Assimilating Irregularly Spaced Sparsely Observed Turbulent Signals with Hierarchical Bayesian Reduced Stochastic Filters

Kristen Brown

Department of Mathematics
North Carolina State University

Collaborator: John Harlim (NC State)

November 15, 2012
Motivation

- Observations from nature are often noisy, temporally irregular and spatially sparse. On the other hand, the typical predictor model is resolved on regularly spaced grid points.
Motivation

• Observations from nature are often noisy, temporally irregular and spatially sparse. On the other hand, the typical predictor model is resolved on regularly spaced grid points.

• Typically, various interpolation techniques are used to merge multiple satellite data
  For example: cloud clearing algorithm [Chahine, 1973].
Motivation

• Observations from nature are often noisy, temporally irregular and spatially sparse. On the other hand, the typical predictor model is resolved on regularly spaced grid points.

• Typically, various interpolation techniques are used to merge multiple satellite data
  For example: cloud clearing algorithm [Chahine, 1973].

• Processing data is often unavoidable
  For example: estimating rainfall given rain-gauge measurements.
Motivation

• Observations from nature are often noisy, temporally irregular and spatially sparse. On the other hand, the typical predictor model is resolved on regularly spaced grid points.

• Typically, various interpolation techniques are used to merge multiple satellite data For example: cloud clearing algorithm [Chahine, 1973].

• Processing data is often unavoidable For example: estimating rainfall given rain-gauge measurements.

• The goal of this talk is to assess the effect of processed data assimilated in the presence of model error.
Motivation

- Observations from nature are often noisy, temporally irregular and spatially sparse. On the other hand, the typical predictor model is resolved on regularly spaced grid points.

- Typically, various interpolation techniques are used to merge multiple satellite data. For example: cloud clearing algorithm [Chahine, 1973].

- Processing data is often unavoidable. For example: estimating rainfall given rain-gauge measurements.

- The goal of this talk is to assess the effect of processed data assimilated in the presence of model error.

- In particular, we will use hierarchical Bayesian framework.
Standard Bayesian Approach

Definition

- $U$: Random variable of the model state
- $\tilde{V}$: Random variable of the irregularly spaced observations
- $P(U, \tilde{V})$: Joint density between two random variables
Standard Bayesian Approach

Definition
- \( U \): Random variable of the model state
- \( \tilde{V} \): Random variable of the irregularly spaced observations
- \( P(U, \tilde{V}) \): Joint density between two random variables

Canonical Discrete-Time Filtering Problem:

\[
\begin{align*}
  u_{m+1} &= f(u_m) + \sigma_{m+1}, \quad \sigma \sim \mathcal{N}(0, r), \\
  \tilde{v}_m &= g(u_m) + \epsilon_m, \quad \epsilon_m \sim \mathcal{N}(0, r^o)
\end{align*}
\]
Standard Bayesian Approach

Definition

- \( U \): Random variable of the model state
- \( \tilde{V} \): Random variable of the irregularly spaced observations
- \( P(U, \tilde{V}) \): Joint density between two random variables

Canonical Discrete-Time Filtering Problem:

\[
\begin{align*}
    u_{m+1} &= f(u_m) + \sigma_{m+1}, \quad \sigma \sim \mathcal{N}(0, r), \\
    \tilde{v}_m &= g(u_m) + \epsilon_m, \quad \epsilon_m \sim \mathcal{N}(0, r^o)
\end{align*}
\]

Solution: Apply the Bayesian Theorem:

\[
P(U|\tilde{V}) \propto P(U)P(\tilde{V}|U).
\]
Consider $\nu \in V$ to be the random variable of interpolated observations at the regular model grid points. Our approach is to apply

$$P(U|\tilde{V}, V) \propto P(U)P(V|\tilde{V}, U)$$

$$\propto P(U)P(\tilde{V}|U, V)P(V|U).$$
Hierarchical Bayesian Approach

Consider \( v \in V \) to be the random variable of interpolated observations at the regular model grid points. Our approach is to apply

\[
P(U|\tilde{V}, V) \propto P(U)P(V|\tilde{V}, U) \\
\propto P(U)P(\tilde{V}|U, V)P(V|U).
\]

Step 1: We apply \( P(\tilde{V}|U, V)P(V|U) \) through an interpolation to obtain \( P(U)P(V|\tilde{V}, U) \). We compare a statistical interpolation called kriging with a deterministic linear interpolation.
Consider \( v \in V \) to be the random variable of interpolated observations at the regular model grid points. Our approach is to apply

\[
P(U | \tilde{V}, V) \propto P(U) P(V | \tilde{V}, U) \propto P(U) P(\tilde{V} | U, V) P(V | U).
\]

Step 1: We apply \( P(\tilde{V} | U, V) P(V | U) \) through an interpolation to obtain \( P(U) P(V | \tilde{V}, U) \). We compare a statistical interpolation called kriging with a deterministic linear interpolation.

Step 2: We apply \( P(U) P(V | \tilde{V}, U) \) through a reduced stochastic Fourier based filter.
Figure: The 2 layer QG model with baroclinic instability, resolved with 128 × 128 grid points in a 2D periodic domain [Smith et al, 2002]. The radius of deformation is chosen to mimic ocean turbulence. The top panels show the barotropic velocity field (arrows) and streamfunction $\Psi$, (contour) and the bottom panels show the baroclinic velocity field and streamfunction $\Psi$ (bottom) at two different times.
Given two-dimensional noisy, sparse observations from the solution to the two-layer quasi-geostrophic model with baroclinic instability, the first task is to interpolate to a regular $6 \times 6$ grid.
Kriging is a maximum likelihood estimator of a random field $Z$ modeled by

$$Z(s) = \mu(s) + \delta(s),$$

assuming Gaussian, stationary noises $\delta(s) \sim \mathcal{N}(0, C(s, s))$. 

The steps of kriging:
1. Estimate the mean $\mu(s)$ using median polishing. [Cressie, 1993]
2. With the deviations $\delta(s)$ build a parametric covariance function. [Cressie, 1993]
3. Compute the conditional mean and covariance at each grid point using the observations and the parametric covariance function.

We compare ordinary kriging with a deterministic linear interpolation.
Kriging is a maximum likelihood estimator of a random field $Z$ modeled by

$$Z(s) = \mu(s) + \delta(s),$$

assuming Gaussian, stationary noises $\delta(s) \sim \mathcal{N}(0, C(s, s))$.

The steps of kriging:

1. Estimate the mean $\mu$ using median polishing. [Cressie, 1993]
2. With the deviations $\delta$ build a parametric covariance function. [Cressie, 1993]
3. Compute the conditional mean and covariance at each grid point using the observations and the parametric covariance function.
Kriging is a maximum likelihood estimator of a random field $Z$ modeled by

$$Z(s) = \mu(s) + \delta(s),$$

assuming Gaussian, stationary noises $\delta(s) \sim \mathcal{N}(0, C(s, s))$.

The steps of kriging:
1. Estimate the mean $\mu$ using median polishing. [Cressie, 1993]
2. With the deviations $\delta$ build a parametric covariance function. [Cressie, 1993]
3. Compute the conditional mean and covariance at each grid point using the observations and the parametric covariance function.

We compare ordinary kriging with a deterministic linear interpolation.
Spatial Interpolation Results

Figure: The true field (left) and the results of a linear interpolation (middle) and kriging interpolation (right).
Figure: The noise covariance in physical space (top) and Fourier space (bottom).
The next step is to apply a reduced stochastic Fourier based filter. The filter approximates the barotropic modes of the 2 layer QG model

\[
\frac{\partial q}{\partial t} + J(\psi, q) + \beta \frac{\partial \psi}{\partial x} + \kappa \nabla^8 q \\
+ \left[ J(\psi^c, q^c) + U \frac{\partial \nabla^2 \psi^c}{\partial x} - \kappa \nabla^2 \psi^c \right] = 0
\]

in Fourier space with

\[
d\hat{\psi}(t) = (-d + i\omega)\hat{\psi}(t)dt + Fdt + \sigma dW(t).
\]

[Madja and Harlim, Chapter 12, 2012]
The next step is to apply a reduced stochastic Fourier based filter. The filter approximates the barotropic modes of the 2 layer QG model

$$\frac{\partial q}{\partial t} + J(\psi, q) + \beta \frac{\partial \psi}{\partial x} + \kappa \nabla^8 q$$

$$+ \left[ J(\psi^c, q^c) + U \frac{\partial \nabla^2 \psi^c}{\partial x} - \kappa \nabla^2 \psi^c \right] = 0$$

in Fourier space with

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

\(\hat{\psi}\): the horizontal Fourier component of the barotropic streamfunction \(\psi\),

[Madja and Harlim, Chapter 12, 2012]
The next step is to apply a reduced stochastic Fourier based filter. The filter approximates the barotropic modes of the 2 layer QG model

\[
\frac{\partial q}{\partial t} + J(\psi, q) + \beta \frac{\partial \psi}{\partial x} + \kappa \nabla^8 q
+ \left[ J(\psi^c, q^c) + U \frac{\partial \nabla^2 \psi^c}{\partial x} - \kappa \nabla^2 \psi^c \right] = 0
\]

in Fourier space with

\[
d\hat{\psi}(t) = (-d + i\omega)\hat{\psi}(t)dt + Fdt + \sigma dW(t).
\]

\(\hat{\psi}\): the horizontal Fourier component of the barotropic streamfunction \(\psi\),

\(W(t)\): a complex-valued Wiener process,

[Madja and Harlim, Chapter 12, 2012]
The next step is to apply a reduced stochastic Fourier based filter. The filter approximates the barotropic modes of the 2 layer QG model

\[
\frac{\partial q}{\partial t} + J(\Psi, q) + \beta \frac{\partial \Psi}{\partial x} + \kappa \nabla^8 q + \left[ J(\Psi^c, q^c) + U \frac{\partial \nabla^2 \Psi^c}{\partial x} - \kappa \nabla^2 \Psi^c \right] = 0
\]

in Fourier space with

\[
d\hat{\Psi}(t) = (-d + i\omega)\hat{\Psi}(t)dt + Fdt + \sigma dW(t).
\]

\(\hat{\Psi}\): the horizontal Fourier component of the barotropic streamfunction \(\Psi\),
\(W(t)\): a complex-valued Wiener process,
\(d\): damping,
\(\omega\): frequency,
\(F\): constant external forcing,
\(\sigma\): noise strength. [Madja and Harlim, Chapter 12, 2012]
The Kalman Filter is a solution to these equations and produces estimates of the mean and covariance prior and posterior to observation.
**Figure:** The RMS errors associated with each step: unfiltered kriging (dashes), filtered kriging (dashes with ‘+’ sign), unfiltered linear interpolation (solid line), and filtered linear interpolation (solid line with circles).
Filtering Results: $M = 36$ and $r^0 = 17.3$

Figure: Filtering results at one particular time. The circles illustrate observation locations.
Filtering Results: \( M = 36 \) and \( r^0 = 17.3 \)

**Figure:** The true barotropic streamfunction (top), interpolated results (middle panels) and filtered results (bottom panels) at one particular time. The circles illustrate observation locations.
Summary

- In every case, kriging outperformed the linear interpolation.

- Filtering further improved the results.

- However, the biggest improvements occurred in the cases of sparser observations or larger noise.

- The Mean Stochastic Model is a very simple one, and we expect the results could be improved with other models.
• In every case, kriging outperformed the linear interpolation.

• Filtering further improved the results.
• In every case, kriging outperformed the linear interpolation.

• Filtering further improved the results.

• However the biggest improvements occurred in the cases of sparser observations or larger noise.
• In every case, kriging outperformed the linear interpolation.

• Filtering further improved the results.

• However the biggest improvements occurred in the cases of sparser observations or larger noise.

• The Mean Stochastic Model is a very simple one, and we expect the results could be improved with other models.


