Bayesian Inference for Ocean Diffusivities

Y.K. Ying, James R. Maddison, Jacques Vanneste

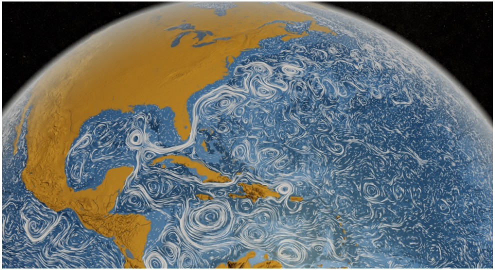
School of Mathematics
The University of Edinburgh

Scales and scaling cascades in geophysical systems,
4th April, 2018

Outline

Background

The turbulent ocean



Eddies in the highly turbulent ocean. Image courtesy: NASA.

Background

What is turbulent diffusion/eddy diffusion?

Simple models for transport in the ocean:

With a 'turbulent' velocity field $\mathbf{u} = \mathbf{u}(x, t), x \in \mathbb{R}^2$:

Passive scalar $c(x, t)$:	Particle position X_t:
$\partial_t c + \mathbf{u}(x, t) \cdot \nabla c = \kappa_m \nabla^2 c,$	$dX_t = \mathbf{u}(X_t, t)dt + \sqrt{2\kappa_m} dW_t,$

- κ_m: small-scale diffusivity.

Background

What is turbulent diffusion/eddy diffusion?

Simple models for transport in the ocean:

With a 'turbulent' velocity field $\mathbf{u} = \mathbf{u}(x,t), x \in \mathbb{R}^2$:

Passive scalar $c(x,t)$:	Particle position X_t:
$\partial_t c + \mathbf{u}(x,t) \cdot \nabla c = \kappa_m \nabla^2 c,$	$dX_t = \mathbf{u}(X_t,t)dt + \sqrt{2\kappa_m}dW_t,$

- κ_m: small-scale diffusivity.

'Averaging'/'Homogenization':

Decompose $\mathbf{u}(x,t) = \mathbf{v}(x) + \mathbf{u}'(x,t)$:

Averaged scalar $\bar{c}(x,t)$:	Particle position $\bar{X}_t$:
$\partial_t \bar{c} + \mathbf{v}(x) \cdot \nabla \bar{c} = \nabla \cdot (\kappa_e(x) \nabla \bar{c}),$	$d\bar{X}_t = \tilde{\mathbf{v}}(\bar{X}_t)dt + \sqrt{2\kappa_e(\bar{X}_t)}dW_t,$

- $\kappa_e(x)$: turbulent/eddy diffusivity tensor field.
- $\tilde{\mathbf{v}}(x) = \mathbf{v}(x) + \nabla \cdot \kappa_e(x)$: 'effective' drift.

Background

Relation between stochastic processes and the advection-diffusion equation

Stochastic Differential Equations

Consider the trajectory of a passive particle X_t at time t:

$$dX_t = [\mathbf{v}(X_t) + \nabla \cdot \kappa_e(X_t)]\, dt + \sqrt{2\kappa_e(X_t)}dW_t,$$

Fokker-Planck Equations

The transition probability density $p(x, t|x_0)$ of a passive particle:

$$\begin{aligned} \partial_t p + \mathbf{v}(x) \cdot \nabla p &= \nabla \cdot (\kappa_e \nabla p), \\ p(x, 0|x_0) &= \delta(x - x_0). \quad \text{(Initial condition)} \end{aligned}$$

Interpretation:

The transition density of an SDE is a passive scalar with an initial Dirac-Delta profile

Description

How to estimate the turbulent diffusivity $\kappa(x)$ in the ocean?

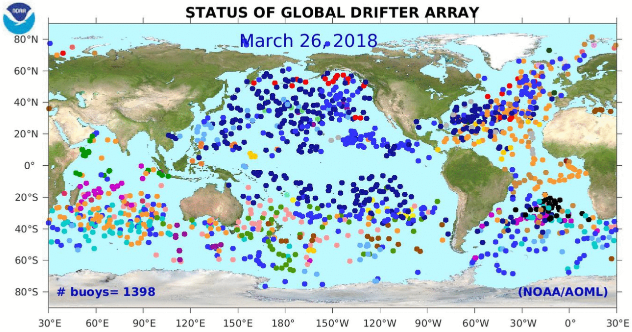


Figure: Some available drifters across the globe.

[Image retrieved from website of *The Global Drifter Program*]

Description

How to estimate the turbulent diffusivity $\kappa(x)$ in the ocean?

Inference

Given the position data $\{X_t\}$, at discrete time $t = t_n$ from the ocean drifters:

1. Seek $\mathbf{u}(x)$ and $\kappa(x)$ of

$$dX_t = \mathbf{u}(X_t)dt + \sqrt{2\kappa(X_t)}dW_t$$

   that 'fits' with the observation data.

2. Quantify the **uncertainty** of the inferred $\mathbf{u}(x)$ and $\kappa(x)$

Methodology: How to diagnose eddy diffusivity?

Conventional Approach:

Given the eddy velocities $\{u'(t)\}$ of drifters,

integrate the velocity autocorrelation function

$$\kappa = \int_0^\infty \left\langle u'(t)u'(t+\tau)\right\rangle d\tau$$

$\Rightarrow$ Obtain one estimate for κ

Methodology: How to diagnose eddy diffusivity?

Conventional Approach:

Given the eddy velocities $\{u'(t)\}$ of drifters,

integrate the velocity autocorrelation function

$$\kappa = \int_0^\infty \left\langle u'(t)u'(t+\tau)\right\rangle d\tau$$

$\Rightarrow$ Obtain one estimate for κ

Bayesian SDE Inference Approach:

Given sequences of positions $\{X_t\}$ at $t = t_n$, $n = 0, 1, \ldots N$,

formulate the posterior distribution

$$p(\mathbf{u}, \kappa | \{X_t\})$$

$\Rightarrow$ Obtain a probability distribution on $\mathbf{u}$ and κ, conditioned on the available data $\{X_t\}$

Bayesian Inference

Formulation

Bayes' Theorem

RECALL: Probability 101:

$$P(A|B)P(B) = P(B|A)P(A)$$
$$\Rightarrow P(A|B) \propto P(B|A)P(A)$$

Bayesian Inference

Formulation

Bayes' Theorem

RECALL: Probability 101:

$$P(A|B)P(B) = P(B|A)P(A)$$
$$\Rightarrow P(A|B) \propto P(B|A)P(A)$$

TARGET: Posterior distribution density $p(\mathbf{u}, \kappa| \{X_t\})$:

$$p(\mathbf{u}, \kappa| \{X_t\}) \propto p(\{X_t\} |\mathbf{u}, \kappa)p(\mathbf{u}, \kappa)$$

Fill in the gaps:

1. Proportional factor $\Rightarrow$ normalisation constant for a p.d.f.
2. *Likelihood* $p(\{X_t\} |\mathbf{u}, \kappa)$ $\Rightarrow$ needs to be computed
3. *Prior* $p(\mathbf{u}, \kappa)$ $\Rightarrow$ subjectively chosen

Bayesian Inference

Formulation

Bayes' Theorem

RECALL: Probability 101:

$$P(A|B)P(B) = P(B|A)P(A)$$
$$\Rightarrow P(A|B) \propto P(B|A)P(A)$$

TARGET: Posterior distribution density $p(\mathbf{u}, \kappa | \{X_t\})$:

$$p(\mathbf{u}, \kappa | \{X_t\}) \propto p(\{X_t\} | \mathbf{u}, \kappa) p(\mathbf{u}, \kappa)$$

Fill in the gaps:

1. Proportional factor $\Rightarrow$ normalisation constant for a p.d.f.
2. *Likelihood* $p(\{X_t\} | \mathbf{u}, \kappa) \Rightarrow$ needs to be computed
3. *Prior* $p(\mathbf{u}, \kappa) \Rightarrow$ subjectively chosen

Bayesian Inference

Computational Aspects

1. How to construct the likelihood $p(\{X_t\}|\mathbf{u}, \kappa)$?

$=$ Probability of reproducing the data $\{X_t\}$, given $(\mathbf{u}, \kappa)$

Evaluation:

1. Pair up the observed positions $\{X_t\}$ consecutively

$$\{X_0, X_1\}, \{X_1, X_2\}, \ldots, \{X_{n-1}, X_n\},$$

   where $h = t_{i+1} - t_i$ is the duration between observation
2. Evaluate transition density $p(X_{i+1}, h|X_i)$
3. Likelihood $p(\{X_t\}|(\mathbf{u}, \kappa)) = \prod_i p(X_{i+1}, h|X_i)$

Bayesian Inference

Computational Aspects

1. How to construct the likelihood $p(\{X_t\}\,|\mathbf{u},\kappa)$?

$=$ Probability of reproducing the data $\{X_t\}$, given $(\mathbf{u},\kappa)$

Evaluation:

1. Pair up the observed positions $\{X_t\}$ consecutively

$$\{X_0, X_1\}, \{X_1, X_2\}, \ldots, \{X_{n-1}, X_n\},$$

   where $h = t_{i+1} - t_i$ is the duration between observation
2. Evaluate transition density $p(X_{i+1}, h|X_i)$
3. Likelihood $p(\{X_t\}\,|(\mathbf{u},\kappa)) = \prod_i p(X_{i+1}, h|X_i)$

Bayesian Inference

Formulation

Bayes' Theorem

RECALL: Probability 101:

$$\begin{aligned} P(A|B)P(B) &= P(B|A)P(A) \\ \Rightarrow P(A|B) &\propto P(B|A)P(A) \end{aligned}$$

TARGET: Posterior distribution density $p(\mathbf{u}, \kappa | \{X_t\})$:

$$p(\mathbf{u}, \kappa | \{X_t\}) \propto p(\{X_t\} | \mathbf{u}, \kappa) p(\mathbf{u}, \kappa)$$

Fill in the gaps:

1. Proportional factor $\Rightarrow$ normalisation constant for a p.d.f.
2. *Likelihood* $p(\{X_t\} | \mathbf{u}, \kappa) \Rightarrow$ needs to be computed
3. *Prior* $p(\mathbf{u}, \kappa) \Rightarrow$ subjectively chosen

Bayesian Inference

Computational Aspects

2. How to take Prior $p(\mathbf{u}, \kappa)$?

$=$ Preference for $(\mathbf{u}, \kappa)$ from experts' opinion

Choice of prior:

$$(\mathbf{u}, \kappa) \sim \text{Uniform distribution on a given range}$$
$$\Rightarrow \text{No preference}$$

Bayesian Inference Machinery

Sampling

What to do with the posterior density $p(\mathbf{u}, \kappa | \{X_t\})$?

Sampling by Metropolis-Hasting Algorithm

$\iff$ produce a histogram

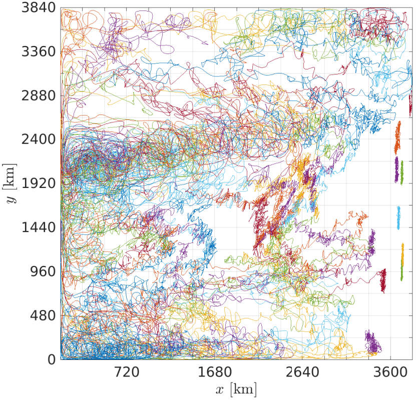
Inference from the posterior density $p(\mathbf{u}, \kappa | \{X_t\})$

1. Estimate $(\mathbf{u}, \kappa)$ via
   - Posterior mean
   - Maximum a posteriori (MAP) estimates
2. Quantify uncertainties via
   - Posterior variance
   - Local modes

[Lost? Examples coming up in a few slides!]

Particles in an Ideal Ocean Model

- *Data input:* Daily positions of 1024 passive particles over 10 years using the quasi-geostrophic equations (similar to *Karabasov et al, Ocean Modelling 2009*)
- 'Spaghetti' plot of trajectories of 64 particles: Click me



Particles in an Ideal Ocean Model

Inference model

1. *Resolution*: Partition the domain into bins of size 240km $\times$ 240km
2. *Fields*: For each bin, impose a velocity and diffusivity field

$$\mathbf{u}(x,y) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11} \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

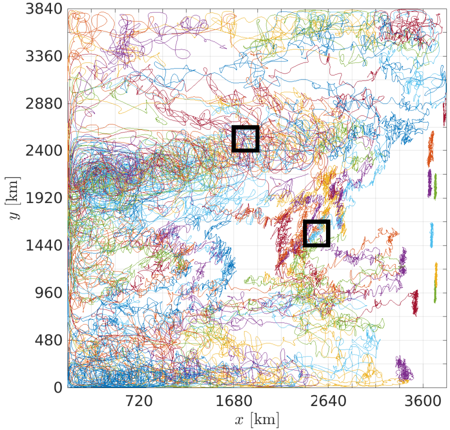
$$\kappa(x,y) = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix}.$$

3. *Locality*: Group the consecutive positions of particles $\{X_i, X_{i+1}\}$ into bins based on their starting points $\{X_i\}$

Passive Particles in an Ideal Ocean Model

Focus on two cells: jet vs not on jet

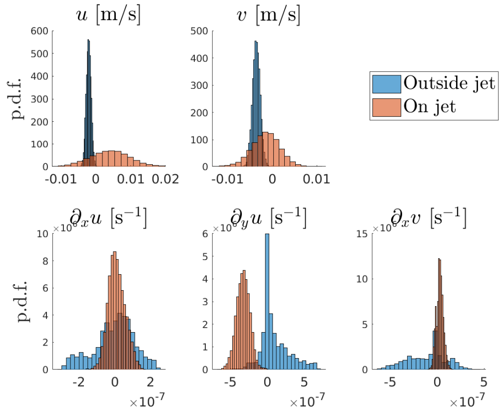
- ☐ Boxes: the chosen two cells:



Passive Particles in an Ideal Ocean Model

Posterior Distributions in two selected cells, sampling interval $h = 64$ days

Velocity fields $u(x, y), v(x, y)$ at cell centres
$\Rightarrow$ faster flow, but more uncertain, on the jet.

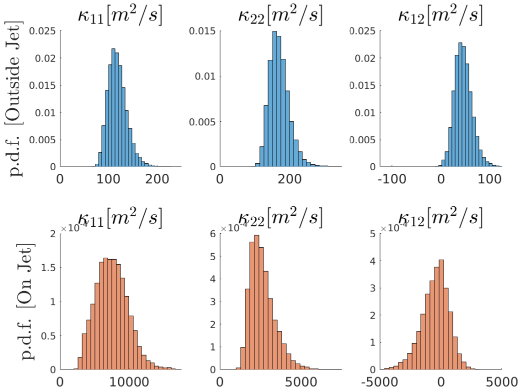


Particles in an Ideal Ocean Model

Posterior Distributions in two selected cells, sampling interval $h = 64$ days

Diffusivity $\kappa = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix}$ at cell centres

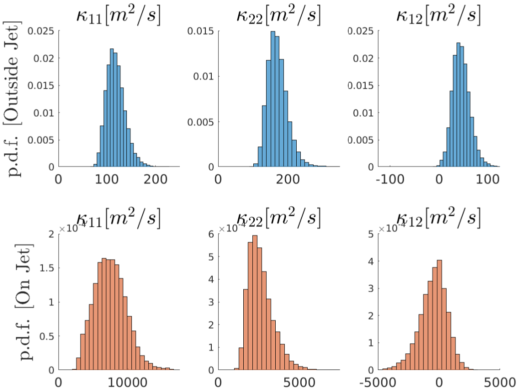
$\Rightarrow$ much larger and much more uncertain diffusivity on the jet.



Particles in an Ideal Ocean Model

A representative number to simplify the analysis?

Locate the mode of the probability distribution ⇒ MAP estimates



Particles in an Ideal Ocean Model

MAP estimates for diffusivity field (Sampling interval $h = 64$ days)

Diagonal components of the diffusivity tensor κ_{11} and κ_{22}:

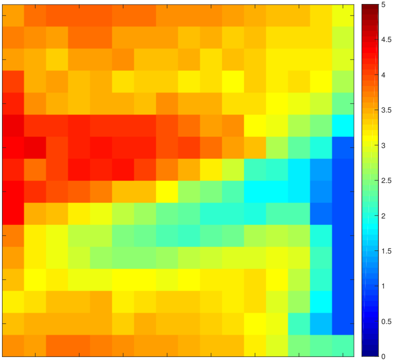


Figure: κ_{11} [m^2/s] on a log10-scale

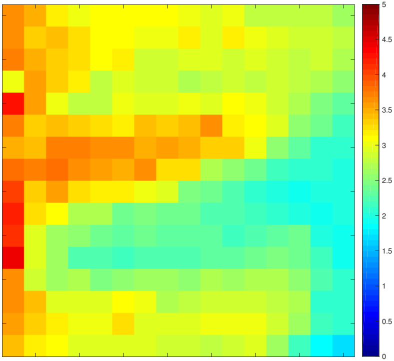


Figure: κ_{22} [m^2/s] on a log10-scale

Take Home Messages

- A Bayesian Inference for Ocean Diffusivities
  - A **probabilistic** inference for $\mathbf{u}$ and κ
  - **Uncertainty quantification** with ease
  - No a priori decomposition of velocity fields required
  - Capable of inferring an anisotropic diffusivity tensor κ
- A physically meaningful diffusivity κ
  - Robust inference: Convergence over sampling interval h
- Importance in resolving the spatial variation in velocity field $\mathbf{u}(x, y)$
  - Robust inference: Convergence requires sufficient resolution of velocity fields

Future work

- Comparisons with existing diffusivity diagnostics
  - Lagrangian methods
  - Passive tracer methods
- Relax the locality assumptions
  - Ensure globally smooth fields are inferred
  - Make diffusivity diagnosis possible in the fast regions
- Inference using real drifters!

Particles in an Ideal Ocean Model

MAP estimates for diffusivity field (Sampling interval $h = 64$ days)

Diagonal components of the diffusivity tensor κ_{11} and κ_{22}:

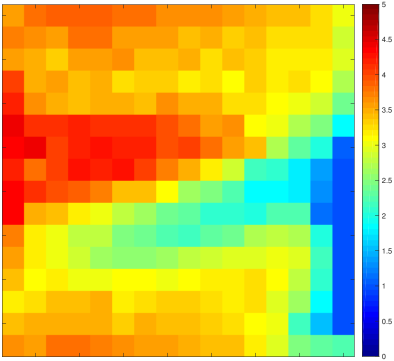


Figure: κ_{11} [m^2/s] on a log10-scale

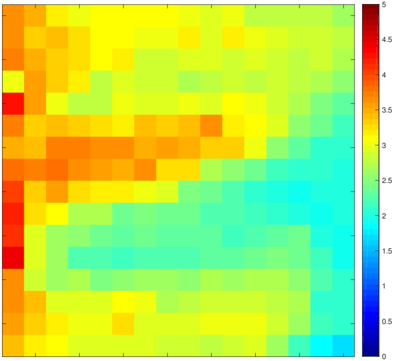


Figure: κ_{22} [m^2/s] on a log10-scale

Particles in an Ideal Ocean Model

MAP estimates for diffusivity field (Sampling interval $h = 1$ days)

Diagonal components of the diffusivity tensor κ_{11} and κ_{22}:

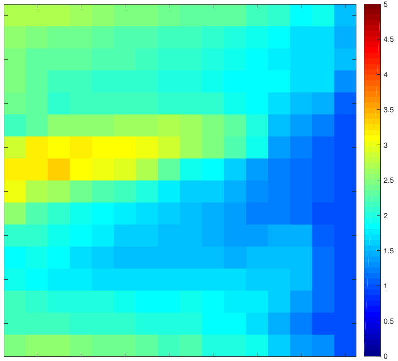


Figure: κ_{11} [m^2/s] on a log10-scale

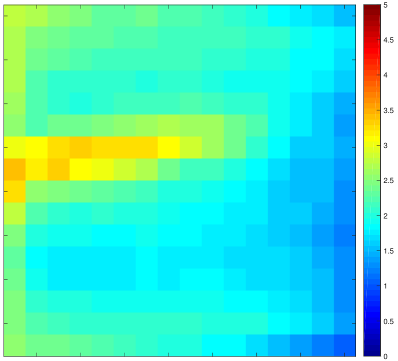


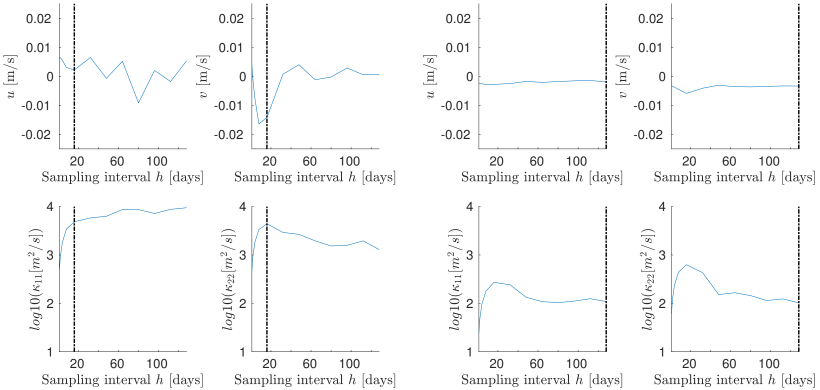
Figure: κ_{22} [m^2/s] on a log10-scale

Particles in an Ideal Ocean Model

What sampling interval h to use?

MAP Fields at cell centres vs sampling interval h

[Left: On the jet; Right: outside the jet]

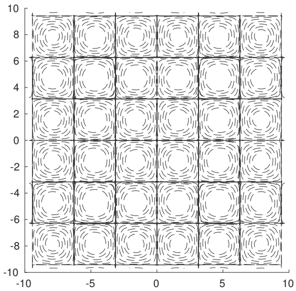


Test Case - Taylor-Green Vortex

Particle advection:

$$dX_t = \mathbf{u}^{TG}(X_t)dt + \sqrt{2\kappa_m}dW_t, \quad \text{where } \mathbf{u}^{TG}(x,y) = \begin{pmatrix} \sin(x)\cos(y) \\ -\cos(x)\sin(y) \end{pmatrix},$$

Figure: Streamlines of Taylor-Green Vortex; Video at [Click me]



Test Case - Taylor-Green Vortex

- For t sufficiently large, coarse-graining the SDE

$$dX_t = \mathbf{u}^{TG}(X_t)dt + \sqrt{2\kappa_m}dW_t \rightarrow dX_t{}^c = \sqrt{2\kappa}dW_t.$$

- Simulated Trajectories
  1. Number of Particles: 1024
  2. Initial conditions: Located at the origin
  3. $\kappa_m = 0.1 \rightarrow \kappa \approx 0.342I$

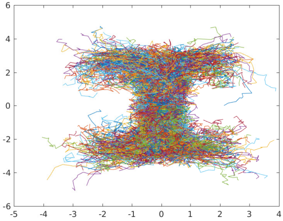


Figure: Simulated Trajectories, up to $t = 50$

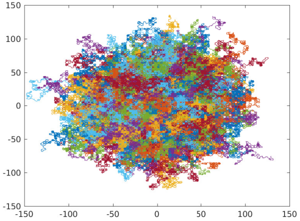


Figure: Simulated Trajectories, up to $t = 2500$

Test Case - Taylor-Green Vortex

- Bayesian Inference: Impose the coarse-grained model

$$dX_t = \begin{pmatrix} u \\ v \end{pmatrix} dt + \sqrt{2 \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix}} dW_t$$

- Plots of mean quantities over sampling interval h

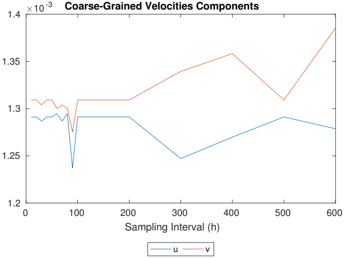


Figure: Mean velocity (u, v) against sampling interval h

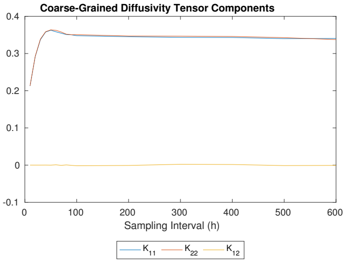


Figure: Mean diffusivity K against sampling interval h