a test model for filtering with model error.

To generate significant model errors as well as to mimic intermittent chaotic instability as often occurs in nature, we allow $\gamma(t)$ to switch between stable ($\gamma > 0$) and unstable ($\gamma < 0$) regimes according to a two-state Markov jump process.
Assume the following observation model:

$$v_m = u(t_m) + \sigma_m^o, \quad \sigma_m^o \sim \mathcal{N}(0, r^o). \tag{1}$$

True Signals for Unforced and Forced cases



Mean Stochastic Model

The prototype one-mode stochastic mean model

$$du(t) = \Big[(-\bar{\gamma} + \mathrm{i}\omega)u(t) + F(t)\Big]dt + \sigma\, dW(t)$$

where one fits the parameters using climatological statistical quantities such as the energy spectrum and correlation time.

Mean Stochastic Model

The prototype one-mode stochastic mean model

$$du(t) = \Big[(-\bar{\gamma} + \mathrm{i}\omega)u(t) + F(t)\Big]dt + \sigma dW(t)$$

where one fits the parameters using climatological statistical quantities such as the energy spectrum and correlation time.

This "poor-man" strategy is discussed in Harlim and Majda Nonlinearity 2008, Comm. Math. Sci. 2010.

Stochastic Parameterized Extended Kalman Filter:

We consider the following canonical model that accounts additive and multiplicative biases:

$$
\begin{aligned}
du(t) &= \Big[(-\gamma(t) + \mathrm{i}\omega)u(t) + F(t) + b(t)\Big]dt + \sigma dW(t) \\
db(t) &= (-\gamma_b + \mathrm{i}\omega_b)b(t)dt + \sigma_b dW_b(t) \\
d\gamma(t) &= -d_\gamma(\gamma(t) - \hat{\gamma})dt + \sigma_\gamma dW_\gamma(t)
\end{aligned}
$$

Stochastic Parameterized Extended Kalman Filter:

We consider the following canonical model that accounts additive and multiplicative biases:

$$
\begin{aligned}
du(t) &= \Big[(-\gamma(t) + \mathrm{i}\omega)u(t) + F(t) + b(t)\Big]dt + \sigma dW(t) \\
db(t) &= (-\gamma_b + \mathrm{i}\omega_b)b(t)dt + \sigma_b dW_b(t) \\
d\gamma(t) &= -d_\gamma(\gamma(t) - \hat{\gamma})dt + \sigma_\gamma dW_\gamma(t)
\end{aligned}
$$

We find stochastic parameters $\{\gamma_b, \omega_b, \sigma_b, d_\gamma, \sigma_\gamma\}$ that are robust for high filter skill beyond the MSM and in many occasions comparable to the perfectly specified filter model.

Stochastic Parameterized Extended Kalman Filter:

We consider the following canonical model that accounts additive and multiplicative biases:

$$\begin{aligned}
du(t) &= \Big[(-\gamma(t) + \mathrm{i}\omega)u(t) + F(t) + b(t)\Big]dt + \sigma dW(t) \\
db(t) &= (-\gamma_b + \mathrm{i}\omega_b)b(t)dt + \sigma_b dW_b(t) \\
d\gamma(t) &= -d_\gamma(\gamma(t) - \hat{\gamma})dt + \sigma_\gamma dW_\gamma(t)
\end{aligned}$$

We find stochastic parameters $\{\gamma_b, \omega_b, \sigma_b, d_\gamma, \sigma_\gamma\}$ that are robust for high filter skill beyond the MSM and in many occasions comparable to the perfectly specified filter model.

This special form has exactly solvable nonlinear solutions and moments and we do not need any linearization as in the standard EKF.

Next, we find the mean $\langle u(t)\rangle$: (Use the calculus tricks in Gershgorin-Majda 2008, 2010)

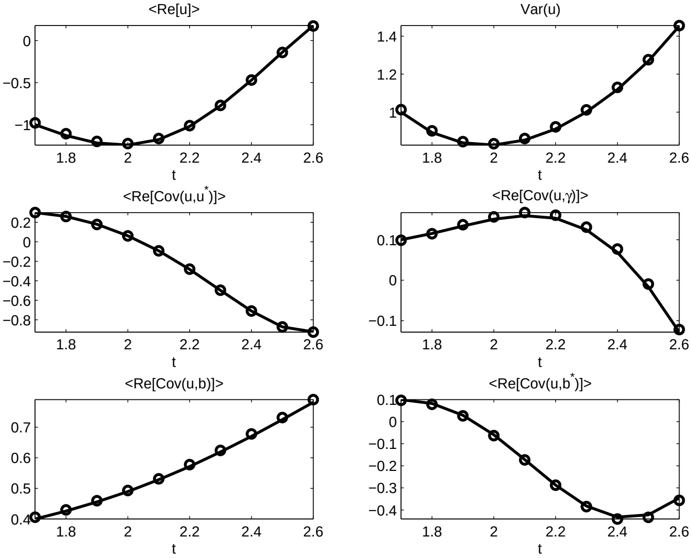
$$\begin{aligned}
\langle u(t)\rangle &= e^{\hat{\lambda}(t-t_0)}\Big(\langle u_0\rangle - Cov(u_0, J(t_0,t))\Big)e^{-\langle J(t_0,t)\rangle+\frac{1}{2}Var(J(t_0,t))} \\
&+\int_{t_0}^{t} e^{\hat{\lambda}(t-s)}\Big(\hat{b}+e^{\lambda_b(s-t_0)}\Big(\langle b_0\rangle-\hat{b}-Cov(b_0,J(s,t))\Big)\Big) \\
&\times e^{-\langle J(s,t)\rangle+\frac{1}{2}Var(J(s,t))}ds \\
&+\int_{t_0}^{t} e^{\hat{\lambda}(t-s)}f(s)e^{-\langle J(s,t)\rangle+\frac{1}{2}Var(J(s,t))}ds \qquad (2)
\end{aligned}$$

where

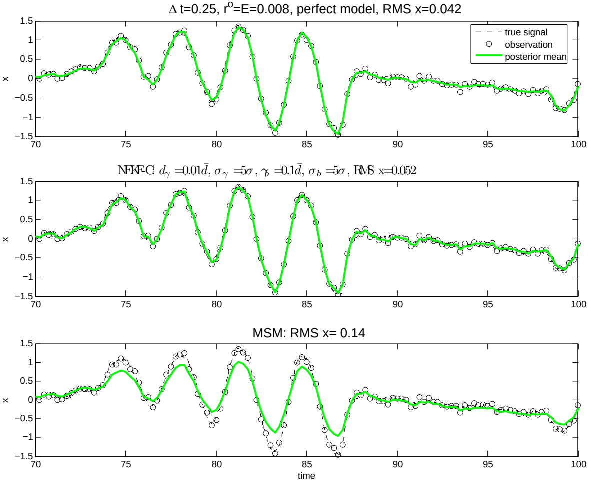
$$\begin{aligned}
\hat{\lambda} &= -\hat{\gamma}+\mathrm{i}\omega, \\
J(s,t) &= \int_{s}^{t}(\gamma(s')-\hat{\gamma})ds',
\end{aligned}$$

and next the cross-covariances ...

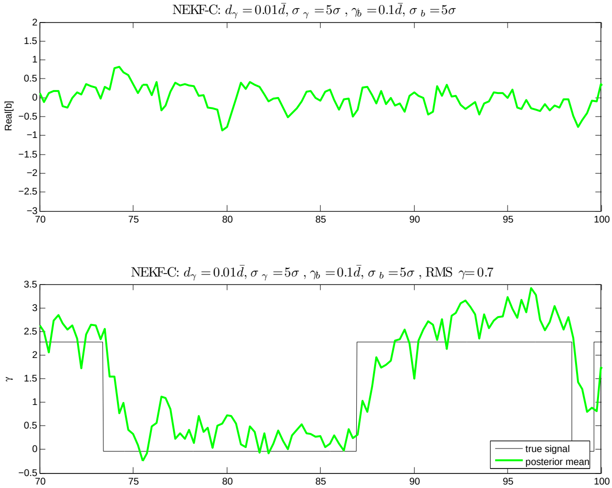
SPEKF: Checking first and second ordered statistics



One mode demonstration of the filtered solution: observed mode



One mode demonstration of the filtered solution: unobserved parameters



Canonical Spatially Extended Turbulent Systems

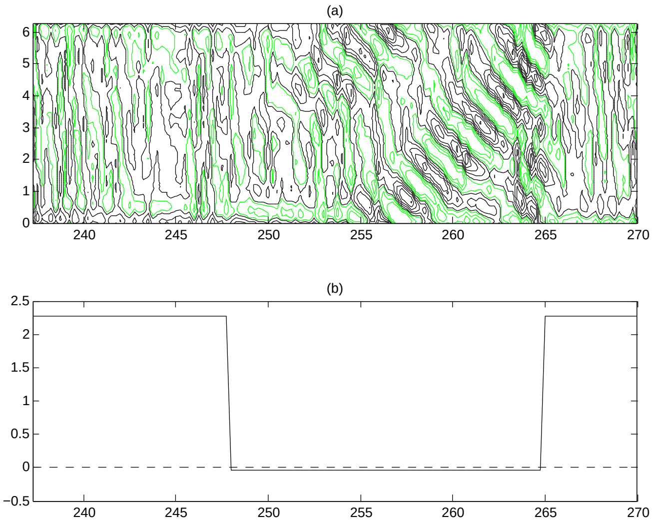
We consider a stochastic PDE with time-dependent damping Langevin equation for the first five Fourier modes, i.e.,

$$\frac{du_k(t)}{dt} = -\gamma_k(t)u_k(t) + \mathrm{i}\omega_k u_k(t) + \sigma_k \dot{W}_k(t) + f_k(t), \quad k = 1, \dots, 5,$$

and linear Langevin equation with constant damping $\bar{d}$ for modes $k > 5$,

$$\frac{du_k(t)}{dt} = -\bar{d}u_k(t) + \mathrm{i}\omega_k u_k(t) + \sigma_k \dot{W}_k(t) + f_k(t), \quad k > 5.$$

Turbulent barotropic Rossby wave equation: $\omega_k = -\beta/k, E_k = k^{-3}$



Incorrectly specified forcings:

Here, we consider a true signal with forcing given by

$$\hat{f}_k(t) = A_{f,k} \exp\Big(\mathrm{i}(\omega_{f,k} t + \phi_{f,k})\Big), \tag{3}$$

for $k = 1, \ldots, 7$ with amplitude $A_{f,k}$, frequency $\omega_{f,k}$, and phase $\phi_{f,k}$ drawn randomly from uniform distributions,

$$\begin{aligned} A_{f,k} &\sim U(0.6, 1), \\ \omega_{f,k} &\sim U(0.1, 0.4), \\ \phi_{f,k} &\sim U(0, 2\pi), \\ \hat{f}_k &= \hat{f}^*_{-k}, \end{aligned}$$

and unforced, $\hat{f}_k(t) = 0$, for modes $k > 7$. However, we do not specify this true forcing to the filter model, i.e., we use $\tilde{f}_k = 0$ for all modes.

Reduced Filter Domain Kalman Filter for regularly spaced sparse observation

We consider regularly spaced sparse observations: $(2M+1)$ observations of $(2N+1)$ model grid points. The Fourier coefficients of the observation model is given as

$$\hat{v}_{\ell,m} = \sum_{k \in \mathcal{A}(\ell)} \hat{u}_{k,m} + \hat{\sigma}^o_m,$$

where

$$\mathcal{A}(\ell) = \{k | k = \ell + (2M+1)q, q \in \mathbb{Z}, |\ell| \leq N\}$$

is the aliasing set of wavenumber ℓ. (Majda-Grote PNAS 2007)

Reduced Filter Domain Kalman Filter for regularly spaced sparse observation

We consider regularly spaced sparse observations: $(2M+1)$ observations of $(2N+1)$ model grid points. The Fourier coefficients of the observation model is given as

$$\hat{v}_{\ell,m} = \sum_{k\in\mathcal{A}(\ell)} \hat{u}_{k,m} + \hat{\sigma}^o_m,$$

where

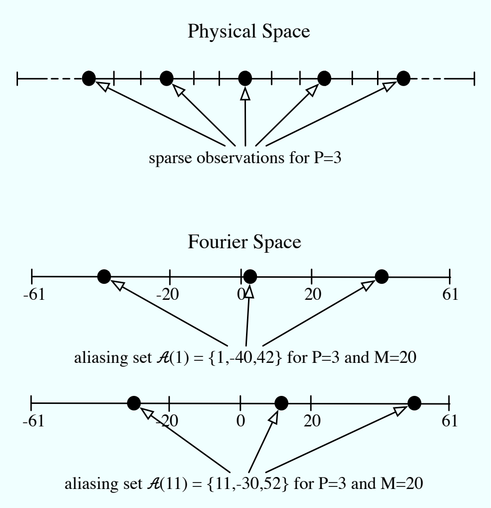
$$\mathcal{A}(\ell) = \{k | k = \ell + (2M+1)q, q \in \mathbb{Z}, |\ell| \leq N\}$$

is the aliasing set of wavenumber ℓ. (Majda-Grote PNAS 2007)

When the energy spectrum is decaying as a function of k, we can use the following reduced observation model

$$\hat{v}'_{\ell,m} \equiv \hat{v}_{\ell,m} - \sum_{k\in\mathcal{A}(\ell),k\neq\ell} \hat{u}_{k,m|m-1} = \hat{u}_{\ell,m} + \hat{\sigma}^o_m.$$

Example: 123 grid pts (61 modes) but only 41 observations (20 modes) available



Incorrectly specified forcings, observed only 15 observations of 105 grid points

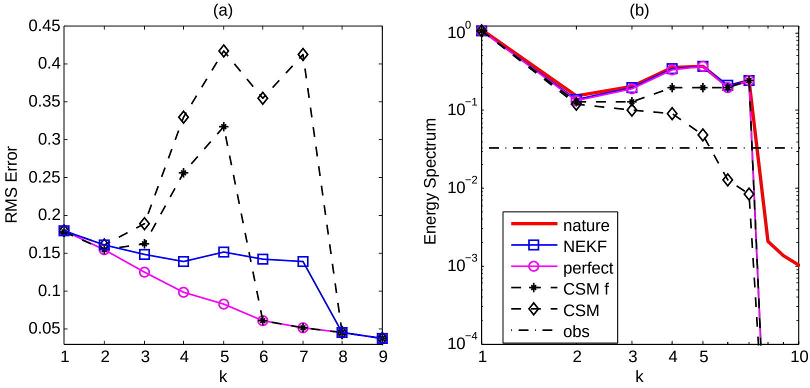


Table of RMSE for the SPDE test case with intermittent burst of instability

Forcing	unforced case	correct forcing	incorrect forcing
r^o	0.2	0.3	0.5
perfect filter	0.35	0.39	0.45
MSM	0.39	0.48	0.73
$\text{MSM}_{f=0}$	-	-	1.17
SPEKF-C	0.38	0.44	0.59
SPEKF-M	0.36	0.42	0.79
SPEKF-A	0.39	0.46	0.60

Canonical Model for Midlatitude Geophysical Flows:

The dynamical equations for the perturbed variables are:

$$\begin{aligned} \frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial q_1}{\partial x} + (\beta + k_d^2 U)\frac{\partial \psi_1}{\partial x} + \nu\nabla^8 q_1 &= 0 \\ \frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U\frac{\partial q_2}{\partial x} + (\beta - k_d^2 U)\frac{\partial \psi_2}{\partial x} + \nu\nabla^8 q_2 + \kappa\nabla^2\psi_2 &= 0 \end{aligned}$$

where q_j is the quasi-geostrophic potential vorticity given as

$$q_j = \beta y + \nabla^2\psi_j + \frac{k_d^2}{2}(\psi_{3-j} - \psi_j)$$

with $\vec{u} = \nabla^{\perp}\psi$, $k_d = \sqrt{8}/L_d$.

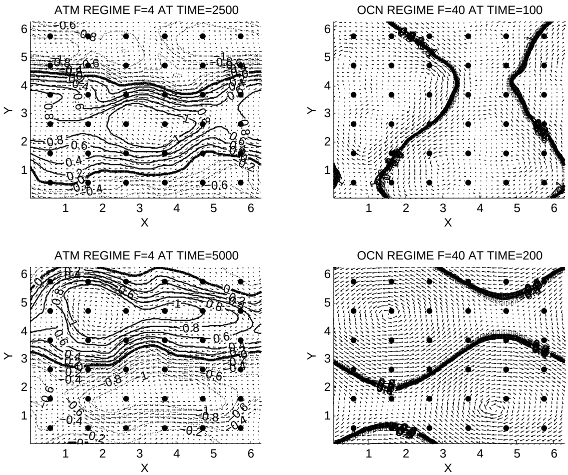
In the two-layer case, the barotropic vertical and baroclinic modes are defined as $\psi_b = (\psi_1 + \psi_2)/2$ and $\psi_c = (\psi_1 - \psi_2)/2$, respectively.

In the two-layer case, the barotropic vertical and baroclinic modes are defined as $\psi_b = (\psi_1 + \psi_2)/2$ and $\psi_c = (\psi_1 - \psi_2)/2$, respectively.
Notice that the barotropic mode dynamical equation,

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \left(J(\psi_c, q_c) + U \frac{\partial \nabla^2 \psi_c}{\partial x} - \kappa \nabla^2 \psi_c \right) = 0$$

is numerically stiff when k_d^2 is large (ocean case).

The 2-layer QG model with baroclinic instability



Stochastic Models for Filtering the barotropic mode:

Recall that

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \Big(\text{baroclinic term}\Big) = 0$$

where $q_b = \beta y + \nabla^2 \psi_b$.

Stochastic Models for Filtering the barotropic mode:

Recall that

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \Big(\text{baroclinic term}\Big) = 0$$

where $q_b = \beta y + \nabla^2 \psi_b$.

Poorman's stochastic models: replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes.

Stochastic Models for Filtering the barotropic mode:

Recall that

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \Big(\text{baroclinic term}\Big) = 0$$

where $q_b = \beta y + \nabla^2 \psi_b$.

Poorman's stochastic models: replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes.

Discrete Fourier Transform:

$$\psi = \sum_{k,\ell} \hat{\psi}_{k,\ell} e^{\mathrm{i}(kx+\ell y)}$$

Thus, each horizontal mode has the following form

$$d\hat{\psi}(t) = (-d + \mathrm{i}\omega)\hat{\psi}(t)dt + f(t)dt$$

Stochastic Models for Filtering the barotropic mode:

Recall that

$$\frac{\partial q_b}{\partial t} + J(\psi_b, q_b) + \beta \frac{\partial \psi_b}{\partial x} + \kappa \nabla^2 \psi_b + \nu \nabla^8 q_b + \Big(\text{baroclinic term}\Big) = 0$$

where $q_b = \beta y + \nabla^2 \psi_b$.

Poorman's stochastic models: replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes.

Discrete Fourier Transform:

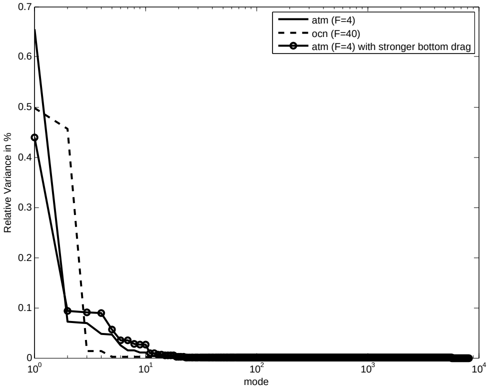
$$\psi = \sum_{k,\ell} \hat{\psi}_{k,\ell} e^{\mathrm{i}(kx+\ell y)}$$

Thus, each horizontal mode has the following form

$$d\hat{\psi}(t) = (-d + \mathrm{i}\omega)\hat{\psi}(t)dt + f(t)dt + \sigma dW(t)$$

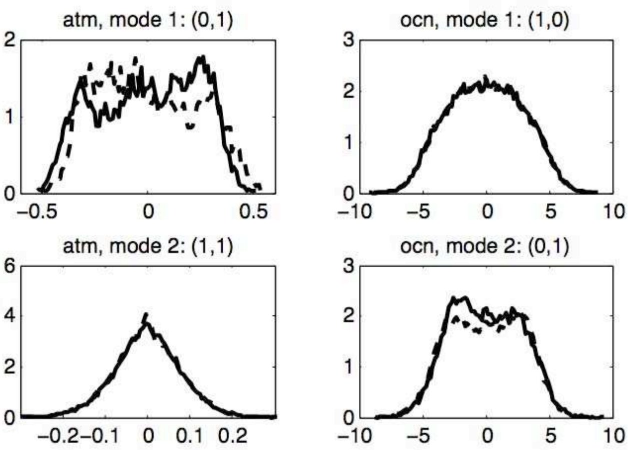
and our task is to parameterize $d, \omega, f(t), \sigma$?

Statistical Quantities: Climatological variances of the barotropic mode

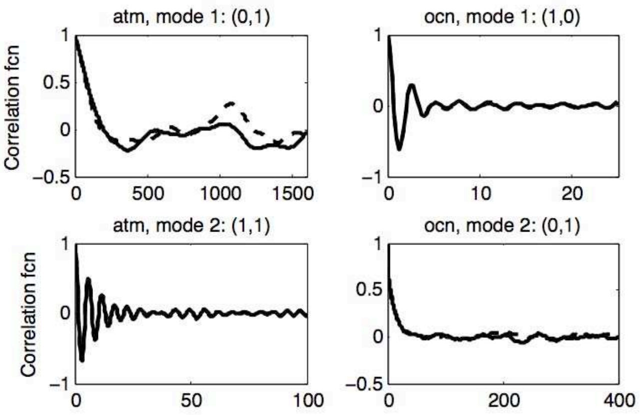


"Atmospheric" case (k_d^2 is small) and "oceanic" case (k_d^2 is large).

Statistical Quantities: Histogram "marginal pdf's"



Statistical Quantities: Correlation functions



Mean Stochastic Models: parameterize d, ω, f, σ

We set $f(t)$ to be a constant equals to the climatological mean $\langle\hat{\psi}\rangle$ (long time average).

Mean Stochastic Models: parameterize d, ω, f, σ

We set $f(t)$ to be a constant equals to the climatological mean $\langle \hat{\psi} \rangle$ (long time average).

MSM2 We use the linear dispersion ω, and we fit the damping and noise strengths to the spectrum and decorrelation time

$$
\begin{aligned}
Var(\hat{\psi}) &= \frac{\sigma^2}{2d} \\
Re[T_{corr}] &\equiv \frac{1}{Var(\hat{\psi})} \int_0^\infty Re[C(\tau)] d\tau = \frac{1}{d}
\end{aligned}
$$

Mean Stochastic Models: parameterize d, ω, f, σ

We set $f(t)$ to be a constant equals to the climatological mean $\langle \hat{\psi} \rangle$ (long time average).

MSM2 We use the linear dispersion ω, and we fit the damping and noise strengths to the spectrum and decorrelation time

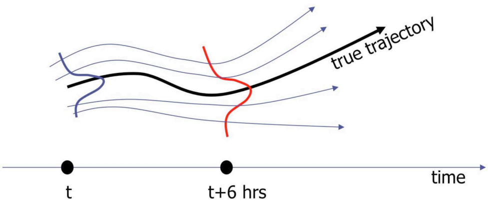
$$
\begin{aligned}
Var(\hat{\psi}) &= \frac{\sigma^2}{2d} \\
Re[T_{corr}] &\equiv \frac{1}{Var(\hat{\psi})} \int_0^\infty Re[C(\tau)] d\tau = \frac{1}{d}
\end{aligned}
$$

MSM1 Ignore the linear dispersion and solve the following

$$
\begin{aligned}
Var(\hat{\psi}) &= \frac{\sigma^2}{2d} \\
T_{corr} &\equiv \frac{1}{Var(\hat{\psi})} \int_0^\infty C(\tau) d\tau = \frac{1}{d + \mathrm{i}\omega}
\end{aligned}
$$

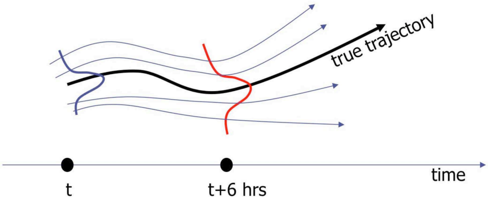
Local least squares EAKF (Anderson 2003)

Approximate the prior error covariance matrix by ensemble covariance.



Local least squares EAKF (Anderson 2003)

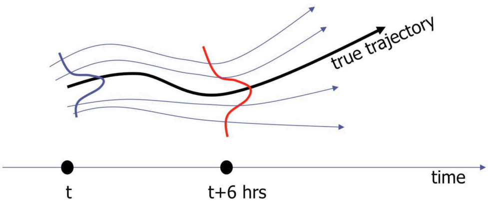
Approximate the prior error covariance matrix by ensemble covariance.



How many ensemble member? How to avoid ensemble collapse and spurious correlations due to finite ensemble size?

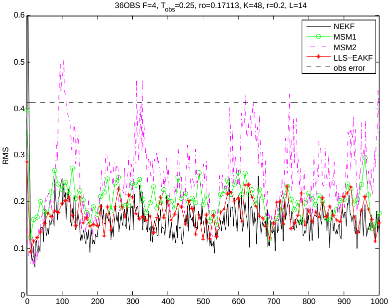
Local least squares EAKF (Anderson 2003)

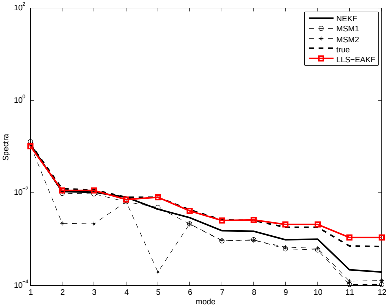
Approximate the prior error covariance matrix by ensemble covariance.

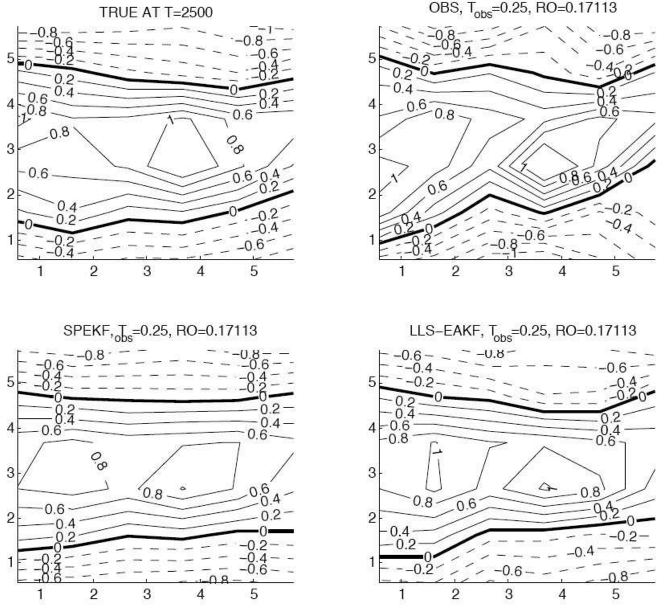


How many ensemble member? How to avoid ensemble collapse and spurious correlations due to finite ensemble size? Computationally, EAKF requires extensive tunings of ensemble size, local box size, covariance inflation, and in the ocean case, integration time step need to be reduced.

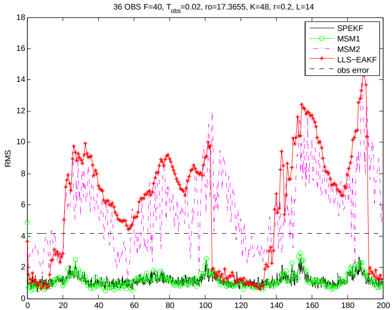
Longer deformation radius case (“atmospheric” regime).

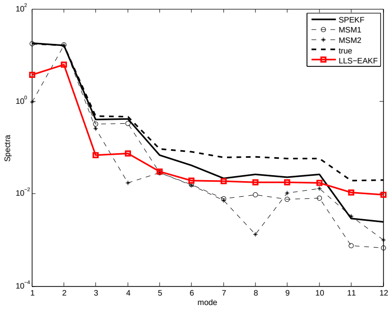


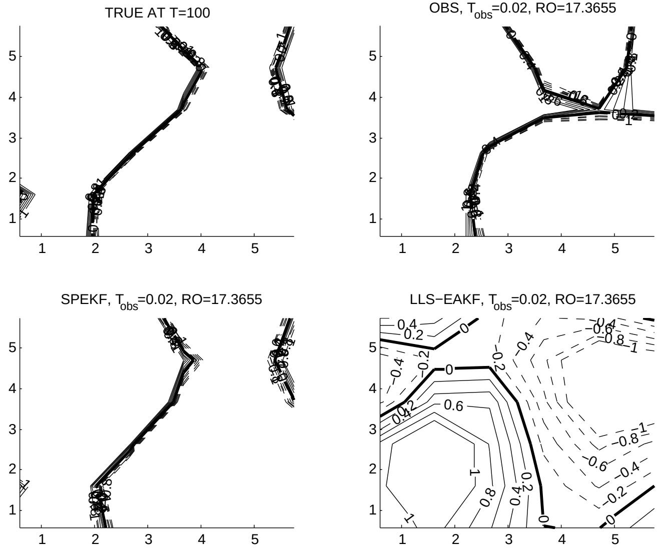



TRUE AT T=2500
OBS, T$_{obs}$=0.25, RO=0.17113
SPEKF, T$_{obs}$=0.25, RO=0.17113
LLS−EAKF, T$_{obs}$=0.25, RO=0.17113

Shorter deformation radius case ("oceanic" regime).







TRUE AT T=100
OBS, Tobs=0.02, RO=17.3655
SPEKF, Tobs=0.02, RO=17.3655
LLS−EAKF, Tobs=0.02, RO=17.3655
1
2
3
4
5

Summary:

1. MSM: We introduce reduced stochastic models through replacing the nonlinearity and baroclinic components with Ornstein-Uhlenbeck process for filtering purpose. This reduced poor man's strategy is numerically very cheap and accurate in a regime when the dynamical systems is strongly chaotic and fully turbulent.

Summary:

1. MSM: We introduce reduced stochastic models through replacing the nonlinearity and baroclinic components with Ornstein-Uhlenbeck process for filtering purpose. This reduced poor man's strategy is numerically very cheap and accurate in a regime when the dynamical systems is strongly chaotic and fully turbulent.
2. SPEKF: We introduce a paradigm model for "online" learning both the additive and multiplicative biases from observations beyond the MSM. This model is analytically solvable such that NO LINEARIZATION is needed when Kalman filter formula is utilized.