Applied Mathematical Tools for Tropical Data Assimilation

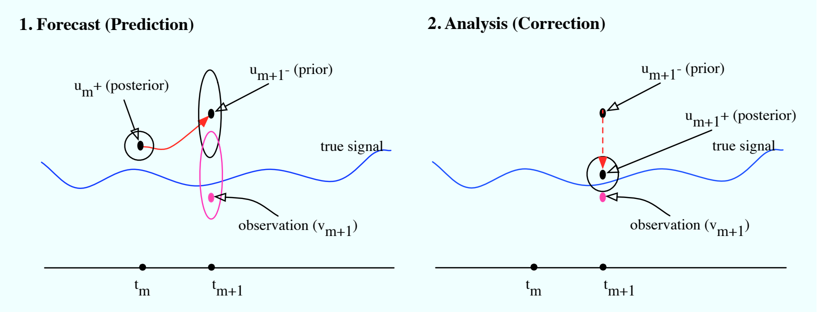
John Harlim [1]

Department of Mathematics
North Carolina State University

May 4, 2011

[1]Support: NC State startup fund and ONR N00014-11-1-1310

What is filtering (or data assimilation)?



The correction step is an application of Bayesian update

$$p(u^+_{m+1}) \equiv p(u^-_{m+1}|v_{m+1}) \propto p(u^-_{m+1})p(v_{m+1}|u^-_{m+1})$$

When Gaussianity and linearity are assumed, one obtains the Kalman filter.

Difficulties of tropical weather prediction:

- The presence of **multiple scales processes without clear scale gap**: cumulus clouds (1-2 km), mesoscale convective systems (5-100 km), equatorial synoptic scale (1000 km), convectively coupled Kelvin waves and two-days waves, and planetary scale such as the MJO.

Difficulties of tropical weather prediction:

- The presence of **multiple scales processes without clear scale gap**: cumulus clouds (1-2 km), mesoscale convective systems (5-100 km), equatorial synoptic scale (1000 km), convectively coupled Kelvin waves and two-days waves, and planetary scale such as the MJO.
- There are **intermittent behavior in the smaller scale processes**; this makes the classical stochastic averaging technique not very useful.

Difficulties of tropical weather prediction:

- The presence of **multiple scales processes without clear scale gap**: cumulus clouds (1-2 km), mesoscale convective systems (5-100 km), equatorial synoptic scale (1000 km), convectively coupled Kelvin waves and two-days waves, and planetary scale such as the MJO.
- There are **intermittent behavior in the smaller scale processes**; this makes the classical stochastic averaging technique not very useful.
- **Sparsely observed wind velocity field** due to limited radio-sounding devices in the tropical region. On the other hand, mass data (temperature, humidity, and pressure) are horizontally plentifully observed from satellite measurement.

Difficulties of tropical weather prediction:

- The presence of **multiple scales processes without clear scale gap**: cumulus clouds (1-2 km), mesoscale convective systems (5-100 km), equatorial synoptic scale (1000 km), convectively coupled Kelvin waves and two-days waves, and planetary scale such as the MJO.
- There are **intermittent behavior in the smaller scale processes**; this makes the classical stochastic averaging technique not very useful.
- **Sparsely observed wind velocity field** due to limited radio-sounding devices in the tropical region. On the other hand, mass data (temperature, humidity, and pressure) are horizontally plentifully observed from satellite measurement.
- Various data assimilation techniques are **successful for midlatitude weather dynamics** but they may not be so successful due to all these issues.

Research Program:

Our goal is to design data assimilation (filtering scheme) that addresses these issues.

- **Complex Dynamical Systems with Model Errors (SPEKF):**
  Majda, Harlim, and Gershgorin, *Mathematical Strategies for Filtering Turbulent Dynamical Systems*, Discrete Contin. Dynam. Syst. A, 27(2), 441-486, 2010.
  Majda and Harlim, *Systematic Strategies for Real Time Filtering of Turbulent Signals in Complex Systems*, Cambridge University Press.

Research Program:

Our goal is to design data assimilation (filtering scheme) that addresses these issues.

- **Complex Dynamical Systems with Model Errors (SPEKF):**
  Majda, Harlim, and Gershgorin, *Mathematical Strategies for Filtering Turbulent Dynamical Systems*, Discrete Contin. Dynam. Syst. A, 27(2), 441-486, 2010.
  Majda and Harlim, *Systematic Strategies for Real Time Filtering of Turbulent Signals in Complex Systems*, Cambridge University Press.
- **Effects of Interpolated Data:**
  Harlim, *Interpolating Irregularly Spaced Observations for Filtering Turbulent Complex Systems*, submitted to SIAM J. Sci. Comput.

Research Program:

Our goal is to design data assimilation (filtering scheme) that addresses these issues.

- **Complex Dynamical Systems with Model Errors (SPEKF):**
  Majda, Harlim, and Gershgorin, *Mathematical Strategies for Filtering Turbulent Dynamical Systems*, Discrete Contin. Dynam. Syst. A, 27(2), 441-486, 2010.
  Majda and Harlim, *Systematic Strategies for Real Time Filtering of Turbulent Signals in Complex Systems*, Cambridge University Press.
- **Effects of Interpolated Data:**
  Harlim, *Interpolating Irregularly Spaced Observations for Filtering Turbulent Complex Systems*, submitted to SIAM J. Sci. Comput.
- **Multiscale dynamical systems with intermittency:**
  Harlim, *Numerical Strategies for Filtering Partially Observed Stiff Stochastic Differential Equations*", J. Comput. Phys., 230(3), 744-762, 2011.
  Kang and Harlim, *A Fast Filtering Framework for Assimilating Partially Observed Multiscale Systems: Macro-Micro-Filter*, submitted to MWR.

Online Model Error Estimation Strategy

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

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

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

$$\frac{d\lambda}{dt} = g(\lambda).$$

Online Model Error Estimation Strategy

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

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

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

$$\frac{d\lambda}{dt} = g(\lambda).$$

Then, perform state estimation (or filtering technique) on (u, λ) using noisy observations v to obtain

$$P(u, \lambda | v) \propto P(u|\lambda) P(\lambda) P(v|u, \lambda)$$

Stochastic Parameterized "Extended" Kalman Filter:

We consider a stochastic model for the evolution of state variable $\hat{\psi}(t)$ together with **combined** additive, $b(t)$, and multiplicative, $\gamma(t)$, bias correction terms:

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

Stochastic Parameterized "Extended" Kalman Filter:

We consider a stochastic model for the evolution of state variable $\hat{\psi}(t)$ together with **combined** additive, $b(t)$, and multiplicative, $\gamma(t)$, bias correction terms:

$$
\begin{aligned}
d\hat{\psi}(t) &= \Big((-\gamma(t) + \mathrm{i}\omega)\hat{\psi}(t) + b(t) + \hat{f}(t)\Big)dt + \sigma dW(t), \\
db(t) &= (-\gamma_b + \mathrm{i}\omega_b)b(t)dt + \sigma_b dW_b(t), \\
d\gamma(t) &= -d_\gamma\Big(\gamma(t) - d\Big)dt + \sigma_\gamma dW_\gamma(t).
\end{aligned}
$$

Here, this nonlinear SDE is exactly solvable and statistics are exactly solvable conditional to Gaussian initial condition. Need to empirically tune $\gamma_b, \omega_b, \sigma_b, d_\gamma, \sigma_\gamma$ but they are quite robust depending of the physical nature of the mode (see GHM-JCP 2010a, 2010b, BGM 2011, KMS 2011).

Application of SPEKF on Geophysical Flows (HM-MWR2010):

The dynamical equations for the perturbed variables about uniform shear with stream function $\Psi_1 = -Uy, \Psi_2 = Uy$:

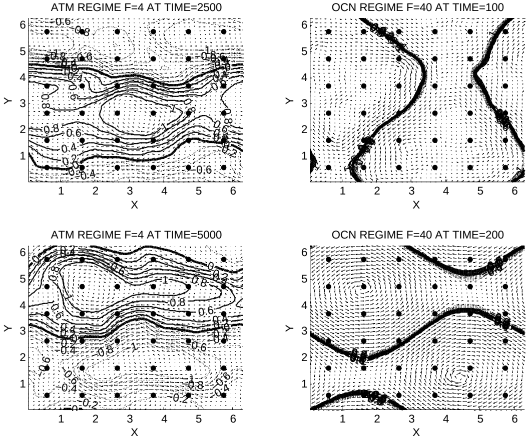
$$\begin{aligned} \frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial q_1}{\partial x} + (\beta + k_d^2 U)\frac{\partial \psi_1}{\partial x} + \nu\nabla^8 q_1 &= 0 \\ \frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U\frac{\partial q_2}{\partial x} + (\beta - k_d^2 U)\frac{\partial \psi_2}{\partial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

q_j is the quasi-geostrophic potential vorticity given as

$$q_j = \nabla^2\psi_j + \beta y + \frac{k_d^2}{2}(\psi_{3-j} - \psi_j), \quad j = 1, 2,$$

with $\vec{u} = \nabla^{\perp}\psi$, $k_d = \sqrt{8}/L_d$ (see Smith et al. 2002).

The 2-layer QG model with baroclinic instability



"Atmosphere" regime, longer deformation radius $F = 1/L_d^2 = 4$ (first column) and "Ocean" regime, $F = 40$ (second column). (see Kleeman and Majda 2005)

Stochastic Models for Filtering the barotropic mode:

Recall that

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

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

Stochastic Models for Filtering the barotropic mode:

Recall that

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

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

Fourier Transform:

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

Thus, each horizontal mode has the following form

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

Stochastic Models for Filtering the barotropic mode:

Replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes (HM Nonlinearity 08, Comm. Math. Sci. 09) or AR(p)-processes (KH, submitted to Phys D). That is,

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

and our task is to parameterize d, ω, σ.

Stochastic Models for Filtering the barotropic mode:

Replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes (HM Nonlinearity 08, Comm. Math. Sci. 09) or AR(p)-processes (KH, submitted to Phys D). That is,

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

and our task is to parameterize d, ω, σ.

Parameterizations:

1. Regressions to empirical statistics from a long time series (Mean Stochastic Models).

Stochastic Models for Filtering the barotropic mode:

Replace the nonlinear terms and all of the baroclinic components by Ornstein-Uhlenbeck processes (HM Nonlinearity 08, Comm. Math. Sci. 09) or AR(p)-processes (KH, submitted to Phys D). That is,

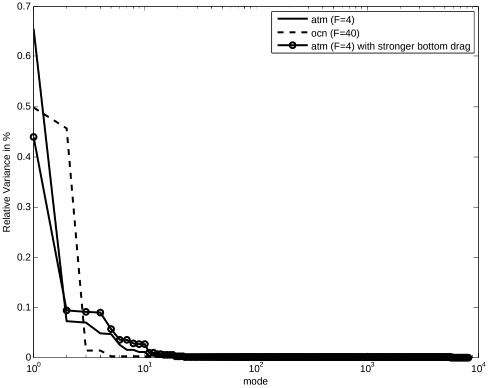
$$d\hat{\psi}(t) = (-d + \mathrm{i}\omega)\hat{\psi}(t)dt + \hat{f}(t)dt + \sigma dW(t)$$

and our task is to parameterize d, ω, σ.

Parameterizations:

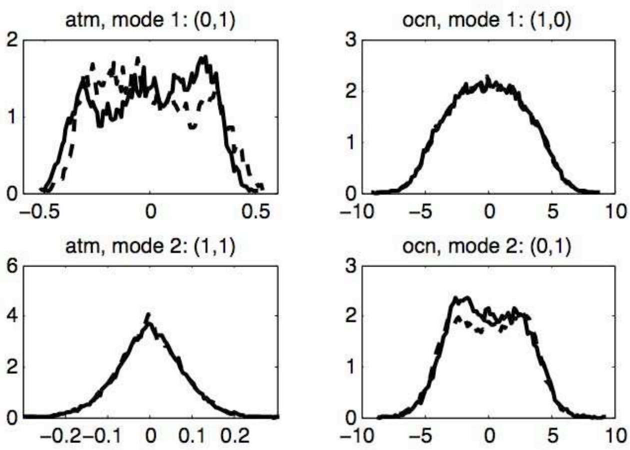
1. Regressions to empirical statistics from a long time series (Mean Stochastic Models).
2. On-the-fly parameterization (SPEKF).

Statistical Quantities: Climatological variances of the barotropic mode

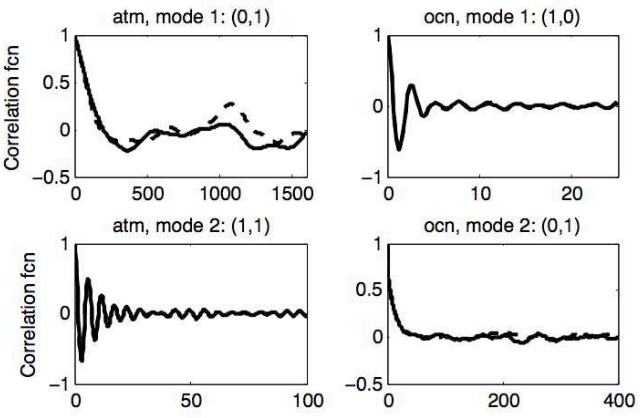


"Atmospheric" case (k_d^2 is small) and "oceanic" case (k_d^2 is large).

Statistical Quantities: Histogram "marginal pdf's"



Statistical Quantities: Correlation functions



Reduced Filters:

Apply the Bayesian framework to these stochastic models (MSM, SPEKF model) to obtain "best" posterior estimate:

$$P(u, \lambda | v) \propto P(\lambda) P(u | \lambda) P(v | u, \lambda)$$

Reduced Filters:

Apply the Bayesian framework to these stochastic models (MSM, SPEKF model) to obtain "best" posterior estimate:

$$P(u, \lambda | v) \propto P(\lambda) P(u|\lambda) P(v|u, \lambda)$$

Remarks:

- If these densities are Gaussian, the analytical solution is known as Kalman filter formula and one can use the posterior mean and covariance as the best estimate and its uncertainty.

Reduced Filters:

Apply the Bayesian framework to these stochastic models (MSM, SPEKF model) to obtain "best" posterior estimate:

$$P(u,\lambda|v) \propto P(\lambda)P(u|\lambda)P(v|u,\lambda)$$

Remarks:

- If these densities are Gaussian, the analytical solution is known as Kalman filter formula and one can use the posterior mean and covariance as the best estimate and its uncertainty.
- The update in SPEKF uses Kalman filter formula but the prior statistics are solutions of a set of **nonlinear equations conditional to Gaussian initial conditions**.

Reduced Filters:

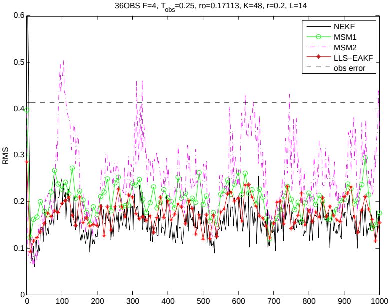
Apply the Bayesian framework to these stochastic models (MSM, SPEKF model) to obtain "best" posterior estimate:

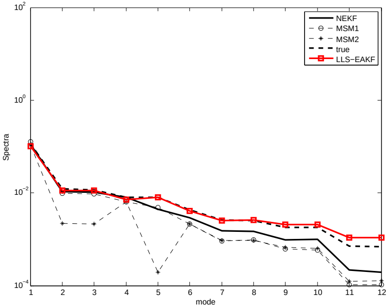
$$P(u,\lambda|v) \propto P(\lambda)P(u|\lambda)P(v|u,\lambda)$$

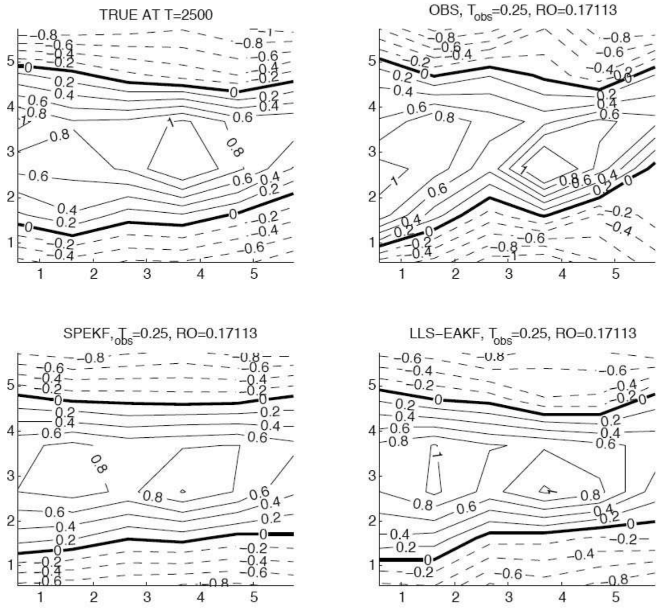
Remarks:

- If these densities are Gaussian, the analytical solution is known as Kalman filter formula and one can use the posterior mean and covariance as the best estimate and its uncertainty.
- The update in SPEKF uses Kalman filter formula but the prior statistics are solutions of a set of **nonlinear equations conditional to Gaussian initial conditions**.
- For special observation network ("plentiful" and regularly spaced sparse network) with i.i.d noise, we have a **reduced filter on each Fourier component independently.**

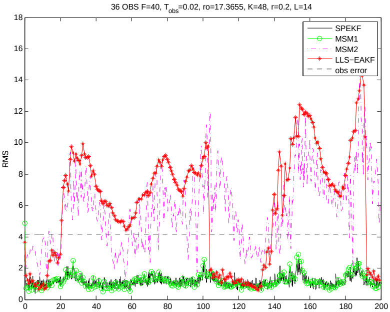
Longer deformation radius case (“atmospheric” regime).

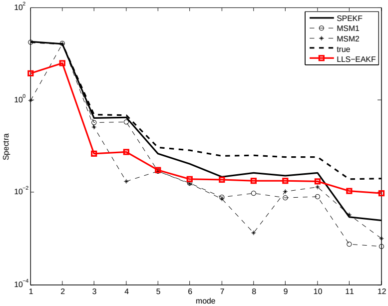


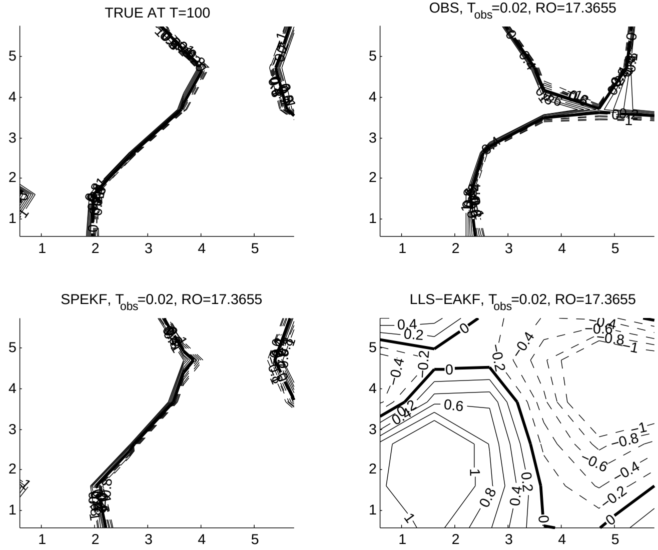



TRUE AT T=2500
OBS, T_{obs}=0.25, RO=0.17113
SPEKF, T_{obs}=0.25, RO=0.17113
LLS−EAKF, T_{obs}=0.25, RO=0.17113

Shorter deformation radius case (“oceanic” regime).






TRUE AT T=100
OBS, T_obs=0.02, RO=17.3655
SPEKF, T_obs=0.02, RO=17.3655
LLS−EAKF, T_obs=0.02, RO=17.3655

Effect of Interpolated data:

1. All results before are for sparse observations that are located at model grid points. How about irregularly spaced observed data?

Effect of Interpolated data:

1. All results before are for sparse observations that are located at model grid points. How about irregularly spaced observed data?
2. The typical observation model is

$$\tilde{u}_m = g(\vec{u}_m) + \tilde{\epsilon}_m, \quad \tilde{\epsilon}_m \sim \mathcal{N}(0, R),$$

   where g is an interpolates model state $\vec{u}$ to irregularly located observations $\tilde{u}$.

Effect of Interpolated data:

1. All results before are for sparse observations that are located at model grid points. How about irregularly spaced observed data?
2. The typical observation model is

$$\tilde{u}_m = g(\vec{u}_m) + \tilde{\epsilon}_m, \quad \tilde{\epsilon}_m \sim \mathcal{N}(0, R),$$

where g is an interpolates model state $\vec{u}$ to irregularly located observations $\tilde{u}$. How about interpolate the irregularly spaced observations to the model grid point?

$$h(\tilde{u}_m) = \vec{u}_m + \epsilon_m, \quad E(|\epsilon_m|^p) = E(|h(\tilde{\epsilon}_m)|^p)$$

What's the justification of doing this?

Effect of Interpolated data:

1. All results before are for sparse observations that are located at model grid points. How about irregularly spaced observed data?
2. The typical observation model is

$$\tilde{u}_m = g(\vec{u}_m) + \tilde{\epsilon}_m, \quad \tilde{\epsilon}_m \sim \mathcal{N}(0, R),$$

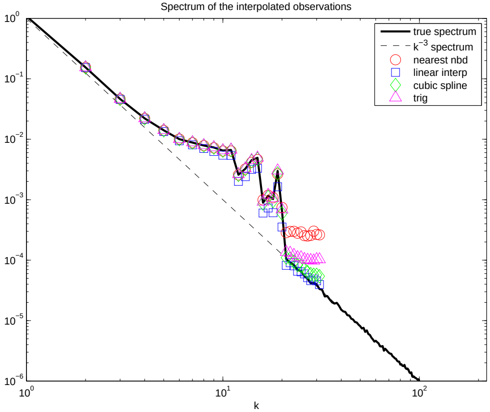
   where g is an interpolates model state $\vec{u}$ to irregularly located observations $\tilde{u}$. How about interpolate the irregularly spaced observations to the model grid point?

$$h(\tilde{u}_m) = \vec{u}_m + \epsilon_m, \quad E(|\epsilon_m|^p) = E(|h(\tilde{\epsilon}_m)|^p)$$

   What's the justification of doing this?
3. How does the uncertainty due to the interpolation errors affects the data assimilation? Which interpolation technique should we use?

Effect of interpolation on energy spectrum:

Interpolated spectrum of a “toy” model for barotropic Rossby waves with intermittent instability (see Ch 5, 8 of MH book or GHM JCP 10b).



The trig interp was considered in (Majda and Grote PNAS 07).

Effect of interpolation on covariance

- if one interpolate observations with i.i.d. noise, then the interpolated noise covariance, $E(|\epsilon_m|^2) = E(|h(\tilde{\epsilon}_m)|^2)$, is **not** diagonal.

Effect of interpolation on covariance

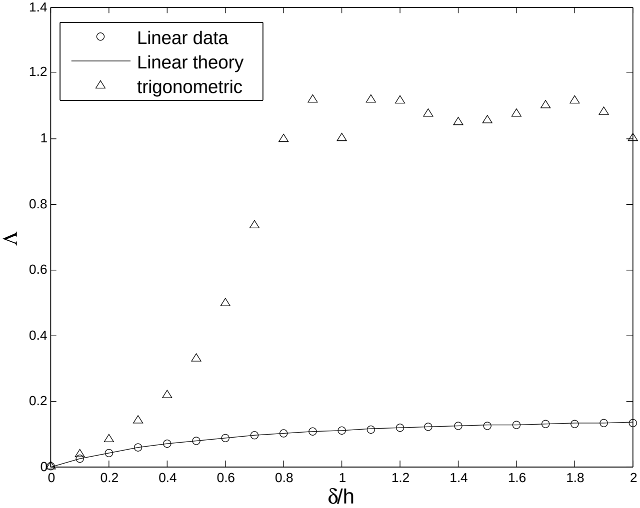
- if one interpolate observations with i.i.d. noise, then the interpolated noise covariance, $E(|\epsilon_m|^2) = E(|h(\tilde{\epsilon}_m)|^2)$, is **not** diagonal.
- The interpolation **operator** h we consider here is **linear**, so the interpolated noises are Gaussian. This may not be true in general.

Effect of interpolation on covariance

- if one interpolate observations with i.i.d. noise, then the interpolated noise covariance, $E(|\epsilon_m|^2) = E(|h(\tilde{\epsilon}_m)|^2)$, is **not** diagonal.
- The interpolation **operator** h we consider here is **linear**, so the interpolated noises are Gaussian. This may not be true in general.
- **Proposition:** Let $\{\sigma_j = \sigma(x_j)\}_{j=0}^{2M}$ be i.i.d. noises with variance r^o at regularly spaced grid points. Let us perturb a single observation site $\tilde{x}_j$ by δ, i.e., $\tilde{x}_j = x_j + \delta$. Then the ratio between the largest off-diagonal term and the smallest diagonal term of the piecewise linearly interpolated covariance matrix is,

$$\Lambda \equiv \frac{\max_{k\neq k'} |R^o_{k,k'}|}{\min_k |R^o_{k,k}|} \leq \frac{2(\delta^2 + 2\delta h)}{(2M+1)(\delta+h)^2 - 2(\delta^2 + 2\delta h)}.$$

Effect of interpolation on covariance

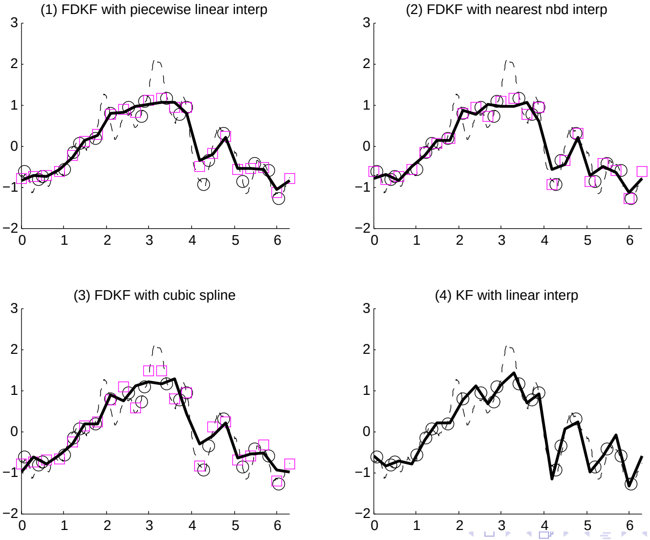


Effect of interpolation on filtered solutions

Table: Weakly irregularly spaced observations: Average RMS errors and spatial correlation for numerical experiments with sparse $2M + 1 = 21$ observations and observation noise error $\sqrt{r^o} = \sqrt{(2M+1)\hat{r}^o} = 0.4583$.

Schemes	RMS error	Spatial corr
1. FDKF with piecewise linear interp	0.3835	0.91
2. FDKF with nearest nbd	0.4417	0.89
3. FDKF with cubic spline	0.4184	0.88
4. Physical space KF with linear interp	0.5136	0.87
5. Coupled FDKF with linear interp	0.4843	0.88
6. Decoupled FDKF with linear interp	0.5089	0.87
7. Coupled FDKF with trig interp	0.4618	0.89
8. Decoupled FDKF with trig interp	0.5010	0.85

Effect of interpolation on filtered solutions: weakly irregularly spaced observations



Effect of interpolation on filtered solutions: weakly irregularly spaced observations

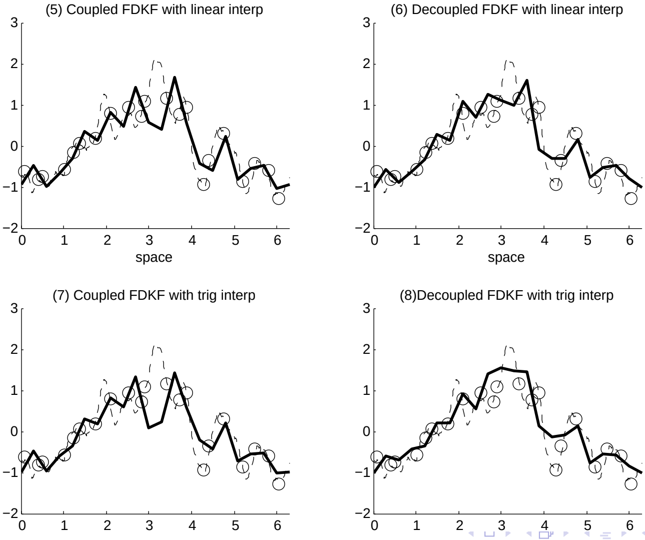
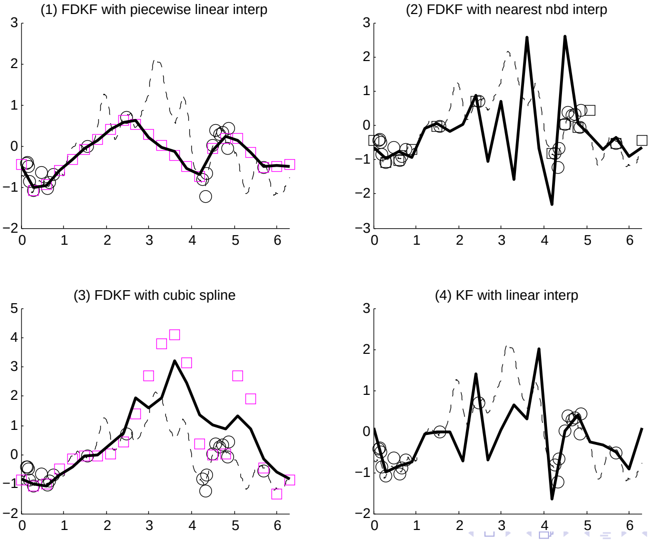


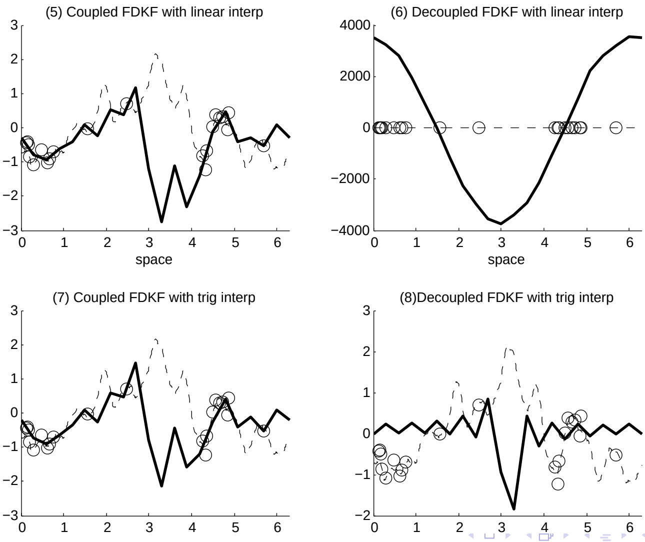
Table: Extremely irregularly spaced and sparse observations: Average RMS errors and spatial correlation for numerical experiments with sparse $2M+1=21$ observations and observation noise error $\sqrt{r^o} = \sqrt{(2M+1)\hat{r}^o} = 0.4583$.

Schemes	RMS error	Spatial corr
1. FDKF with piecewise linear interp	0.6774	0.83
2. FDKF with nearest nbd	1.4507	0.61
3. FDKF with cubic spline	1.0161	0.47
4. Physical space KF with linear interp	1.5488	0.57
5. Coupled FDKF with linear interp	0.9160	0.78
6. Decoupled FDKF with linear interp	3507.9	0
7. Coupled FDKF with trig interp	0.9198	0.77
8. Decoupled FDKF with trig interp	1.7558	0

Effect of interpolation on filtered solutions: extremely irregularly spaced observations



Effect of interpolation on filtered solutions: extremely irregularly spaced observations



Summary:

1. We introduce systematic stochastic parameterization filtering strategy with non-Gaussian statistics that corrects model errors on-the-fly.
2. We study the effects of interpolated observations on data assimilation: recommend lower order interpolation technique relative to higher order one.

References:

- [HM 2010] Harlim and Majda, "Filtering Turbulent Sparsely Observed Geophysical Flows", MWR, 138(4), 1050-1083.
- [BGM 2011] Branicki, Gershgorin, and Majda, "Filtering skill for turbulent signals for a suite of nonlinear and linear extended Kalman filters", submitted to J. Comp. Phys.
- [KMS 2011] Keating, Majda, and Smith, "New methods for estimating poleward eddy heat transport using satellite altimetry", submitted to MWR.
- [KH 2011] Kang and Harlim, "Filtering nonlinear spatio-temporal chaos with autoregressive linear stochastic model, submitted to Physica D.
- [MG 2007] Majda and Grote, "Explicit off-line criteria for stable accurate time filtering of strongly unstable spatially extended systems", Proc. Nat. Acad. Sci, 104:1124-1129, 2007.

Discussions:

1. Reduced Stochastic Filters (MSM, SPEKF). Future consideration: Applying this technique on Three-Cloud Models (Khouider and Majda 2007). How to extend SPEKF to vector valued field? Use the theoretical based MJO model to filter the simulated MJO solutions from the appropriate GCMs.

Discussions:

1. Reduced Stochastic Filters (MSM, SPEKF). Future consideration: Applying this technique on Three-Cloud Models (Khouider and Majda 2007). How to extend SPEKF to vector valued field? Use the theoretical based MJO model to filter the simulated MJO solutions from the appropriate GCMs.
2. Effects of Interpolated Observations on Data Assimilations. Different or more sophisticated statistical interpolators? Extension to two-dimensional field? If we have smaller scale observations, how do we assimilate this data on-the-fly?

Discussions:

1. Reduced Stochastic Filters (MSM, SPEKF). Future consideration: Applying this technique on Three-Cloud Models (Khouider and Majda 2007). How to extend SPEKF to vector valued field? Use the theoretical based MJO model to filter the simulated MJO solutions from the appropriate GCMs.
2. Effects of Interpolated Observations on Data Assimilations. Different or more sophisticated statistical interpolators? Extension to two-dimensional field? If we have smaller scale observations, how do we assimilate this data on-the-fly?
3. Filtering multiscale systems with small-scale intermittency (Macro-Micro-Filtering framework): Future consideration: Apply superparameterization on MMF.