Data Assimilation in an Incompressible Viscous Fluid

Simon Cotter

Warwick University, March 20th 2009,

Outline

Motivation

Huge Eulerian And Lagrangian Data Sets

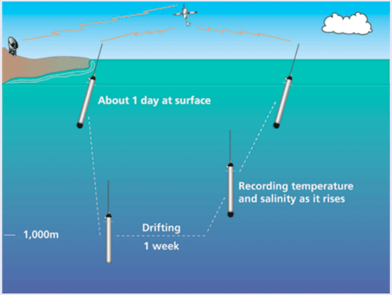
- Large amounts of data
- Blending with sophisticated PDE models will lead to:
- Better weather forecasting
- Better understanding of ocean flows

Motivation

Huge Eulerian And Lagrangian Data Sets

- Large amounts of data
- Blending with sophisticated PDE models will lead to:
- Better weather forecasting
- Better understanding of ocean flows

Example Of Lagrangian Data Collection



General Setup

- Observation operator $\mathcal{G}$ on input to the dynamical system
- Finite dimensional data given by

$$y = \mathcal{G}(u) + \xi, \quad \xi \sim \mathcal{N}(0, \Sigma)$$

- Use MCMC to make inference about u
- u can be
  - $u = v_0$ initial condition of dynamical system
  - $u = (v_0, f)$ initial condition and time-dependent forcing

The (Navier) Stokes Equation

- Stokes operator $\mathcal{A} = -\mathcal{P}\Delta$, where projection $\mathcal{P}$ puts us into an appropriate divergence-free function space
- For $u \in H$, (Navier) Stokes flow can also be given by an ODE on $H = L^2 \cap \{\text{Divergence-free fields}\}$:

$$\begin{aligned} \frac{dv}{dt} + \mathcal{A}v + \gamma B(v,v) &= \mathcal{P}f, \\ v(0) &= v_0 \in H. \end{aligned}$$

- Theory developed for $\gamma = 0, 1$. Numerics for $\gamma = 0$:
    - A fully spectral method can be implemented
    - Code much less computationally expensive
    - Given v_0 and f, solution calculable by FFT. For now let $f \equiv 0$.

The Observation Operators

Eulerian Case: $G_{\mathrm{E}}(v_0)$

Given a set of observation points in space $\{x_j\}_{j=1}^{J} \subset \mathbb{T}^2$, and set of observations times $\{t_k\}_{k=1}^{K}$, then

$$G_{\mathrm{E}}(v_0, f) = \{v(x_j, t_k)\}_{j,k=1}^{J,K},$$

where v is the solution of the Stokes equation with initial condition v_0.

The Observation Operators

Eulerian Case: $\mathcal{G}_{\mathrm{E}}(v_0)$

Given a set of observation points in space $\{x_j\}_{j=1}^{J} \subset \mathbb{T}^2$, and set of observations times $\{t_k\}_{k=1}^{K}$, then

$$\mathcal{G}_{\mathrm{E}}(v_0, f) = \{v(x_j, t_k)\}_{j,k=1}^{J,K},$$

where v is the solution of the Stokes equation with initial condition v_0.

Lagrangian Case: $\mathcal{G}_{\mathrm{L}}(u = v_0)$

Given a set of initial positions for J passive tracers $\{x_j\}_{j=1}^{J} \subset \mathbb{T}^2$, we consider the paths of the tracers governed by the ODEs

$$\begin{aligned} \frac{dz_j}{dt} &= v(z_j(t), t), \quad \forall t > 0, \\ z_j(0) &= x_j, \end{aligned}$$

where v is the solution of the Stokes equation with initial condition v_0. Given a set of observation times $\{t_k\}_{k=1}^{K}$, we define

$$\mathcal{G}_{\mathrm{L}}(v_0, f) = \{z_j(t_k)\}_{j,k=1}^{J,K}.$$

Observational Noise

- As stated before, $y \in \mathbb{R}^{2JK}$ satisfies

$$y = \mathcal{G}(v_0, f) + \xi, \quad \xi \sim \mathcal{N}(0, \Sigma).$$

- Likelihood that y was created with $v(x, 0) = v_0 \in H$:

$$\mathbb{P}(y|v_0) \propto \exp\left(-\frac{1}{2}\|y - \mathcal{G}(v_0, f)\|_\Sigma^2\right).$$

Outline

Bayes' Theorem and the Posterior

- Prior distribution $\mu_0 = \mathbb{P}(v_0)$.
- Likelihood:

$$\mathbb{P}(y|v_0) \propto \exp\left(-\frac{1}{2}\|y - \mathcal{G}(v_0, f)\|_\Sigma^2\right).$$

- Posterior distribution $\mu = \mathbb{P}(v_0|y)$
- Bayes Theorem:

$$\frac{d\mu}{d\mu_0} \propto \exp\left(-\frac{1}{2}\|y - \mathcal{G}(v_0, f)\|_\Sigma^2\right).$$

Prior Measure

- Prior distribution: Gaussian random field

$$v_0 \sim \mathcal{N}(0, \mathcal{A}^{-\alpha}), \quad \alpha \in \mathbb{R}^+.$$

- Construct a sample $v_0 \sim \mathcal{N}(0, \mathcal{A}^{-\alpha})$ via Karhunen-Loeve expansion:

$$v_0 = \sum_{k \in \mathbb{K}} \lambda_k^{-\alpha/2} \phi_k \xi_k, \quad \xi_k \sim \mathcal{N}(0,1) iid,$$

Absolute Continuity of the Posterior

Theorem

Assume that $f \in L^2(0, T, H^r)$ *for some* $r > 0$. *Define a Gaussian measure* μ_0 *on* H*, with mean* $v_b \in H^\alpha$ *and covariance operator* $\mathcal{A}^{-\alpha}$. *If for any* $\alpha > 1$*, then the density*

$$\frac{d\mu}{d\mu_0}(v) \propto \exp\left(-\frac{1}{2}|y - \mathcal{G}_E(v)|_\Sigma^2\right) \quad (1)$$

is μ_0*-a.s. non-zero,* μ_0-*measurable and* μ_0 *integrable. Thus the posterior measure* μ *on* H*, defined by the Radon-Nikodym derivative* (1)*, is absolutely continuous with respect to the prior measure* μ_0. [a]

[a] Data Assimilation Problems In Fluid Mechanics: Bayesian Formulation In Function Space SC,MD,JR,AS

Outline

Metropolis-Hastings in Function Space

- We seek to create a Markov chain of samples from the posterior distribution. Given accepted state of the initial condition u, we propose

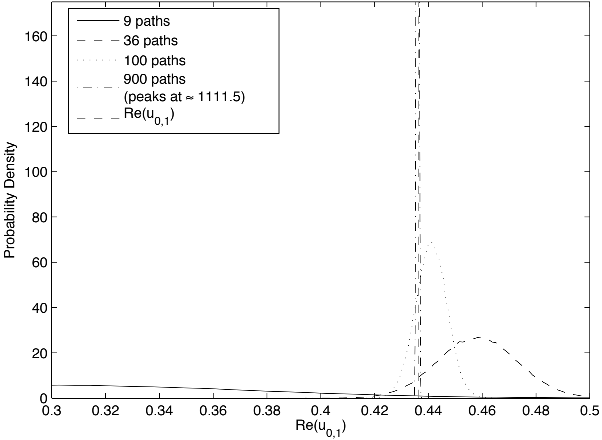
$$v = (1-\beta^2)^{1/2}u + \beta w, \quad w \sim \mu_0 = \mathcal{N}(0, \mathcal{A}^{-\alpha}).$$

- We accept this sample with probability given by Metropolis-Hastings formula:

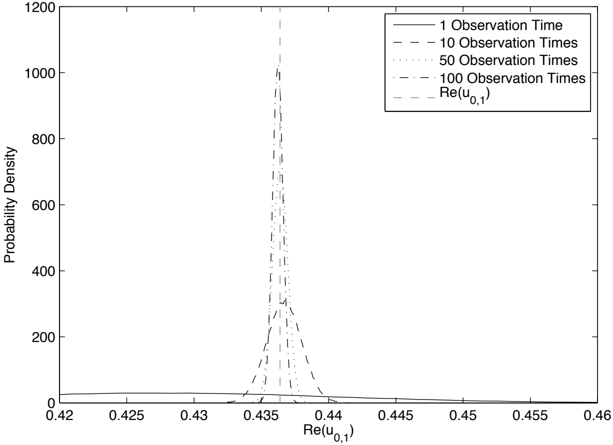
$$a(u,v) = 1 \wedge \exp(\Phi(u) - \Phi(v)), \quad \Phi(\cdot) = \frac{1}{2}\|\mathcal{G}(\cdot) - y\|_{\Sigma}^2.$$

- Acceptance probability independent of β
- Appropriate version of random walk in infinite dimensions with Gaussian priors
- Goldilocks' principle
- Adaptive burn-in to find sensible β

MCMC Results



MCMC Results



Outline

General Setup

- Observation operator $\mathcal{G}$ on initial condition of system *AND* forcing $f(x,t)$
- Data given by

$$y = \mathcal{G}(u = (v_0, f)) + \xi, \quad \xi \sim \mathcal{N}(0, \Sigma)$$

- Use data to infer on $v(0) = v_0$, initial condition of system *AND* forcing or "Model Error" f

Metropolis-Hastings for Model Error

- Very similar proposal:

$$\begin{pmatrix} v \\ g \end{pmatrix} = (1-\beta^2)^{1/2} \begin{pmatrix} u \\ f \end{pmatrix} + \beta \begin{pmatrix} w \\ \psi \end{pmatrix},$$

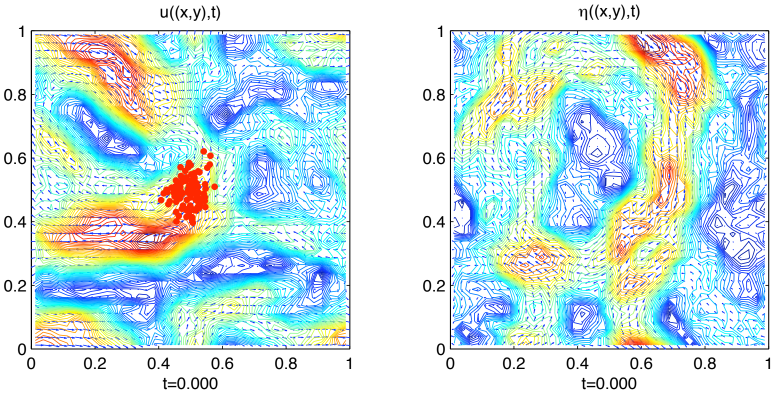
$$\begin{pmatrix} w \\ \psi \end{pmatrix} \sim \mu_0 = \mathcal{N}(0, \mathcal{A}^{-\alpha}) \times \nu_0(\psi)$$

- ν_0 is space-time GRF prior
- Identical acceptance probability with $\Phi(\cdot) = \frac{1}{2}\|\mathcal{G}(\cdot) - y\|_\Sigma^2$:

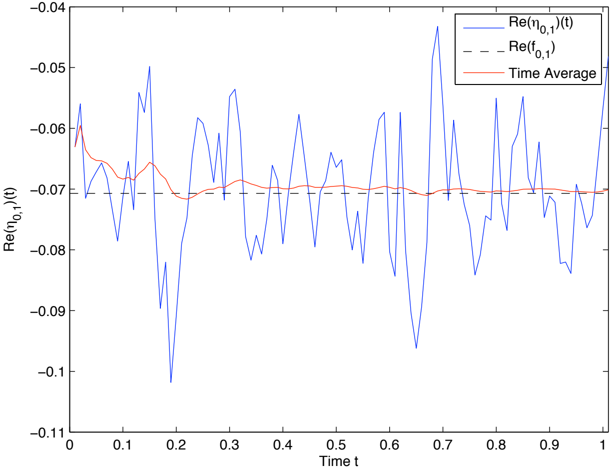
$$a\left(\begin{pmatrix} u \\ f \end{pmatrix}, \begin{pmatrix} v \\ g \end{pmatrix}\right) = 1 \wedge \exp\left(\Phi\begin{pmatrix} u \\ f \end{pmatrix} - \Phi\begin{pmatrix} v \\ g \end{pmatrix}\right).$$

- Need sufficient regularity in the prior to ensure we can make sense of $\mathcal{G}$ with probability 1

What on Earth do these things look like?



Results


−0.04
−0.05
−0.06
−0.07
−0.08
−0.09
−0.1
−0.11
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time t
$Re(\eta_{0,1})(t)$
$Re(\eta_{0,1})(t)$
$Re(f_{0,1})$
Time Average

Eulerian Inference on Model Error

Theorem

Suppose our observation operator returns observations at a sequence of times $0 < t_1 \leq t_2 \leq \ldots \leq t_N \leq T$. *Then given an initial condition and forcing* $(v_0, \eta) \in H \times L^2((0,T), H)$*, there exists an infinite number of alternative* $\eta' \in \times L^2((0,T), H)$ *with* $\eta \neq \eta'$ *almost everywhere, such that*

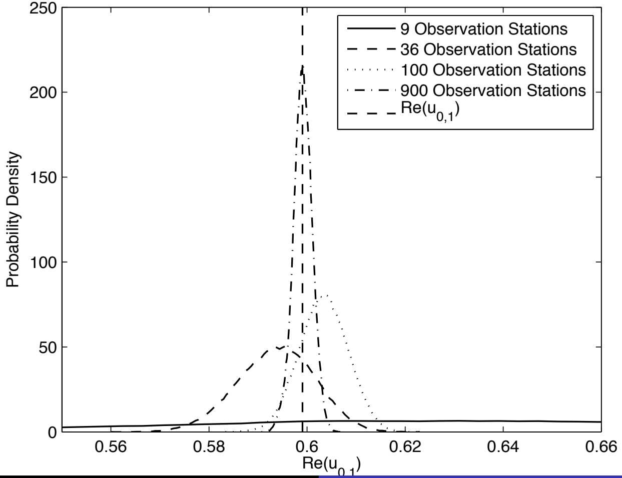
$$\mathcal{G}_E(v_0, \eta') = \mathcal{G}_E(v_0, \eta).$$

- Can only infer on:

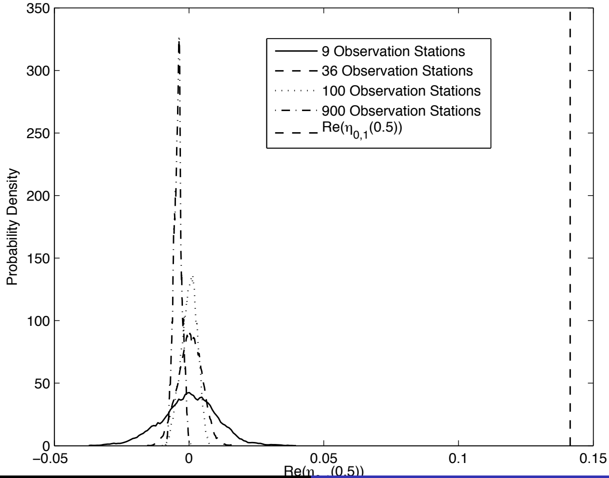
$$F_k(t_j) = \int_0^{t_j} e^{-A(t_j - s)} \eta(s) ds$$

  for each $k \in \{1 \ldots N\}$.
- Not applicable in Lagrangian case

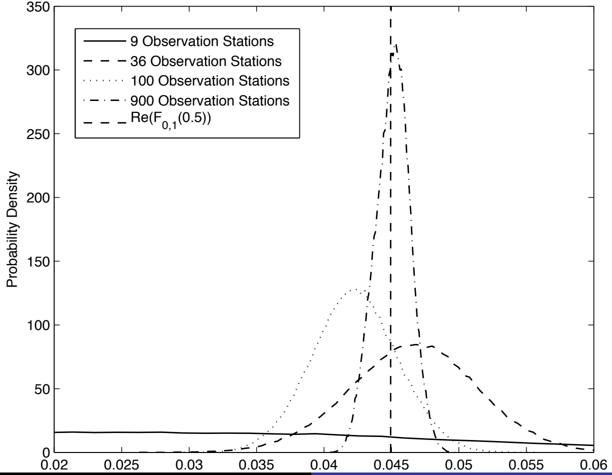
Results


9 Observation Stations
36 Observation Stations
100 Observation Stations
900 Observation Stations
Re(u0,1)
Probability Density
Re(u0,1)
250
200
150
100
50
0
0.56
0.58
0.6
0.62
0.64
0.66

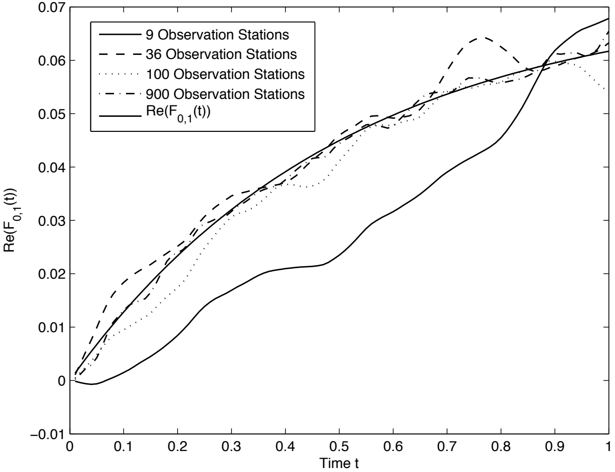
Results


9 Observation Stations
36 Observation Stations
100 Observation Stations
900 Observation Stations
Re($\eta_{0,1}$(0.5))
Probability Density
350
300
250
200
150
100
50
0
−0.05
0
0.05
0.1
0.15

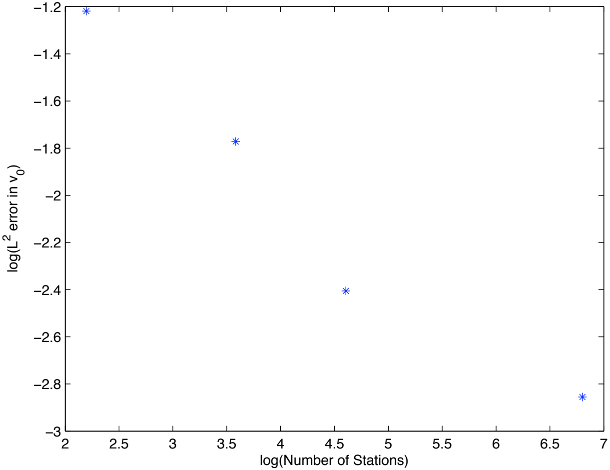
Results


9 Observation Stations
36 Observation Stations
100 Observation Stations
900 Observation Stations
$Re(F_{0,1}(0.5))$
Probability Density
350
300
250
200
150
100
50
0
0.02
0.025
0.03
0.035
0.04
0.045
0.05
0.055
0.06

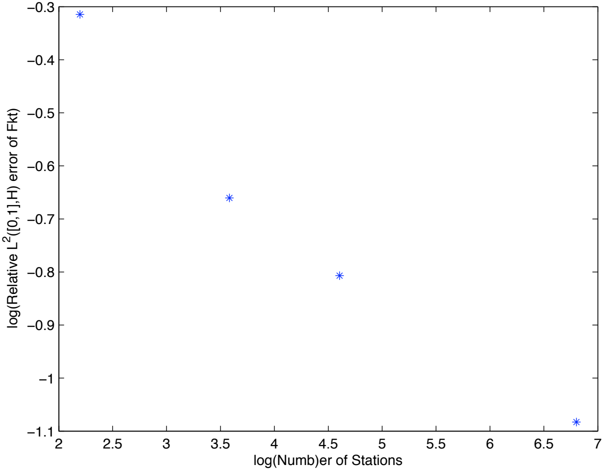
Results


9 Observation Stations
36 Observation Stations
100 Observation Stations
900 Observation Stations
Re(F0,1(t))
Re(F0,1(t))
Time t
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Results


−1.2
−1.4
−1.6
−1.8
−2
−2.2
−2.4
−2.6
−2.8
−3
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
log(L^2 error in v_0)
log(Number of Stations)

Results


−0.3
−0.4
−0.5
−0.6
−0.7
−0.8
−0.9
−1
−1.1
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
log(Relative L2([0,1],H) error of Fkt)
log(Numb)er of Stations

Outline

Summary

- Data assimilation formulated on function space
- This leads to MCMC methods on function space
- These sampling methods are robust under discretization
- We can use data to infer on both the initial condition and forcing of the system
- We require minimum amounts of regularity in the prior distribution to make sense of the observation operator $\mathcal{G}$
- MCMC methods allow us to sample from well-defined posterior distributions

References/Acknowledgements

References

- Data Assimilation Problems In Fluid Mechanics: Bayesian Formulation In Function Space Simon Cotter, Masoumeh Dashti, James Robinson, Andrew Stuart (Submitted to Inverse Problems) available at www.simoncotter.com
- Data Assimilation Problems In Fluid Mechanics: Sampling Methods In Function Space, Simon Cotter, Masoumeh Dashti, James Robinson, Andrew Stuart (In preparation)

References/Acknowledgements

References

- Data Assimilation Problems In Fluid Mechanics: Bayesian Formulation In Function Space Simon Cotter, Masoumeh Dashti, James Robinson, Andrew Stuart (Submitted to Inverse Problems) available at www.simoncotter.com
- Data Assimilation Problems In Fluid Mechanics: Sampling Methods In Function Space, Simon Cotter, Masoumeh Dashti, James Robinson, Andrew Stuart (In preparation)

Acknowledgements

Thanks go to

- Andrew Stuart, Masoumeh Dashti and James Robinson
- Centre for Scientific Computing for use of cluster
- EPSRC for SLC PhD funding