Experimental Design for Parameter Estimation and Uncertainty Quantification in Complex Computer Models: Issues and Examples

Derek Bingham
Statistics and Actuarial Science
Simon Fraser University

Matt Pratola – The Ohio State U.
Josh Rigler – US Geological Survey
Steve Sain – Jupiter Intelligence
Mike Wiltberger - US National Center for Atmospheric Research

SFU
Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Outline

- Some background
- An approach to parameter estimation
- Some examples
- Experiment design
- Wrap-up

Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Experiments are often performed on computational models

- Many scientific applications use mathematical models to describe physical systems

- Rapid growth in computer power has made it possible to study complex physical phenomena that might otherwise be too time consuming or expensive to observe

- To understand how inputs to the computer code impact the system, scientists adjust the inputs to computer simulators and observe the response … they conduct a computer experiment

Computer experiments

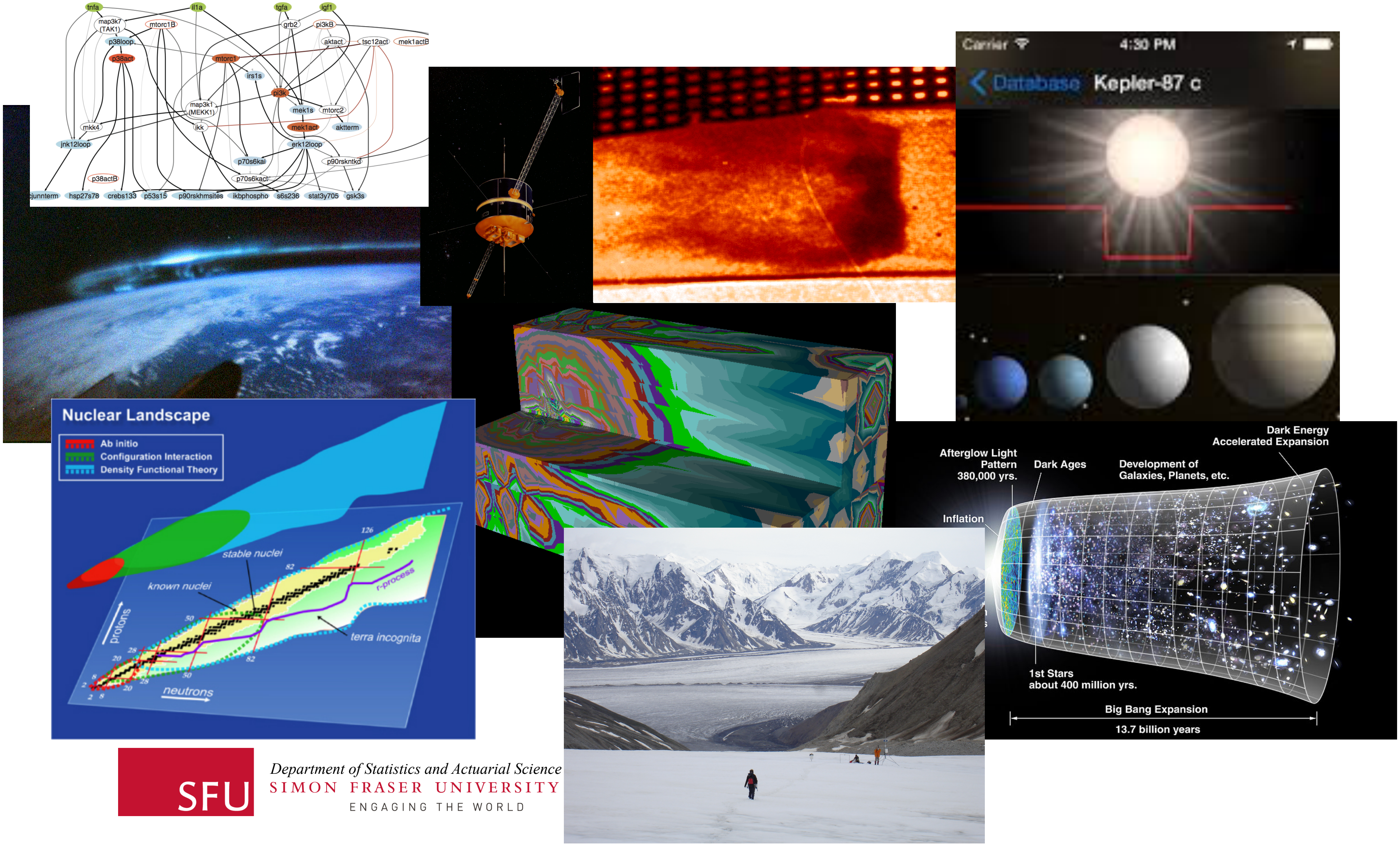
The computer models frequently:

1. Not available on your desktop

2. So computationally intensive that only a limited number of evaluations of the model can be done

- Heitman et al. (2016) describe a campaign that requires roughly 2 years to obtain 100 evaluations of a cosmology model

Some examples of applications



SFU

Department of Statistics and Actuarial Science

SIMON FRASER UNIVERSITY

ENGAGING THE WORLD

Experiment design

Often parameters

- In computer experiments, as in many other types of experiments considered, there are usually some **factors**, ***x***, that can be adjusted and some **response variable**, ***y***, that is impacted by (some) of the factors
- The experiment is conducted, for example to see which/how the factors impact the response
- Generally, the **experimental design** is the set of inputs settings that are used in your simulation experiments (sometimes called the input deck)

Experiment design

- For a computer experiment, the experiment design is the set of d-dimensional inputs where the computational model is evaluated

- **Notation:** ***X*** is an n x d, design matrix; $y(X)$ is the n x 1 vector of responses is have scalar response

- Experimental region is usually the unit hypercube $[0,1]^d$ *(not always)*

Experiment design should meet your goals

- **Experiment goals:**
  - Computing the mean response
  - Response surface estimation or computer model emulation
  - Sensitivity analysis
  - Optimization
  - Parameter estimation
  - Estimation of contours and percentiles

.

.

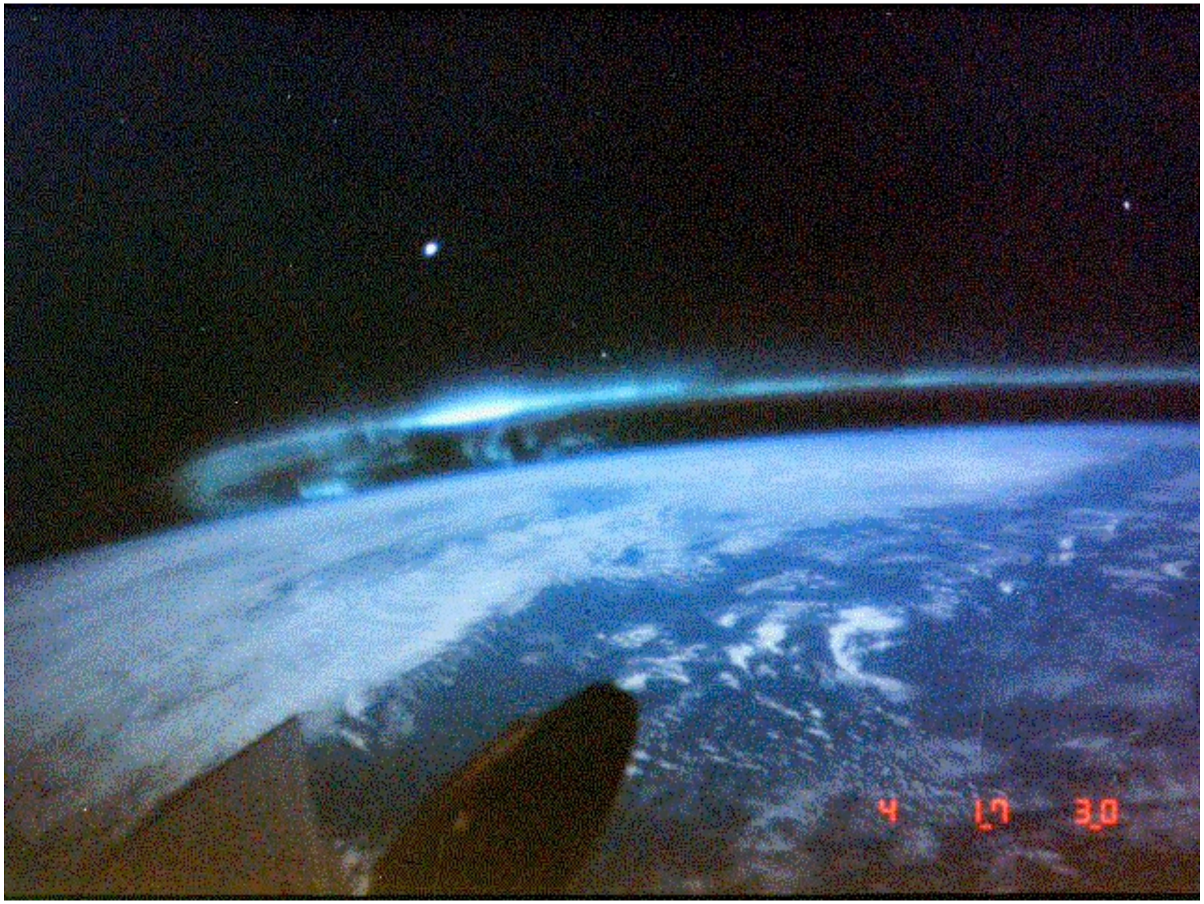
.

To know which ensembles to perform need to know how you are doing parameter estimation

- The Lyon-Fedder-Mobary (LFM) model simulates the interaction of solar wind plasma and the magnetosphere
- We see this as the Aurora Borealis
- Computer model is slow to run... 4 runs ≈ 1 month
- Outputs are two large 3-dimensional space-time fields
- Field observations available from the Polar Ultraviolet Imager satellite (UVI)
- Scientific problem: estimate $\theta = (\alpha, \beta, R)$

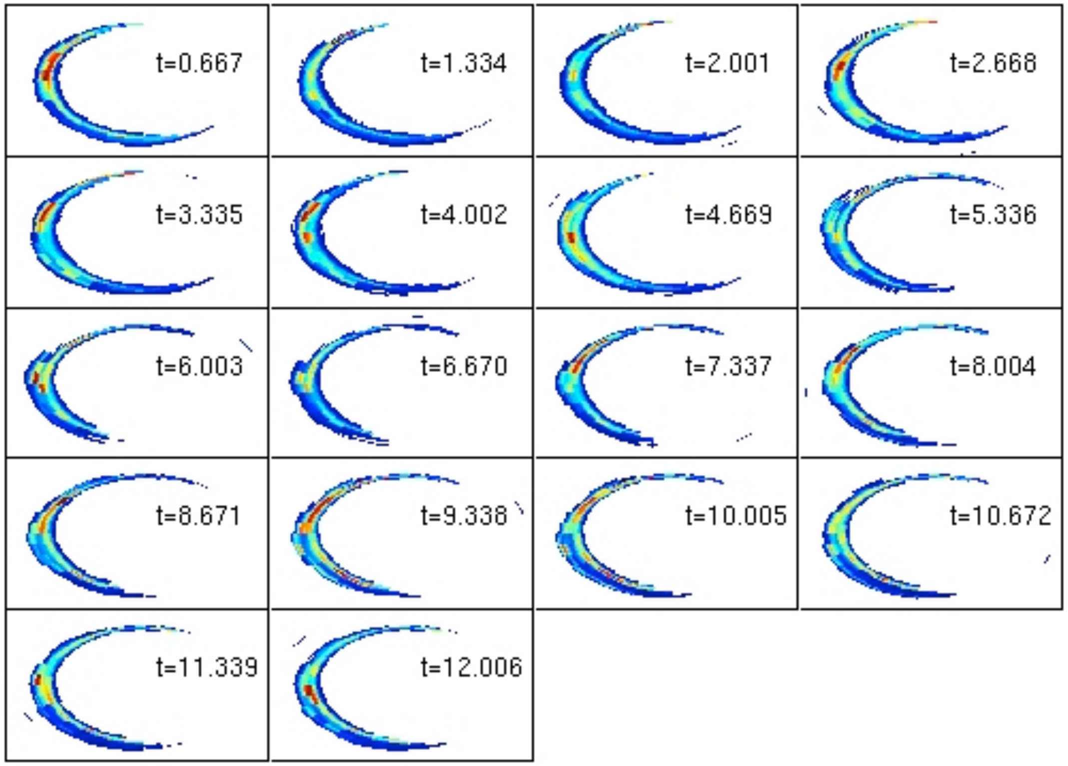
SFU
Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Example



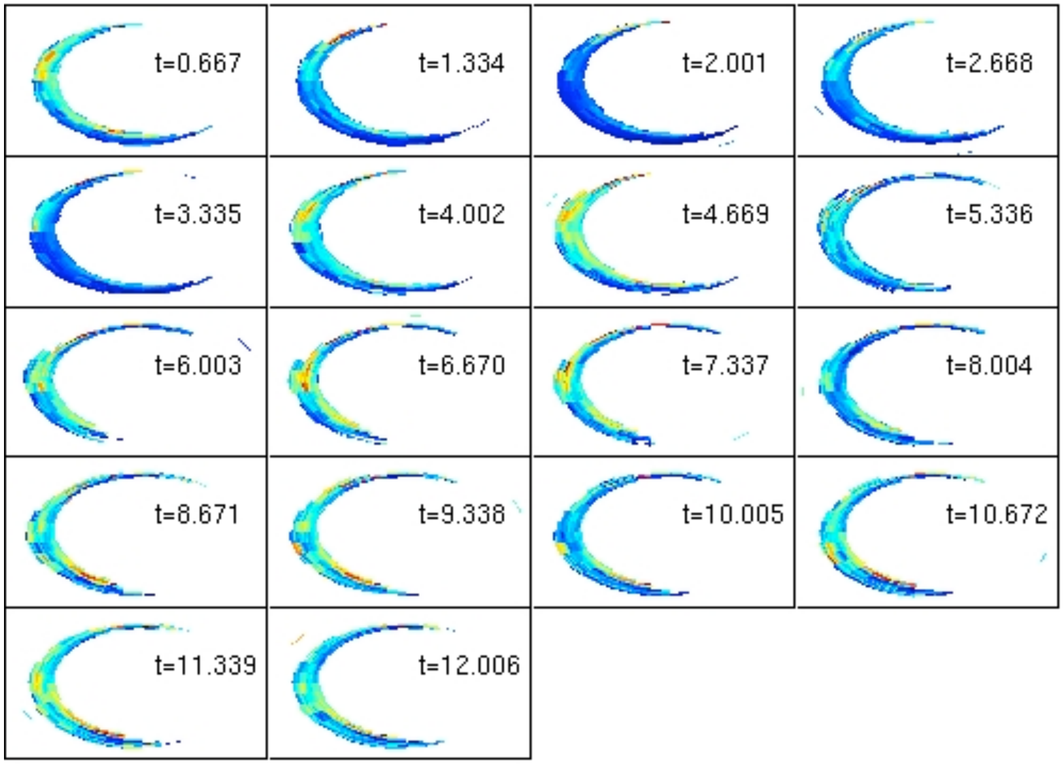
UVI measurements

(Flux – January 10, 1997)



UVI measurements

(Energy– January 10, 1997)



LFM simulations

Want to estimate

- The LFM model has three inputs: $\theta = (\alpha, \beta, R)$

- Model is treated as essentially unbiased:

$$Y_f(\theta) = \eta(\theta) + \epsilon_i$$
$$Y_c(t) = \eta(t)$$

- One approach is to **emulate** the computer model and minimize some distance between the field and computer model as a function of the inputs

LFM simulations

Sampled via Latin hypercube sampling in 3-d

- 20 simulations from the deterministic computer model
- Each simulation gives a space-time field on a 1,944 grid for each response variable (field and model)
- Have one field observation from UVI for both responses on the same grid (i.e., have observed one storm)
- Leads to roughly 40,000 × 40,000 covariance matrix to calibrate one response if, say, a Gaussian process (GP) model is used

LFM simulations

- LFM model output exhibits behavior that is not well represented by a non-stationary GP as a function of the inputs

- **Challenges:** Large data structure and non-stationary covariance model is required

LFM simulations

- Will use a sequential improvement-based approach
- Need a (well designed) collection of initial model runs
- Attempts to measure the discrepancy between the computer model run at a each input setting and the field observation
- **Model the criterion as a function of the (calibration) parameter (will use a Gaussian process model)**
- Estimate of the parameter is where the criterion surface is minimized

Approach

Let:

– $Y_f(\theta) = \eta(\theta) + \epsilon$ field observation

– $Y_c(t) = \eta(t)$ computer model response

Then:

– $\delta(t) = Y_f(\theta) - Y_c(t) = \eta(\theta) - \eta(t) + \epsilon$

Approach

$$\delta(t) = Y_f(\theta) - Y_c(t) = \eta(\theta) - \eta(t) + \epsilon$$

Idea:

- **Restricted model:**

$$t = \theta,\ \delta(\theta) \sim N(0, \sigma^2 I)$$

- **Unrestricted Model:**

$$\delta(t) \sim N(\mu_t, \sigma_t^2 R + \sigma^2 I)$$ ← GP

SFU
Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Use the likelihoods as discrepancy model

- Define the discrepancy criterion:

$$\Delta(t_i) = -2log\left(\frac{L_r^*(\delta(t_i))}{L_u^*(\delta(t_i)}\right)$$

where L_r^* denotes the maximized likelihood for the restricted model, L_u^* denotes the likelihood for the unrestricted model

We model this surface as a GP and minimize

Would be nice to have infinite amounts of simulations and observations, but…

- In practice, we only have limited runs of the computer model at different settings of t

- Model the criterion as a function of t, (use a GP)

$$\Delta(t) \sim N(\mu_\Delta, \sigma^2_\Delta R_\Delta)$$

Now have a response surface that can be minimized

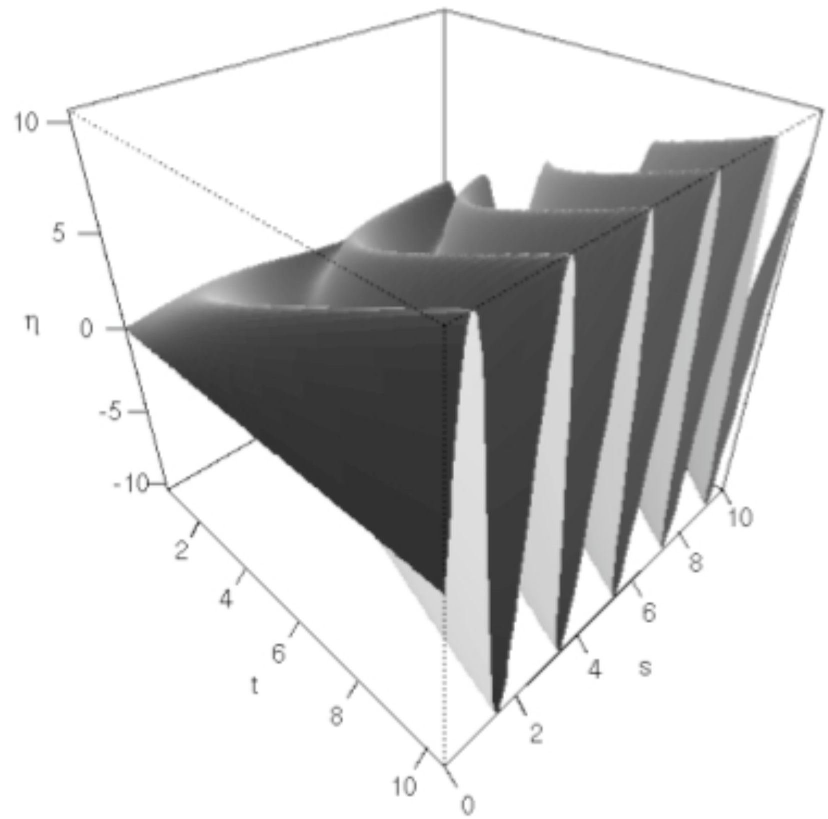
- Let $\hat{\theta} = \arg\min_t \hat{\Delta}(t)$

Toy example

- Consider the simple harmonic oscillator of an object with unit mass

- A simplified form of the solution to the differential equation is (t is the calibration parameter):

$$\eta(t) = t \times sin(ts)$$


10
5
η
0
-5
-10
2
4
6
t
8
10
0
2
4
s
6
8
10

Toy example

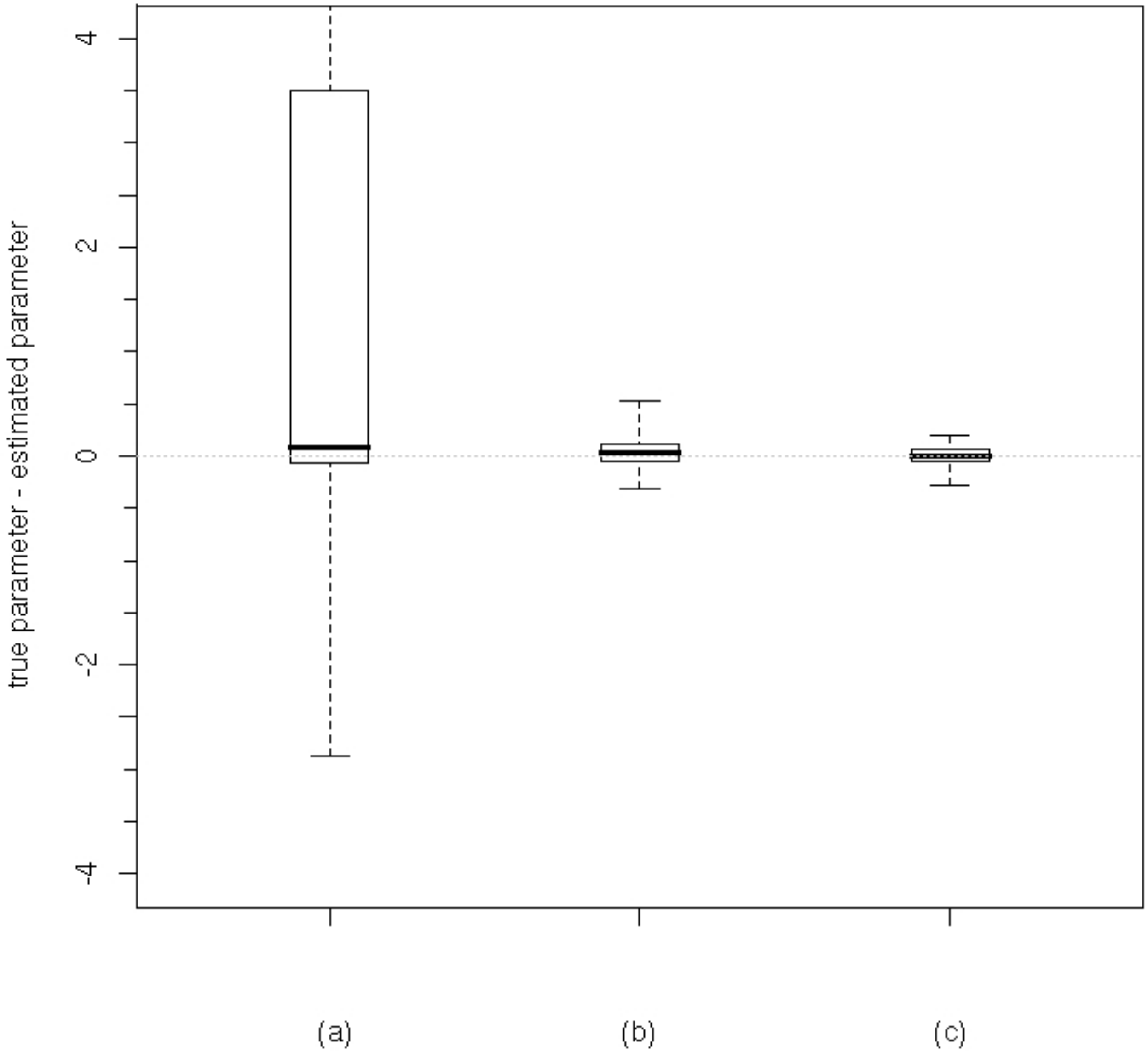
- Simulate from this system with

$$s \in [0, 10] \qquad t \sim U[0, 11]$$

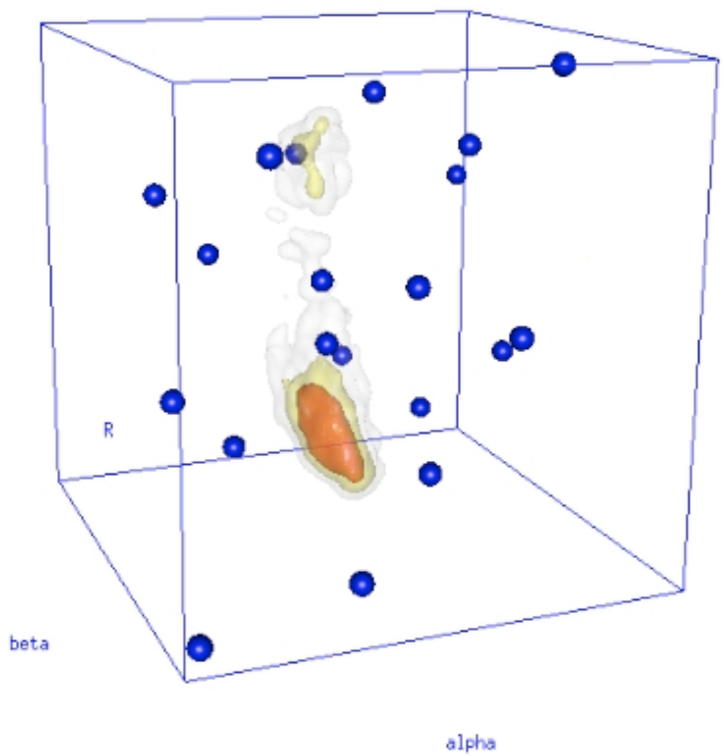
- Error for field observations:

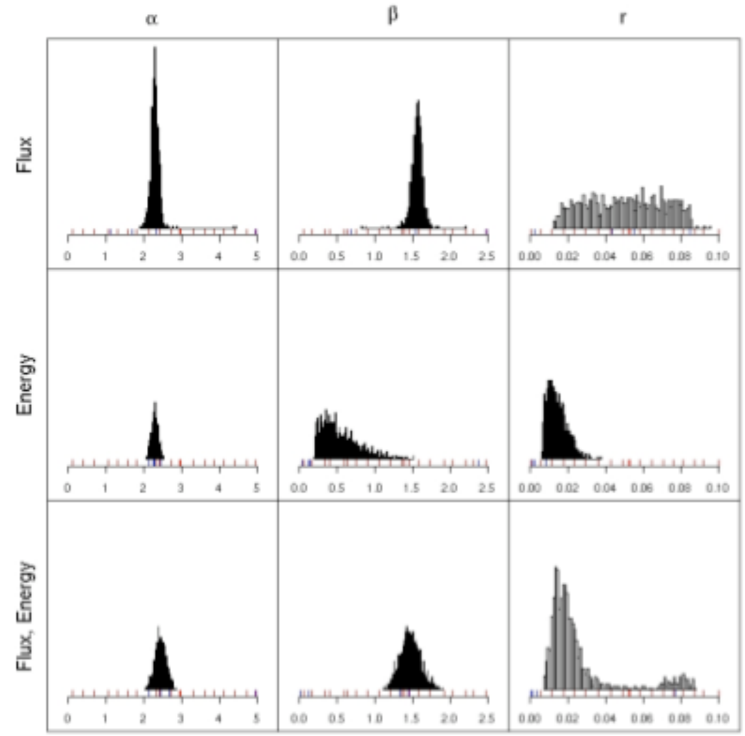
$$\epsilon \sim N(0, \rho Var(\eta(\theta))$$

- Did this 1000 times, each time applying our estimation method


true parameter - estimated parameter
4
2
0
-2
-4
(a)
(b)
(c)

Back to LFM model





SFU
Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Attempt to minimize the discrepancy

- If we are lucky enough to have run the computer model at the correct value of the calibration parameter, then criterion will be small (zero)

- Can show:
  - At $t = \theta$ → $\Delta(t)$ → 0
  - Otherwise discrepancy goes to infinity
  - Can show that this is related to minimizing the Kullback-Leibler divergence

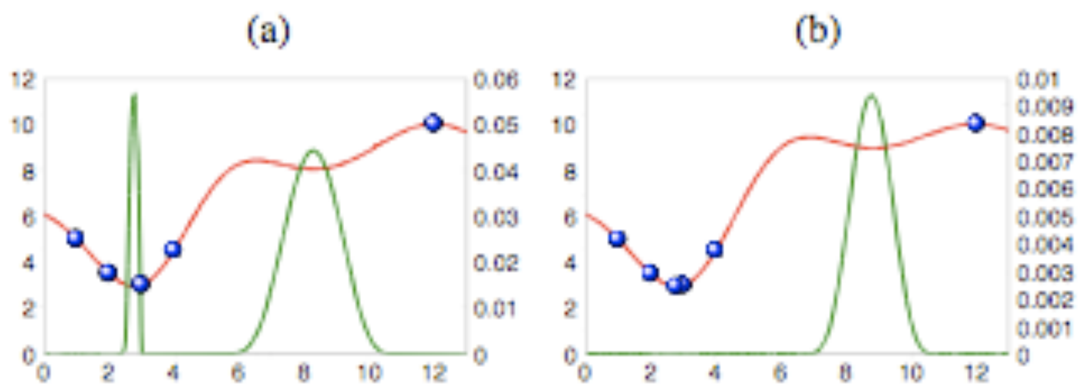
Sequentially choosing model runs

- Criterion allows for simple sequential design strategy based on the Expected Improvement of Jones et al (1998)

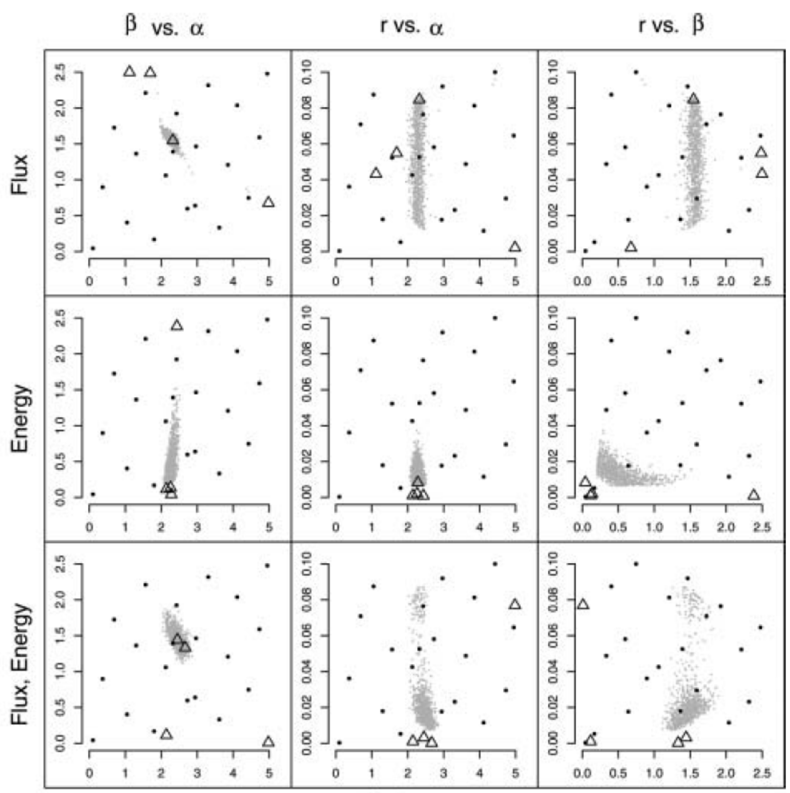
- Define the improvement as $I(t) = max(min\Delta(t) - \hat{\Delta(t)}, 0)$

- New trials are chosen so that E(*I(t)*) is maximized

$$E[I(\boldsymbol{t}^*)] = (\min(\boldsymbol{\Delta}) - \widehat{\Delta}(\boldsymbol{t}^*))\Phi\left(\frac{\min(\boldsymbol{\Delta}) - \widehat{\Delta}(\boldsymbol{t}^*)}{s_\Delta(\boldsymbol{t}^*)}\right) + s_\Delta(\boldsymbol{t}^*)\phi\left(\frac{\min(\boldsymbol{\Delta}) - \widehat{\Delta}(\boldsymbol{t}^*)}{s_\Delta(\boldsymbol{t}^*)}\right),$$



Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD


β vs. α
r vs. α
r vs. β
Flux
Energy
Flux, Energy

Model calibration

- **Main goals:** Use field data and computer model output to
  1. estimate parameters that govern the model (parameter estimation)
  2. build a predictive model for the physical system with estimates of uncertainty

- Can attempt to use the model and “noisy” physical data to estimate parameters that govern the system and make prediction

- Hopefully the predictive model is “better” than just using the code alone or the field data alone

- Would like assessment of uncertainty for parameters and predictions

Model calibration – Statistical formulation

$$y_s(x,t) = \eta(x,t)$$

$$y_f(x,\theta) = \eta(x,\theta) + \delta(x) + \epsilon$$

Kennedy and O'Hagan, 2001
Higdon et al., 2008

Computer model evaluated at x and θ

Typically modelled using Gaussian processes (GPs)

- Where,
  - x model or system inputs
  - y_f system response
  - y_s simulator response
  - θ calibration parameters
  - ϵ random error

Have data from 2 separate sources – field observations and computer model outputs

Also have a model for systematic discrepancy, $\delta(x)$

Thank you for your time

SFU
Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD