Introduction to computer experiments, and the challenges of expensive models

Jonathan Rougier

Department of Mathematics
University of Bristol UK

RSS SCS, Bath, Wed 28 Oct 2009

Two types of experiment

System experiment

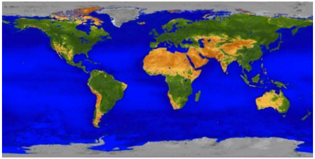
- Multiple experimental units may not be available!
- Many uncontrolled sources of variation
- *Difficulty of doing the experiment you want*

Two types of experiment

System experiment

- Multiple experimental units may not be available!
- Many uncontrolled sources of variation
- *Difficulty of doing the experiment you want*

Computer experiment



- Can do more-or-less any experiments we want
- Have to account for limitations in the simulator
- *Difficulty of interpreting the experiment you do*

Illustration: the Greenland ice-sheet

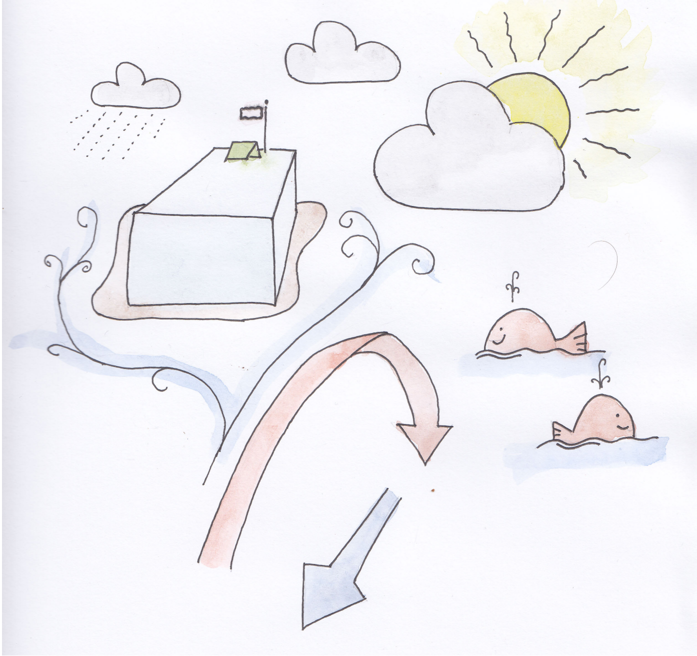
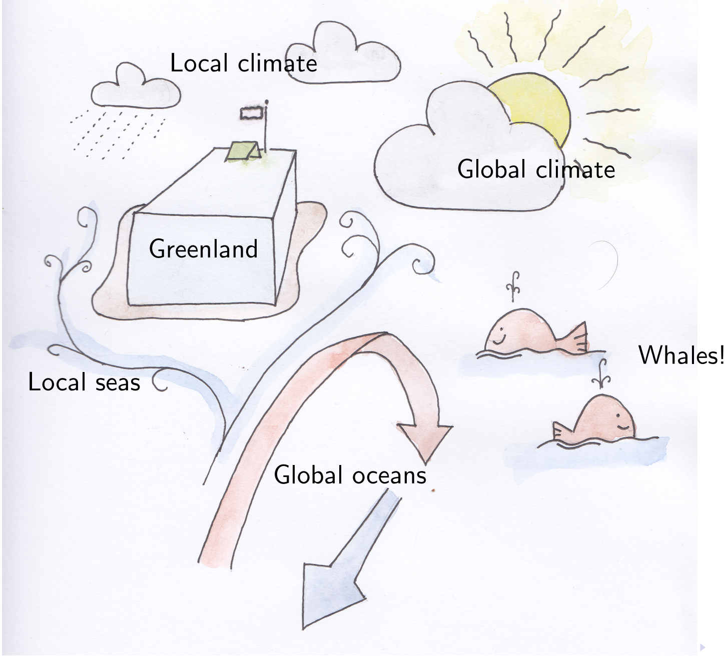


Illustration: the Greenland ice-sheet



Current practice in Environmental Science

System Complex interdependencies:

- W Global oceans and climate, biosphere
- X Local seas, local climate
- Y Volume and dynamics of Greenland ice-sheet

Current practice in Environmental Science

System Complex interdependencies:

- W Global oceans and climate, biosphere
- X Local seas, local climate
- Y Volume and dynamics of Greenland ice-sheet

Simulator $y = s(x, \theta)$, deterministic, ignores W, asserts causal direction from X to Y, contains unknown parameters $\tilde{\theta}$

Current practice in Environmental Science

System: Complex interdependencies:

- W Global oceans and climate, biosphere
- X Local seas, local climate
- Y Volume and dynamics of Greenland ice-sheet

Simulator: $y = s(x, \theta)$, deterministic, ignores W, asserts causal direction from X to Y, contains unknown parameters $\tilde{\theta}$

Observations: $Z = z^{\text{obs}}$ where usually $Z_i \sim \mathrm{N}\big(m_i \cdot Y, \sigma_i^2\big)$

Current practice in Environmental Science

System Complex interdependencies:

- W Global oceans and climate, biosphere
- X Local seas, local climate
- Y Volume and dynamics of Greenland ice-sheet

Simulator $y = s(x, \theta)$, deterministic, ignores W, asserts causal direction from X to Y, contains unknown parameters $\tilde{\theta}$

Observations $Z = z^{\text{obs}}$ where usually $Z_i \sim \mathrm{N}\big(m_i \cdot Y, \sigma_i^2\big)$

'Tuning' Select a representative value $\hat{x}$ for X and then

$$\hat{\theta} = \operatorname*{argmin}_{\theta} \sum_i \frac{1}{\sigma_i^2} \Big\{ z_i^{\text{obs}} - m_i \cdot s(\hat{x}, \theta) \Big\}^2 .$$

Current practice in Environmental Science

System Complex interdependencies:

- W Global oceans and climate, biosphere
- X Local seas, local climate
- Y Volume and dynamics of Greenland ice-sheet

Simulator $y = s(x, \theta)$, deterministic, ignores W, asserts causal direction from X to Y, contains unknown parameters $\tilde{\theta}$

Observations $Z = z^{\text{obs}}$ where usually $Z_i \sim \mathrm{N}\big(m_i \cdot Y, \sigma_i^2\big)$

'Tuning' Select a representative value $\hat{x}$ for X and then

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum\nolimits_i \frac{1}{\sigma_i^2} \Big\{ z_i^{\text{obs}} - m_i \cdot s(\hat{x}, \theta) \Big\}^2 .$$

'Prediction' $\hat{y} = s(\hat{x}, \hat{\theta})$.

A quick inventory

- In the previous outline, there are two sources of uncertainty unaccounted for:
  - The true value of X – using $\hat{x}$ as a plug-in may mask a large impact from variability in X.
  - The limitations of the simulator – implicitly it is treated as perfect.

A quick inventory

- In the previous outline, there are two sources of uncertainty unaccounted for:
  - The true value of X – using $\hat{x}$ as a plug-in may mask a large impact from variability in X.
  - The limitations of the simulator – implicitly it is treated as perfect.
- There is an additional source of information that has been ignored:
  - Physical interpretation of $\tilde{\theta}$, for which some components may relate to observable features of the system.
- Ideally, we would like to incorporate all of these, and we can do this within a Bayesian framework in which learning about $(X, Y, \tilde{\theta})$ takes the form of estimating or sampling from

$$\Pr\{X, Y, \tilde{\theta} \mid Z = z^{\text{obs}}\}.$$

Structuring the joint distribution

We factorise the joint distribution of $(X, Y, \tilde{\theta})$ as

$$\pi(X, Y, \tilde{\theta}) = \underbrace{\pi(Y \mid X, \tilde{\theta})}_{\text{structural,}} \times \underbrace{\pi(X)}_{\text{input,}} \times \underbrace{\pi(\tilde{\theta})}_{\text{parametric uncertainty}}$$

where $X \perp\!\!\!\perp \tilde{\theta}$ seems reasonable.

The three sources of uncertainty (terminology):

Parametric Not knowing the 'correct' value of the simulator parameters;

Input Not knowing the true value of X;

Structural Not knowing the true value of Y, even were we to know X and $\tilde{\theta}$.

Structural uncertainty

Including structural uncertainty is a big advance.

- The 'perfect simulator' has

$$\pi(Y \mid X, \tilde{\theta}) = \delta\big(Y - s(X, \tilde{\theta})\big),$$

where $\delta(\cdot)$ is the Dirac delta function.

Structural uncertainty

Including structural uncertainty is a big advance.

- The 'perfect simulator' has

$$\pi(Y \mid X, \tilde{\theta}) = \delta\big(Y - s(X, \tilde{\theta})\big),$$

where $\delta(\cdot)$ is the Dirac delta function.

- The current SOTA ('best input' approach) has

$$\pi(Y \mid X, \tilde{\theta}) = \pi_\epsilon\big(Y - s(X, \tilde{\theta})\big)$$

where, typically,

$$\epsilon \sim \mathrm{N}(\mathbf{0}, \Sigma);$$

ϵ is known as the discrepancy and Σ as the discrepancy variance. In the case $\Sigma \to \mathbf{0}$ we are back at the perfect simulator, so this is a *friendly generalisation*.

The main challenges

At its simplest, the Bayesian approach to calibration requires us to 'score' samples from the prior distribution of $\tilde{\theta}$ using the *likelihood function*

$$L(\theta) \propto \phi\big(z^{\text{obs}}; Ms(\hat{x}, \theta), M\Sigma M^T + D\big)$$

where ϕ is the Gaussian density function, M is the incidence matrix, D is the diagonal matrix of measurement error variances.

(For simplicity: not integrating out X, although we could.)

The main challenges

At its simplest, the Bayesian approach to calibration requires us to 'score' samples from the prior distribution of $\tilde{\theta}$ using the *likelihood function*

$$L(\theta) \propto \phi\big(z^{\text{obs}}; Ms(\hat{x}, \theta), M\Sigma M^T + D\big)$$

where ϕ is the Gaussian density function, M is the incidence matrix, D is the diagonal matrix of measurement error variances.

(For simplicity: not integrating out X, although we could.)

Challenges

1. Specifying $\pi(X)$ and the discrepancy variance, Σ. We may choose to revert to $\pi(X) = \delta(X - \hat{x})$, as above, if uncertainty about X is not thought to be substantial.

The main challenges

At its simplest, the Bayesian approach to calibration requires us to 'score' samples from the prior distribution of $\tilde{\theta}$ using the *likelihood function*

$$L(\theta) \propto \phi\big(z^{\text{obs}}; Ms(\hat{x}, \theta), M\Sigma M^T + D\big)$$

where ϕ is the Gaussian density function, M is the incidence matrix, D is the diagonal matrix of measurement error variances.

(For simplicity: not integrating out X, although we could.)

Challenges

1. Specifying $\pi(X)$ and the discrepancy variance, Σ. We may choose to revert to $\pi(X) = \delta(X - \hat{x})$, as above, if uncertainty about X is not thought to be substantial.

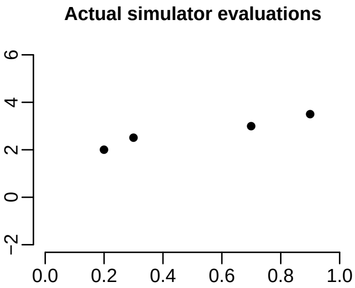
2. Doing the inferential calculation. This is a problem whenever the simulator is sufficiently expensive that it is unrealistic to span the parameter space with evaluations.

Enter *The Emulator*

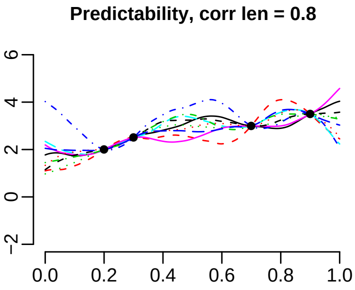
An emulator is a Bayesian statistical framework for predicting the simulator output, which augments information from an ensemble of evaluations with additional judgements about:

- Smoothness and differentiability
- Monotonicity and explicit functional relationships
- Important low-order interactions.

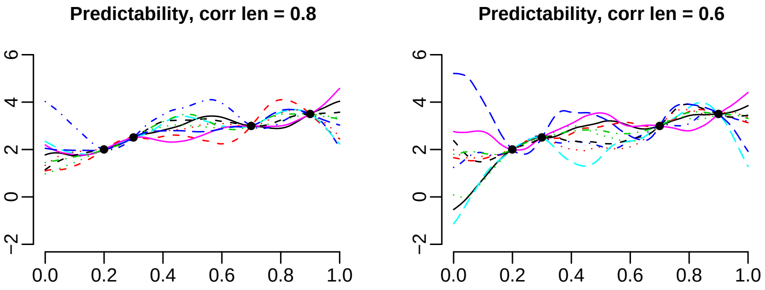
Enter *The Emulator*



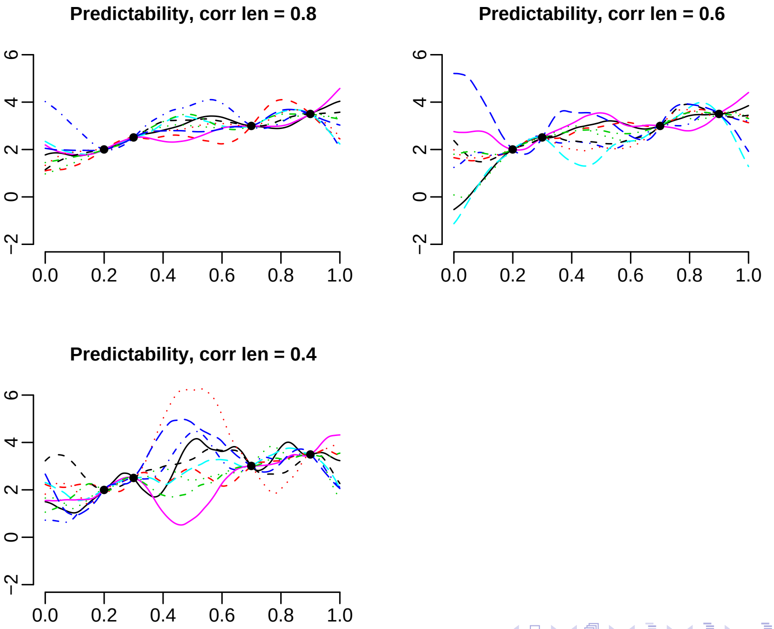
Enter *The Emulator*



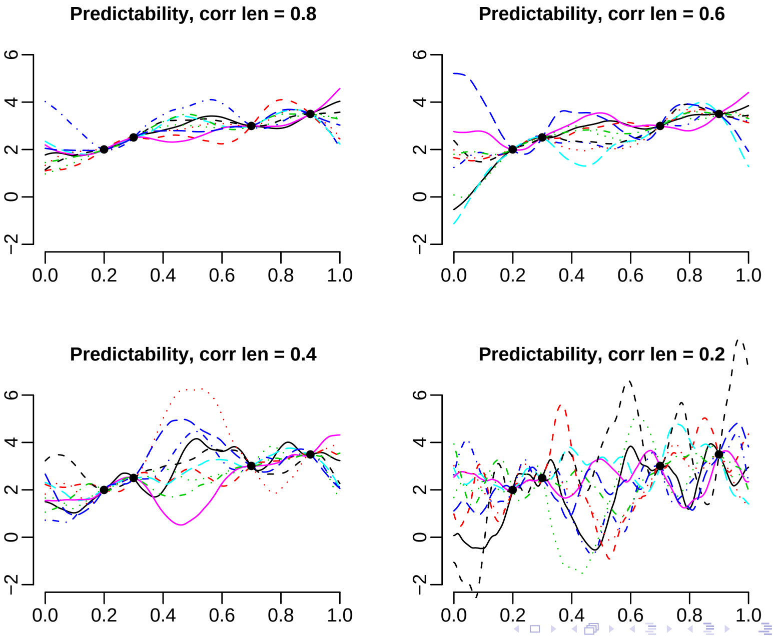
Enter *The Emulator*



Enter *The Emulator*



Enter *The Emulator*



Enter *The Emulator* (cont)

- Emulators are usually fitted as Gaussian Processes (possibly after transformation), in which case the output from an emulator is a mean function $\mu(\theta) := E\{s(\hat{x}, \theta)\}$ and variance function $\Psi(\theta) := \mathrm{Var}\{s(\hat{x}, \theta)\}$.
- The likelihood function in this case is

$$L(\theta) \propto \phi\big(z^{\text{obs}}; M\mu(\theta), M(\Psi(\theta) + \Sigma)M^{T} + D\big),$$

i.e. $\mu(\theta)$ has replaced $s(\hat{x}, \theta)$, and there is an extra contribution to the variance of $\Psi(\theta)$.

Enter *The Emulator* (cont)

- Emulators are usually fitted as Gaussian Processes (possibly after transformation), in which case the output from an emulator is a mean function $\mu(\theta) := E\{s(\hat{x}, \theta)\}$ and variance function $\Psi(\theta) := \mathrm{Var}\{s(\hat{x}, \theta)\}$.
- The likelihood function in this case is

$$L(\theta) \propto \phi\big(z^{\text{obs}}; M\mu(\theta), M(\Psi(\theta) + \Sigma)M^T + D\big),$$

  i.e. $\mu(\theta)$ has replaced $s(\hat{x}, \theta)$, and there is an extra contribution to the variance of $\Psi(\theta)$.
- Much of the skill in emulator construction is choosing a good set of training data, using screening and experimental design, often sequentially. One useful way to present the emulator is using probabilistic sensitivity analysis.

  There are all sorts of additional challenges when the parameter space becomes large.

Summary

- Inevitably, when we model complicated systems, we will ignore some aspects, and simplify others. Our simulators of such systems are always imperfect.
- Quantifying uncertainty does not happen at the end of the analysis: it occurs right at the start, when we describe the informativeness of our simulator in terms of parametric, input, and structural uncertainty.
- Expensive simulators can only be run a limited number of times; in this case, the choice of runs has to be made carefully, and the information in those runs can be augmented with additional judgements, through constructing an emulator.