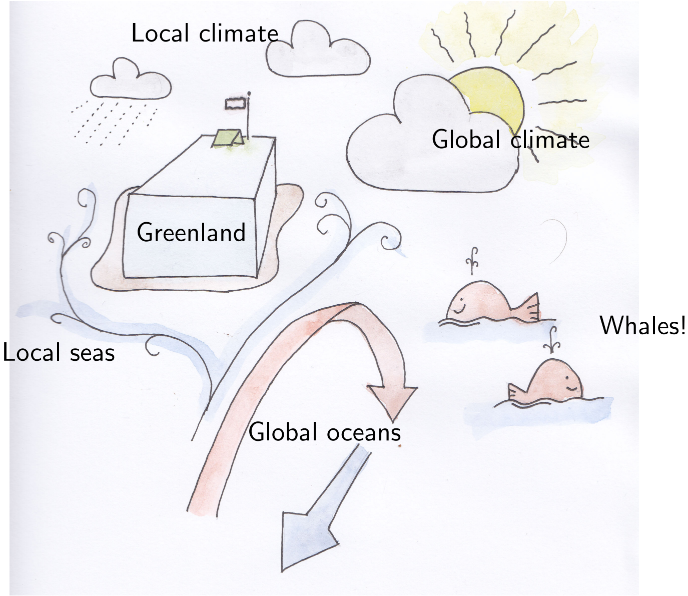
Complex systems: Accounting for model limitations

Jonathan Rougier

Department of Mathematics
University of Bristol, UK

Research Students' Conference, Warwick,
Monday 12 April 2010

Illustration: the Greenland ice-sheet



Simplest interesting example

Conditional on θ:

$$x_0 \sim \pi_{x_0}(\cdot;\theta) \qquad \text{(init. cond. unc.)}$$

$$x_t = g(x_{t-1}, \omega_t; \theta) \qquad \text{(state eqn.)}$$

$$y_t = f(x_t;\theta) + \nu_t \qquad \text{(obs. eqn.)}$$

where

$$\omega_t \overset{\text{iid}}{\sim} \mathrm{N}(0, I) \qquad \text{(structural uncertainty)}$$

$$\nu_t \overset{\text{iid}}{\sim} \mathrm{N}(0, v^2) \qquad \text{(measurement unc.)}$$

and then let $\theta \sim \pi_\theta(\cdot)$, to account for parametric uncertainty. The functions f and g are given, likewise the measurement uncertainty standard deviation, v.

Sampling from $\{x_0, x_1, \ldots, x_T, \theta\} \mid \{y_1, \ldots, y_T\}$ "intractable and unsolved" (C. Andrieu)

The calibration problem

To learn about θ, typically by summarising samples from the distribution $\pi(\theta \mid \mathbf{y})$, where $\mathbf{y} = (y_1, \ldots, y_T)$. We'll treat x_0 as known, for simplicity.

- Ideally, we would run an MCMC chain with proposal $q(\theta \to \theta')$ and acceptance probability

$$\alpha(\theta, \theta') = 1 \wedge \frac{\pi(\mathbf{y} \mid \theta')\, \pi(\theta')}{\pi(\mathbf{y} \mid \theta)\, \pi(\theta)}\, \frac{q(\theta' \to \theta)}{q(\theta \to \theta')}.$$

The calibration problem

To learn about θ, typically by summarising samples from the distribution $\pi(\theta \mid \mathbf{y})$, where $\mathbf{y} = (y_1, \ldots, y_T)$. We'll treat x_0 as known, for simplicity.

- Ideally, we would run an MCMC chain with proposal $q(\theta \to \theta')$ and acceptance probability

$$\alpha(\theta, \theta') = 1 \wedge \frac{\pi(\mathbf{y} \mid \theta')\, \pi(\theta')}{\pi(\mathbf{y} \mid \theta)\, \pi(\theta)}\, \frac{q(\theta' \to \theta)}{q(\theta \to \theta')}.$$

- The catch is that we need to integrate out $\boldsymbol{\omega}$ in order to evaluate $\pi(\mathbf{y} \mid \theta)$:

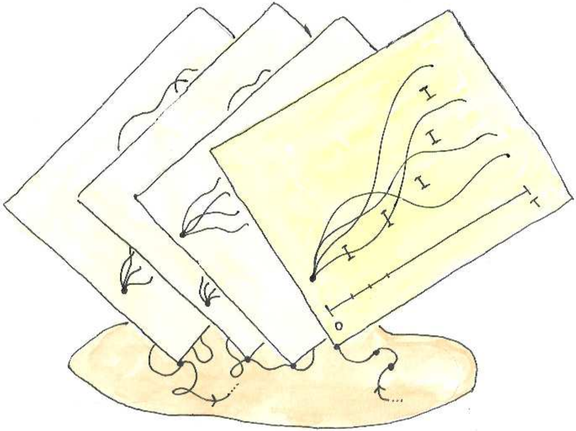
$$\pi(\mathbf{y} \mid \theta) = \int \pi(\mathbf{y} \mid \boldsymbol{\omega}, \theta)\, \pi(\boldsymbol{\omega})\, \mathbf{d}\boldsymbol{\omega}$$

  where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_T)$ — *ouch!!*

The calibration problem (cont.)

In a picture . . .

$(\boldsymbol{\omega} \mid \theta)$-space (high-dimensional)



θ-space (low-dimensional)

The calibration problem (cont.)

- We *could* approximate this tricky density with an Importance Sampler estimate

$$\tilde{\pi}(\mathbf{y} \mid \theta) = N^{-1} \sum_{i=1}^{N} \frac{\pi(\mathbf{y} \mid \boldsymbol{\omega}^i, \theta)\, \pi(\boldsymbol{\omega}^i \mid \theta)}{q_\omega(\boldsymbol{\omega}^i; \mathbf{y}, \theta)} \quad \boldsymbol{\omega}^i \overset{\text{iid}}{\sim} q_\omega(\cdot; \mathbf{y}, \theta),$$

  but how would this affect the MCMC chain?

The calibration problem (cont.)

- We *could* approximate this tricky density with an Importance Sampler estimate

$$\tilde{\pi}(\mathbf{y} \mid \theta) = N^{-1} \sum_{i=1}^{N} \frac{\pi(\mathbf{y} \mid \boldsymbol{\omega}^i, \theta)\, \pi(\boldsymbol{\omega}^i \mid \theta)}{q_\omega(\boldsymbol{\omega}^i; \mathbf{y}, \theta)} \quad \boldsymbol{\omega}^i \overset{\text{iid}}{\sim} q_\omega(\cdot; \mathbf{y}, \theta),$$

  but how would this affect the MCMC chain?

- Mark Beaumont's (2003) result. The θ-marginal of the MCMC equilibrium distribution with $\pi(\mathbf{y} \mid \theta)$ replaced by $\tilde{\pi}(\mathbf{y} \mid \theta)$ is *still* $\pi(\theta \mid \mathbf{y})$, for all $N \geq 1$.
  1. One has to accept/reject $\{\boldsymbol{\omega}^i\}$ along with θ.
  2. Small N normally implies a sticky chain.

- The general result was stated by Andrieu et al (2007): the θ-marginal is $\pi(\theta \mid \mathbf{y})$ for any unbiased estimator of $\pi(\mathbf{y} \mid \theta)$.

The stochastic van der Pol oscillator

Has a 'slow' response x and a 'fast' response x', related as

$$x'' + x' + (\alpha - x^2)x = \sigma x \, \mathrm{d}W,$$

where W is a Brownian motion. It is the basis for several phenomenological models of glacial cycles, where x is 'ice volume' and x' is 'temperature' (or, effectively, 'CO_2'). Here $\theta = (\alpha, \sigma)$.

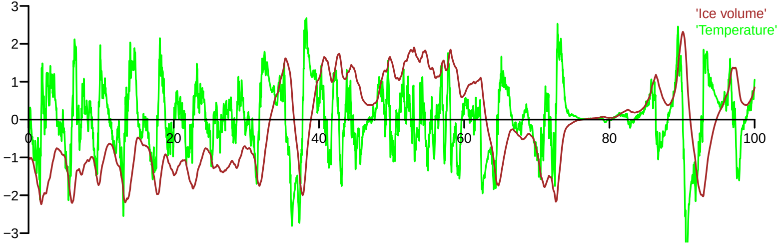
The stochastic van der Pol oscillator

Has a 'slow' response x and a 'fast' response x', related as

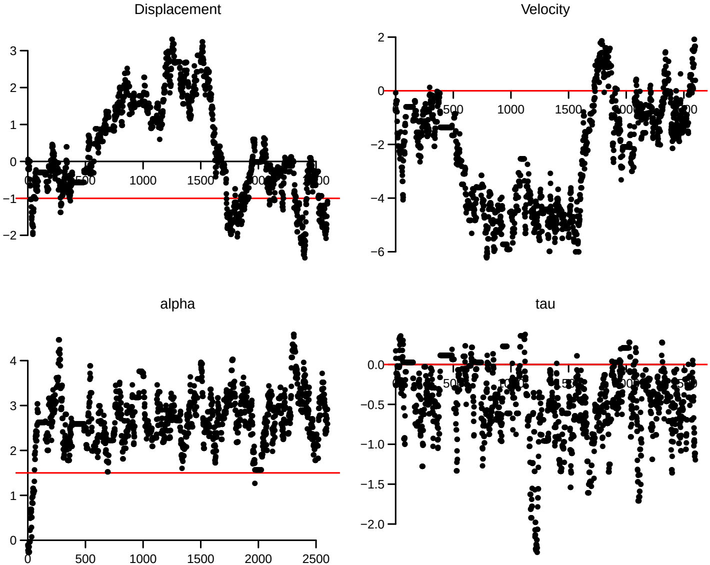
$$x'' + x' + (\alpha - x^2)x = \sigma x \, \mathrm{d}W,$$

where W is a Brownian motion. It is the basis for several phenomenological models of glacial cycles, where x is 'ice volume' and x' is 'temperature' (or, effectively, 'CO_2'). Here $\theta = (\alpha, \sigma)$.

One realisation:



The evidence (24 hours)



Summary

- In inference for environmental systems, model limitations require us to account for both parametric and structural uncertainty.
- The generic problem for dynamical systems is therefore *non-linear data assimilation with uncertain static parameters.*
- Learning about the parameters involves integrating out the high-dimensional state vector; this can be done 'exactly' (in the MCMC sense) using Beaumont's result.
- Recent developments not mentioned here have generalised this approach (e.g. pseudo-marginal approach, particle-MCMC).
- This is an exciting time to be working as a statistician in environmental science!