Dimension Reduction, Stochastic Parametrization and Data Assimilation for Transport in the Ocean

Clint Dawson
Juan Restrepo
Steven Rosenthal
Shankar Venkataramani
Oil in the environment

- Evaporation
- Photo-oxidation
- Spreading
- Dispersion
- Dissolution
- Sedimentation
- Emulsification
- Biodegradation
Tracking
Feature based (Lagrangian) data assimilation

(a) (b)

(c) (d)

Classic

Contour Analysis
Canonical transformations for displacement assimilation

Area preserving flows correspond to symplectic maps.

Find $M$ such that

$$\min ||q(M(x)) - q_0||_2^2.$$

here $(x, y)^M \rightarrow (X, Y)$.

In 2-Dimensions, the generating function is $G(X, y) = Xy + f(X, y)$.

$$x = \frac{\partial G}{\partial y} = X + f_y(X, y)$$

$$Y = \frac{\partial G}{\partial X} = y + f_x(X, y).$$

invertible if $f_{yx} > -1$.

Regularize!
Displacement map
Displacement assimilation: Ensemble Kalman filter
Transport
Nearshore sticky waters

A red tide event, off the coast of Florida. The event occurs nearly annually along the state’s Gulf Coast.

Image courtesy of P. Schmidt, Charlotte Sun.
Model I

- A thin oil slick sits atop the ocean.
- The ocean’s mixed layer of thickness $P$ is laden with oil droplets.
- $x$ is the distance from the shore.
- The break zone extends to $x = L$.
- The ocean surface is at $z = 0$ and bottom topography is fixed and described by $z = -H(x)$. 
Model II

- Bottom topography:

\[ H(x) = H_0 + mx, \quad 0 \leq x \leq X, \]

- Mixed layer thickness \( \xi(x) \approx \min(H(x), P) \).

- Bulk oil concentration \( B(x, t) \). Equivalent thickness of oil layer \( b(x, t) = B(x, t)\xi(x) \).

**Parabolic profile for mean flow:**

\[ U = U^{St} \left( 1 + \frac{4z}{H(x)} + \frac{3z^2}{H(x)^2} \right) \]

**Bulk flux:**

\[ \Gamma_B = b(x, t)U^{St} \left[ 1 - \frac{\xi(x)}{H(x)} \right]^2 . \]
Model III

Diffusivity:

\[
D(x) = D_{eddy} + S(x)D_L.
\]

\[
S(x) = \frac{1}{1+\exp[(x-L)/w]}
\]

Surface/bulk mixing

\[
\text{Flux: } \Gamma_{SB} = \frac{1}{\tau(x)} [(1 - \gamma)s - \gamma PB]
\]

\[
\tau(x) \sim \frac{P^2}{D(x)}
\]

Vertical equilibrium:

\[
\frac{s}{PB} = \frac{\gamma}{1 - \gamma}.
\]
Model IV

\[
\frac{\partial s}{\partial t} + \frac{\partial [u_s(x)s]}{\partial x} = -\frac{(1 - \gamma)s - \gamma PB}{\tau(x)} + \frac{\partial}{\partial x} \left[ D(x) \frac{\partial s}{\partial x} \right],
\]

\[
\frac{\partial b}{\partial t} + \frac{\partial [u_B(x)b]}{\partial x} = \frac{(1 - \gamma)s - \gamma PB}{\tau(x)}
\]

\[
+ \frac{\partial}{\partial x} \left[ \xi(x)D(x) \frac{\partial}{\partial x} B \right],
\]

\[
\Rightarrow \quad \frac{\partial b}{\partial t} + \frac{\partial [v(x)b]}{\partial x} = \frac{(1 - \gamma)s - \gamma PB}{\tau(x)} + \frac{\partial}{\partial x} \left[ D(x) \frac{\partial b}{\partial x} \right],
\]
Finite dimensional reduction

\[ q = b + s \quad \implies \quad s \approx \frac{\gamma P}{\gamma P + (1 - \gamma)\xi} q, \quad b \approx \frac{(1 - \gamma)\xi}{\gamma P + (1 - \gamma)\xi} q. \]

\[
\frac{\partial q}{\partial t} = \frac{\partial}{\partial x} \left[ D(x) \frac{\partial q}{\partial x} - u_e(x)q \right],
\]

\[
u_e(x) = \frac{\gamma P u_S + (1 - \gamma)\xi v(x)}{\gamma P + (1 - \gamma)\xi}.
\]

- Independent of \( \tau(x) \).
Asymptotic solutions

- $q \approx \frac{1}{\sqrt{2\pi \sigma^2(t)}} \exp \left[ -\frac{(x - \mu(t))^2}{2\sigma^2(t)} \right]$ describes an evolving unit mass Gaussian pulse

$$\frac{\partial}{\partial t} \mu = u_e(\mu) + \frac{\partial D}{\partial x}(\mu), \quad \frac{\partial}{\partial t} \sigma^2 = 2D(\mu),$$

This model works best at early times, when $\sigma(t)/\mu(t) \ll 1$, so the pulse does not “feel” the boundary condition at the shore $x = 0$.

- Steady state

$$q \rightarrow q_\infty = C \exp \left[ \int \frac{u_e(x)}{D(x)} \, dx \right], C \text{ is a normalizing constant.}$$
Qualitative behavior of solutions

- Maxima of the steady state solution are at points where $u_e = 0$.
- In what circumstances is the maximum of the steady state distributions “significantly” away from the shore, i.e. the location of the maximum is on the scale of $L$, the width of the break zone?

Dimensionless parameters

\[
\beta = \left( \frac{P - H_0}{H_\infty - H_0} \right) \frac{X}{L}.
\]

\[
\delta = \frac{D_L}{D_{threshold}} = \frac{(1 - \gamma)D_L(H_\infty - H_0)}{\gamma P|\mathcal{U}^{St}|X}.
\]
Patched solutions

When $\mu(t) = a\sigma(t)$, where $a$ is a $O(1)$ constant, switch from the description in terms of a moving Gaussian pulse to the description

$$q(x, t) \approx q_\infty(x) + f(x)e^{-\lambda_1 t},$$

where

$$\frac{\partial}{\partial x} \left[ D(x) \frac{\partial f}{\partial x} - u_e(x)f \right] = -\lambda_1 f$$

$t_e$ is the switching time:

$$q(x, t) \approx q_\infty + \left[ \sqrt{\frac{2}{\pi \sigma^2(t_e)}} \frac{1}{1 + \text{Erf} (\sqrt{a/2})} \exp \left[ -\frac{(x - \mu(t_e))^2}{2 \sigma^2(t_e)} \right] - q_\infty \right] e^{-\lambda_1 (t - t_e)},$$
Comparison of full and patched solutions

\[ P = 1, \gamma = 0.9 \]
Increase $\beta$

Increase $\delta$
Aging
Oil as a composite

- **Oil in the environment is a composite** with tens to hundreds of distinct individual species.

- Biological, Chemical and Physical processes act with different rates on the components.

- Modeling is complicated: **Composition dependence** implies **History dependence** in the dynamics of the composite.

- **Goal:** Develop stochastic, autonomous and low dimensional models for environmental processes acting on oil, and couple with models of large scale flows in the ocean.
Setup

“Evaporation” – a linear decay process with rates that can vary (widely!) among the different components.

For species $i$: $\frac{\partial c_i(t)}{\partial t} = -\alpha_i c_i(t)$.

The species are indexed by $1 \leq i \leq N$ with $N \gg 1$.

Choose the indexing so that $\alpha_{i+1} > \alpha_i$.

Observable: $M(t) = \sum_i \beta_i c_i(t)$.

Change of variables + “Continuum limit” $N \to \infty$:

$\frac{\partial \rho(w,t)}{\partial t} = -w \rho(w,t), \quad M(t) = \int_0^1 \rho(w) \, dw$. 
Mori-Zwanzig formalism

Full dynamics: $\partial_t \rho = L \rho$.
$P$ is projection onto the observables, $Q = I - P$.
Decompose state into observables and noise: $\xi = P \rho, \eta = \rho - \xi = Q \rho$.

$$\dot{\eta} = QL \xi + QL \eta$$

$$\eta(t) = \int_0^t e^{(t-s)QL}QL\xi(s)ds + e^{tQL}\eta(0)$$

$$\dot{\xi} = PL \xi + PL \eta$$

$$\xi = \underbrace{PL \xi}_{\text{Markovian}} + \underbrace{\int_0^t PLe^{(t-s)QL}QL\xi(s)ds}_{\text{memory}} + \underbrace{PLe^{tQL}\eta(0)}_{\text{noise}}$$

This is an example of a Generalized Langevin equation.
Analytic results

\[ M = \int \rho dw, \quad P = 1 \otimes \int \text{ yields} \]

\[ \dot{M}(t) = -\frac{1}{2}M(t) + \int_0^t h(t-s)M(s)ds + \beta(t). \]

**History kernel:** \( h(t) = \mathcal{L}^{-1} \left[ (\log(1+s^{-1}))^{-1} \right] \) is given by an inverse Laplace transform.

\( \beta \) is "noise" resulting from the uncertainties in the (unobservable) density distribution \( \rho(w,t) \).

The \( s \to 0 \) behavior of the Laplace transform is **non-analytic** \( \implies \) **fat-tail** for the memory kernel.

\[ h(t) \sim \frac{1}{t(\log(t))^2}. \]
Mean and fluctuations
Filtering, Prediction and Data assimilation

"Invariant measure" is $\rho = 0$! We thus don’t have a Fluctuation-dissipation theorem. How to characterize the noise process, which is now non-stationary?

State space model: $M_n = \sum_{j=1}^{n} h_j M_{n-j} + \beta_n$.

Noisy measurements: $\hat{M}_n = M_n + \sigma \gamma_n$ where $\gamma_n$ are uncorrelated normal variates.

If we assume $\text{var}(\beta_n) \gg \sigma^2$, we have the prediction/filtering algorithm $\tilde{M}_n \approx \sum_{j=1}^{n} h_j \hat{M}_{n-j}$.

**Question:** How to truncate the sum?
Filter error

![Graph showing filter error over time with different filtering methods: No filtering, Truncated FIR, and Empirical filter. The x-axis represents time, and the y-axis represents inferred error on a logarithmic scale. The graph illustrates the performance of each method over time, with the Empirical filter showing the lowest error.]
Conclusions

- Many interesting PDE problems arise from study of oil in the environment.

- Combining observational data with models.

- New methods for analysis: Rigorous derivation of low dimensional models.