Bayesian inference of ocean diffusivity from Lagrangian trajectory data

Y-K Ying, J R Maddison & J Vanneste

School of Mathematics and Maxwell Institute
University of Edinburgh, UK
www.maths.ed.ac.uk/˜vanneste

Motivation

Coarse-graining turbulent transport

The scalar concentration $C(\boldsymbol{x}, t)$ of passive scalars obeys the advection–diffusion equation

$$\partial_t C + \boldsymbol{u} \cdot \nabla C = \kappa \nabla^2 C \,,$$

with $\nabla \cdot \boldsymbol{u} = 0$ and molecular diffusivity κ.

The velocity field $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t)$ is

- turbulent, complex, multiscale,
- generates fine scales in $C(\boldsymbol{x}, t)$,
- only partially resolved in GCMs.

Need to coarse grain the advection–diffusion equation, so that large-scale features of $C(\boldsymbol{x}, t)$ are well reproduced.

Motivation

Coarse-graining turbulent transport

Coarse-grained model: advection–difffusion equation

$$\partial_t C + \boldsymbol{U} \cdot \nabla C = \nabla \cdot (K \nabla C),$$

with

- $\boldsymbol{U} = \boldsymbol{U}(\boldsymbol{x})$ an averaged velocity,
- $K = K(\boldsymbol{x})$ an eddy diffusivity tensor.

Justified by:

- simplicity,
- asymptotic results (e.g. for $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t/\epsilon)$),
- central limit theorem.

Applies for $t \gg$ correlation time of $\boldsymbol{u}$.

How to estimate $\boldsymbol{U}$ and K?

Inferring $\boldsymbol{U}$ and K

Using numerical models or satellite-altimetry-based $\boldsymbol{u}$,

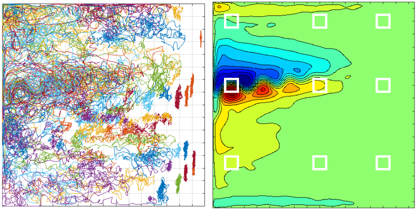
- $\boldsymbol{U}$ as (Eulerian) time average, $\boldsymbol{U} = \overline{\boldsymbol{u}}$,
- K from flux-gradient relations in numerical experiments:
  $\overline{\boldsymbol{u}'C'} = K\nabla\bar{C}$, eg Abernathey et al 2013
- effective diffusivity. Nakamura 1996, Haynes & Shuckburgh 2000

Want to exploit Lagrangian trajectory data (e.g. Global Drifter Program, Argo. . .):

$$\dot{\boldsymbol{X}} = \boldsymbol{u}(\boldsymbol{X}, t) + \sqrt{2\kappa}\dot{\boldsymbol{W}}$$

coarse-grained as

$$\dot{\boldsymbol{X}} = \boldsymbol{U}(\boldsymbol{X}) + \sqrt{2K(\boldsymbol{X})} \diamond \dot{\boldsymbol{W}}.$$



Infer $\boldsymbol{U}$ and K from set of trajectories $\{\boldsymbol{X}^{(i)}(t_j),\ i, j = 1, 2, \cdots\}$.

Inferring $\boldsymbol{U}$ and K

Standard approach: consider particles starting near $\boldsymbol{x}$:

- $\boldsymbol{U}(\boldsymbol{x})$ as average velocity,

$$\boldsymbol{U}(\boldsymbol{x}) = \frac{1}{t}\langle \boldsymbol{X}^{(i)}(t) - \boldsymbol{X}^{(i)}(0) \,|\, \boldsymbol{X}^{(i)}(0) \approx \boldsymbol{x}\rangle,$$

- $K(\boldsymbol{x})$ from covariance,

$$K(\boldsymbol{x}) = \frac{1}{2t}\langle \Big(\boldsymbol{X}'^{(i)}(t) - \boldsymbol{X}'^{(i)}(0)\Big) \otimes \Big(\boldsymbol{X}'^{(i)}(t) - \boldsymbol{X}'^{(i)}(0)\Big) \,|\, \boldsymbol{X}^{(i)}(0) \approx \boldsymbol{x}\rangle,$$

- or from the velocity correlation

$$K(\boldsymbol{x}) = \frac{1}{2}\langle \int_0^t \boldsymbol{u}'^{(i)}(s)\boldsymbol{u}'^{(i)}(0)\,\mathrm{d}s \,|\, \boldsymbol{X}^{(i)}(0) \approx \boldsymbol{x}\rangle,$$

Taylor 1922, Davies 1991, Griesel et al 2010

Assumption of locality, $K(\boldsymbol{X}_i(t)) \approx \text{const}$, problematic since $t \gg T_{\text{corr}}$ needed.

Bayesian inference

The correct approach to infer any parameter. Here,

$$\boldsymbol{\theta} = (\boldsymbol{U}, K),$$

discretised in applications, to be inferred from a single trajectory $\{\boldsymbol{X}_i = \boldsymbol{X}(t_i),\ i = 1, 2, \cdots\}$.

Bayes's formula:

$$p(\boldsymbol{\theta}|\{\boldsymbol{X}_i\}) \propto p(\{\boldsymbol{X}_i\}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}).$$

- $p(\boldsymbol{\theta}|\{\boldsymbol{X}_i\})$: posterior, pdf of values of $\boldsymbol{\theta}$ given data,
- $p(\{\boldsymbol{X}_i\}|\boldsymbol{\theta})$: likelihood, solution of the advection–diffusion equation,
- $p(\boldsymbol{\theta})$: prior.

Bayesian inference

We can compute the likelihood:

$$p(\{\boldsymbol{X}_i\}|\boldsymbol{\theta}) = \prod_i \phi_{\boldsymbol{\theta}}(\boldsymbol{X}_i, h|\boldsymbol{X}_{i-1}), \quad h = t_i - t_{i-1}$$

where $\phi_{\boldsymbol{\theta}}(\boldsymbol{x}, t|\boldsymbol{y})$ is the fundamental solution of the advection–diffusion equation,

$$\partial_t \phi_{\boldsymbol{\theta}} + \boldsymbol{U} \cdot \nabla \phi_{\boldsymbol{\theta}} = \nabla \cdot (K \nabla \phi_{\boldsymbol{\theta}}), \qquad \phi_{\boldsymbol{\theta}}(x, 0|y) = \delta(x - y).$$

Problem solved?

Bayesian inference

Challenge:

- $p(\boldsymbol{\theta}|\{\boldsymbol{X}_i\})$ is a pdf in high dimension:
  for 16×16 grid, piece-uniform $(\boldsymbol{U}, K)$, $\dim \boldsymbol{\theta} = 1280$,
- sample from $p(\boldsymbol{\theta}|\{\boldsymbol{X}_i\})$ to estimate

$$\int f(\boldsymbol{\theta}) p(\boldsymbol{\theta}|\{\boldsymbol{X}_i\})\, \mathrm{d}\boldsymbol{\theta}, \quad \text{eg} \ \langle K(\boldsymbol{x})\rangle = \int K(\boldsymbol{x}) p(\boldsymbol{\theta}|\{\boldsymbol{X}_i\})\, \mathrm{d}\boldsymbol{\theta},$$

  using Markov Chain Monte Carlo: $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_{10^6}$,
- require $\phi_{\boldsymbol{\theta}_j}(X_i, h|X_{i-1})$ for each t_i and each $\boldsymbol{\theta}_j$,
- $10^5 \times 10^6 = 10^{11}$ numerical solutions of the advection–diffusion equation.

Infeasible.

Bayesian inference

Two methods to make the inference tractable:

1. Locality assumption: with $(\boldsymbol{U}, K)$ taken as constant along particle trajectories, closed form for $\phi_{\boldsymbol{\theta}}(\boldsymbol{x}, t|\boldsymbol{y})$,

   Ying, Maddison & V 2019

2. Coarse-grained inference: replace exact starting position by starting cell,

$$\phi_{\boldsymbol{\theta}}(X_i, h|X_{i-1}) \quad \rightsquigarrow \quad \bar{\phi}_{\boldsymbol{\theta}}(X_i, h|B_k) = |B_k|^{-1} \int_{B_k} \phi_{\boldsymbol{\theta}}(X_i, h|\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}.$$

   Solve 1 advection-diffusion for all $X_{i-1} \in B_k$:

$$\partial_t \bar{\phi}_{\boldsymbol{\theta}} + \boldsymbol{U} \cdot \nabla \bar{\phi}_{\boldsymbol{\theta}} = \nabla \cdot \left(K \nabla \bar{\phi}_{\boldsymbol{\theta}} \right), \quad \bar{\phi}_{\boldsymbol{\theta}}(x, 0|B_k) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in B_k \\ 0 & \text{if } \boldsymbol{x} \notin B_k \end{cases}.$$

Bayesian inference

Coarse-grained inference

For $N \to \infty$ data points, drawn from exact model with $\boldsymbol{\theta} = \boldsymbol{\theta}_*$,

$$p(\{X_i\}|\boldsymbol{\theta}) \asymp \mathrm{e}^{-ND(\theta)}, \quad \bar{p}(\{X_i\}|\boldsymbol{\theta}) \asymp \mathrm{e}^{-N\bar{D}(\theta)},$$

with

- relative entropies (Kullback–Leibler divergences)

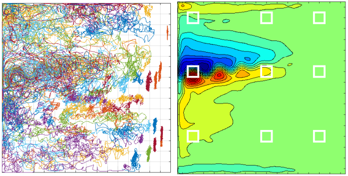
$$D(\boldsymbol{\theta}) = \int \phi_{\boldsymbol{\theta}_*} \log \frac{\phi_{\boldsymbol{\theta}}}{\phi_{\boldsymbol{\theta}_*}} \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}, \quad \bar{D}(\boldsymbol{\theta}) = \sum_k \int \bar{\phi}_{\boldsymbol{\theta}_*} \log \frac{\bar{\phi}_{\boldsymbol{\theta}}}{\bar{\phi}_{\boldsymbol{\theta}_*}} \,\mathrm{d}\boldsymbol{x},$$

- minimised for $\boldsymbol{\theta} = \boldsymbol{\theta}_*$: Maximum A Posteriori (MAP) estimate is exact,
- loss of information, $D(\boldsymbol{\theta}) \geq \bar{D}(\boldsymbol{\theta})$.

QG flow

Application to simulated trajectories:

- 3-layer quasi-geostrophic model,
- wind-driven, Atlantic-like,
- 513^2 resolution.

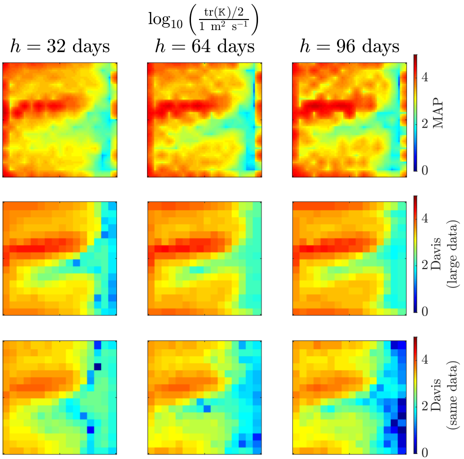


Data & MCMC

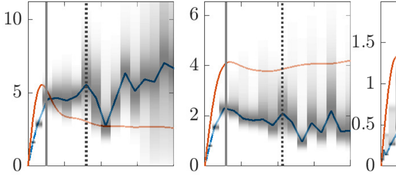
- 10-year trajectories of 676 drifters,
- exact numerical integration for approximate velocity field,
- 16×16 cells B_k,
- 283,920 MCMC samples,
- standard Metropolis–Hastings–Gibbs sampler,
- advection–diffusion solved by finite volume.

QG flow

Diffusivity: MAP estimate vs sampling time h

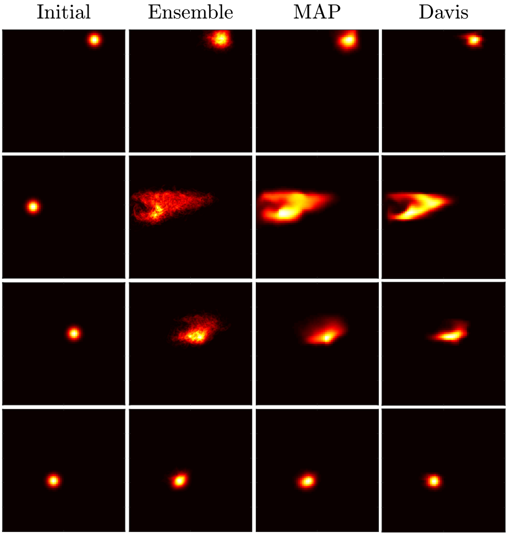


With locality assumption



QG flow

Comparing coarse-grained with high-resolution advection–diffusion



Conclusion

Bayesian inference of mean velocity and diffusivity

- clear interpretation of $\boldsymbol{U}$ and K,
- made computationally tractable using coarse-grained inference,
- removes the need for locality assumption: no need that $h \ll U/\text{cell length}$, just $h \ll T_{\text{corr}}$,
- performs well with moderate data volume,
- gives $p(\boldsymbol{U}, K|\{X_i\})$, uncertainty estimate.

Future work

- computational challenge: increase spatial resolution, improve MCMC sampler, exploit localisation,
- divergent flow, non-uniform distribution of drifters,
- application to real drifters.