# Quantified uncertainties: the good, the bad and the plain ugly

Nick Dudley Ward

OCMO

University of Otago,
7-9 January 2013

# Outline

# The Book

## Motivation, 10 years ago

- Didn't think the people I was working with really knew much about probability (but that didn't stop them using the language)
- Markets exist independently of each other but there is both commonality and exclusivity:
  - Lloyds (catastrophic physical risks)
  - London Stock Exchange
  - Derivative securities
- Lest we forget–excess of loss spiral in Lloyds, 2008 crash

# New motivation

# New motivation

# From the Otago Daily Times, August 2012

**Otago Regional Council's water quality plan changes**

Proposed limit of 10kg of nitrogen per hectare of land per year, on average, going into groundwater.

# From the Otago Daily Times, August 2012

**Otago Regional Council's water quality plan changes**

Proposed limit of 10kg of nitrogen per hectare of land per year, on average, going into groundwater.

**Disgruntled farmer**

'...let's base it on science rather than assumptions or modelling...'. North Otago farmer Robert Borst.

# From the Otago Daily Times, August 2012

### Otago Regional Council's water quality plan changes

Proposed limit of 10kg of nitrogen per hectare of land per year, on average, going into groundwater.

### Disgruntled farmer

'...let's base it on science rather than assumptions or modelling...'. North Otago farmer Robert Borst.

### ORC response

'...modelling was science in action.' ORC policy director Fraser McRae.

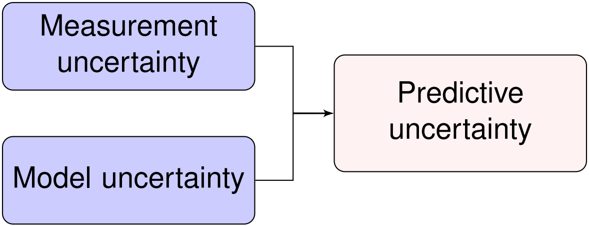
# The Challenges

- How do you sell inverse problems and UQ?
- Regularisation does very nice, thank you!
- If you play the UQ card you open up a can of worms
- Skakeholder buy-in

# The Challenges

- How do you sell inverse problems and UQ?
- Regularisation does very nice, thank you!
- If you play the UQ card you open up a can of worms
- Skakeholder buy-in

First selling point:




Therefore, need to quantify envelope of uncertainty

# Aims

- to be disagreeable!
- to remind us of our responsibilities to engage a broader debate (Sir Paul Callaghan, FRS)
- to invite robust comment

# Outline

# Chapter 1

## Hayman's question

'What is that number $p$ between 0 and 1?' (Prof W.K. Hayman, FRS, Imperial College 1960's)

### Hayman's question

'What is that number $p$ between 0 and 1?' (Prof W.K. Hayman, FRS, Imperial College 1960's)

### Law(s) of averages

- Basic intuition wrapped up in ideas of repeatability and regularity: repeat an experiment indefinitely with $n$ different outcomes (or events). Then the proportion of occurrences of a particular event will 'approach' the probability of that event.
- This is what makes probability a useful tool at *all* levels.
- In particular: 'return period' (engineers, underwriters)
- No good trying to hide under Kolmogorov's shirt tails

# Littlewood on probability

## From a Mathematician's Miscellany:

'Mathematics (by which I shall mean pure mathematics) has no grip on the real world; if probability is to deal with the real world it must contain elements outside mathematics; the *meaning* of probability must relate to the real world, from which we can then proceed deductively (i.e. mathematically). We will suppose (as we may by lumping several primitive propositions together) that there is just one primitive proposition, the 'probability' axiom, and we call it $A$ for short. Although it has got to be true, $A$ is by the nature of the case incapable of deductive proof, for the sufficient reason that it is about the real world ...'

# Littlewood on probability

## From a Mathematician's Miscellany:

'Mathematics (by which I shall mean pure mathematics) has no grip on the real world; if probability is to deal with the real world it must contain elements outside mathematics; the *meaning* of probability must relate to the real world, from which we can then proceed deductively (i.e. mathematically). We will suppose (as we may by lumping several primitive propositions together) that there is just one primitive proposition, the 'probability' axiom, and we call it *A* for short. Although it has got to be true, *A* is by the nature of the case incapable of deductive proof, for the sufficient reason that it is about the real world . . . '

## Back to: 'What is that number *p* between 0 and 1?'

- Mathematical theory gives no prescription for assigning probabilities
- Sufficient that some self evident axioms are satisfied

# Outline

# Earthquake forecasting

## USGS earthquake forecast in SF Bay area

David Freedman's analysis in 'What is the probability of an earthquake?'

# Earthquake forecasting

## USGS earthquake forecast in SF Bay area

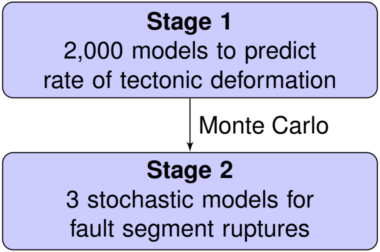
David Freedman's analysis in 'What is the probability of an earthquake?'

**Stage 1**
2,000 models to predict
rate of tectonic deformation

# Earthquake forecasting

## USGS earthquake forecast in SF Bay area

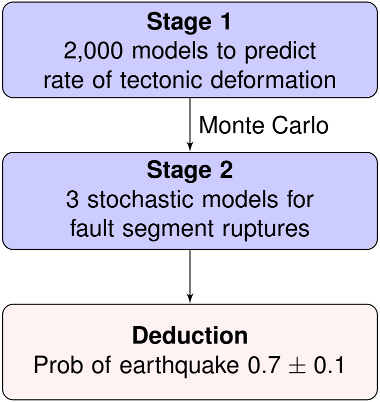
David Freedman's analysis in 'What is the probability of an earthquake?'



# Earthquake forecasting

## USGS earthquake forecast in SF Bay area

David Freedman's analysis in 'What is the probability of an earthquake?'



# Earthquake forecasting: comments

## Freedman's views

- 'by a process we do not understand, those uncertainties [estimated in stage 1] were propagated through stage 2 to estimate the uncertainty of the estimated probability of a large earthquake. If this view is correct, 0.1 is a gross underestimate of the uncertainty'
- 10 sources of error overlooked

# Earthquake forecasting: comments

## Freedman's views

- 'by a process we do not understand, those uncertainties [estimated in stage 1] were propagated through stage 2 to estimate the uncertainty of the estimated probability of a large earthquake. If this view is correct, 0.1 is a gross underestimate of the uncertainty'
- 10 sources of error overlooked

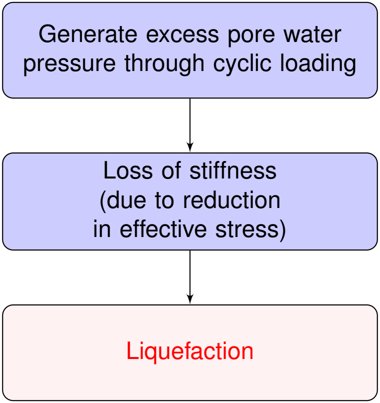
## Morals

- Don't get seduced by Monte Carlo sampling!
- Action is a poor apology for thought (Milne Anderson, UCL)
- Assigning a probability gives a cloak of respectability

# Liquefaction assessment

## What is the risk of liquefaction?

- Simplified method
- Process:



# Liquefaction assessment: simplified method

EQ load

**Cyclic stress ratio**

$$\text{CSR} = \frac{\tau_{\text{cyc}}}{\sigma'_v} = 0.65 \frac{a_{\max}}{g} \frac{\sigma_v}{\sigma'_v} r_d$$

amplitude, (frequency?)

Liquefaction resistance

**Cyclic resistance ratio**

$$\text{CRR} = \text{CRR}_{M=7.5} \times MSF \times K_\sigma$$

If

$$\text{F}_L = \frac{\text{CRR}}{\text{CSR}} < 1$$

then liquefaction

# Sources of uncertainty

## Obvious sources of uncertainty

Even if you accept the procedure there are numerous uncertainties with the scalings:

- 0.65
- Stress reduction factor, $r_d$ (many empirical formulae)
- Magnitude scaling factor, MSF
- Effects of overburden stress $K_\sigma$

- If based on lab tests:

$$\mathrm{CRR}_{\mathrm{field}} = C_1 C_2 C_3 C_4 C_5 \mathrm{CRR}_{\mathrm{triaxial}}$$

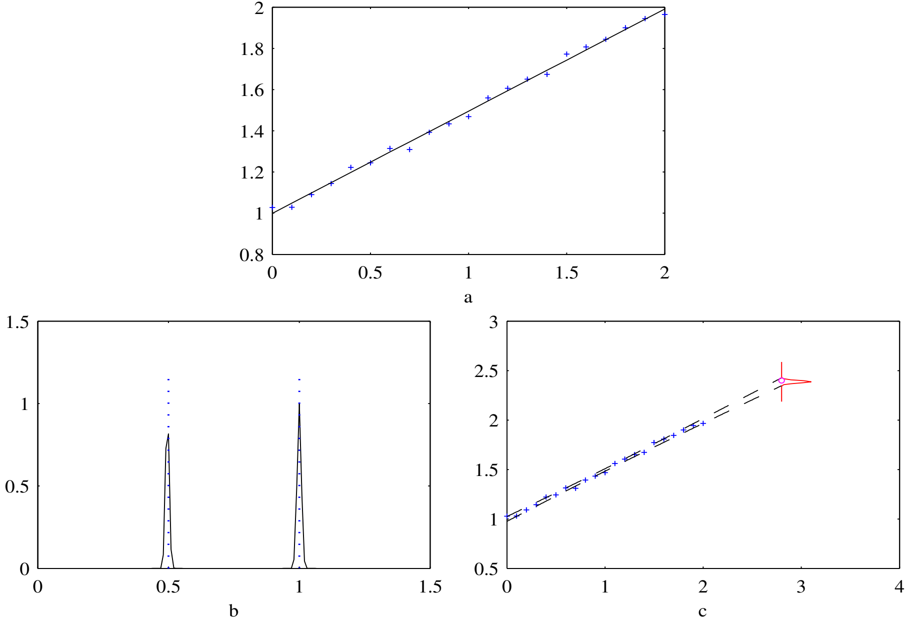
As well as with other corrections for soil type, sloping ground &c., and the all important interpretation of test data

## Then there are harder questions about the underlying methodology:

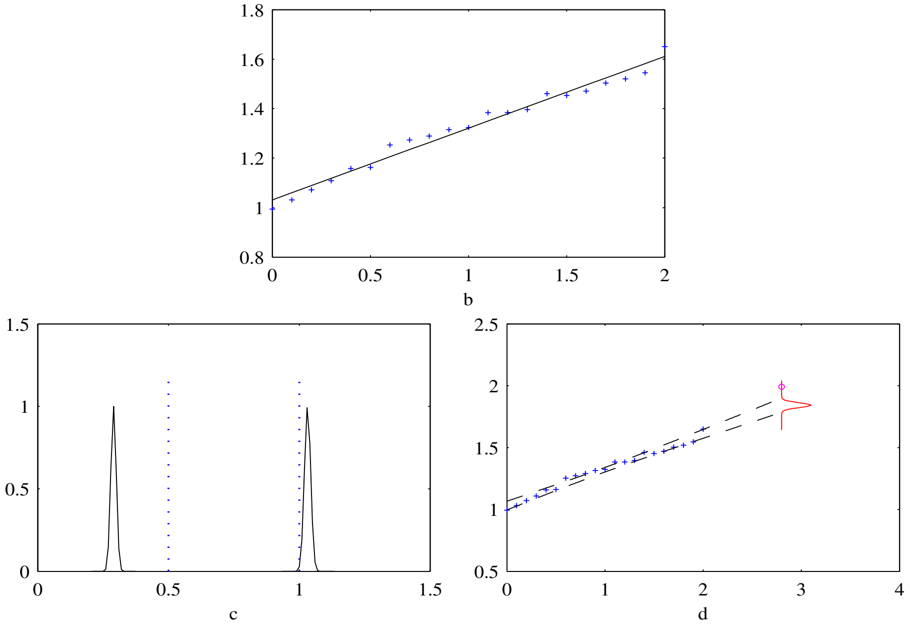
- Lab inspired approach to analysis, over regularising (?), the ‘uniform assumption’, equivalent number of cycles

# Outline

# Example 1: Linear least squares fit with good UQ



# Example 2: Linear least squares fit with poor UQ



# What went wrong?

# What went wrong?

## Example 1

- Model calibrated, i.e. ‘fitted to data’, confidence intervals reasonable ✔
- Parameter uncertainty reasonable, i.e. capture true values ✔
- Predictive uncertainty reasonable, i.e. capture reality ✔

## Example 2

- Model calibrated ✔
- But
  - poor parameter uncertainty, i.e. don’t capture true values ✗
  - poor predictive uncertainty ✗

# What went wrong?

## Example 1

- Model calibrated, i.e. ‘fitted to data’, confidence intervals reasonable ✔
- Parameter uncertainty reasonable, i.e. capture true values ✔
- Predictive uncertainty reasonable, i.e. capture reality ✔

## Example 2

- Model calibrated ✔
- But
  - poor parameter uncertainty, i.e. don’t capture true values ✗
  - poor predictive uncertainty ✗

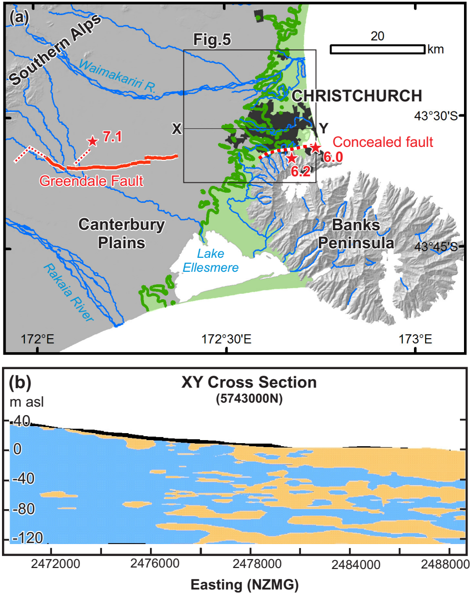
## Distinction

- Good and bad UQ

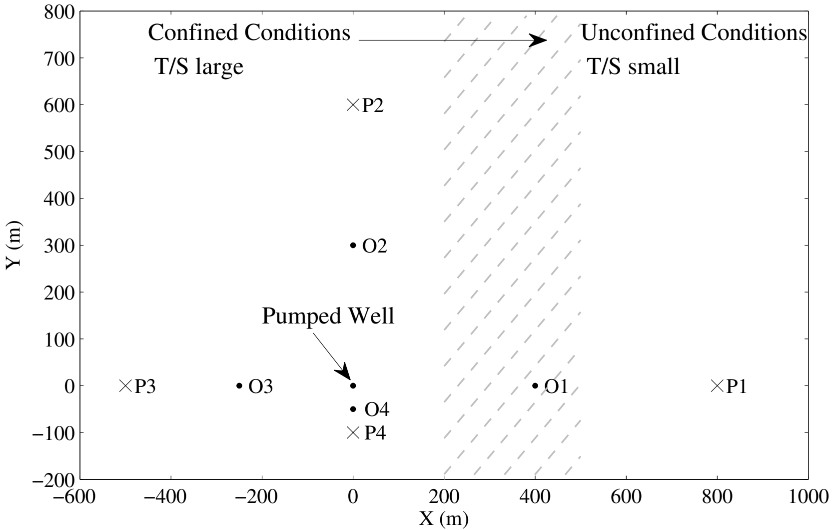
# Outline

# Scenario

# Scenario


(a)
Southern Alps
Waimakariri R.
Fig.5
20
km
CHRISTCHURCH
43°30'S
X
Y
Concealed fault
7.1
6.0
6.2
Greendale Fault
Canterbury
Plains
Banks
Peninsula
43°45'S
Lake
Ellesmere
Rakaia River
172°E
172°30'E
173°E
(b)
XY Cross Section
(5743000N)
m asl
40
0
-40
-80
-120
2472000
2476000
2478000
2484000
2488000
Easting (NZMG)

# Test case


Confined Conditions
T/S large
Unconfined Conditions
T/S small
P2
O2
Pumped Well
P3
O3
O1
P1
O4
P4
Y (m)
X (m)
800
700
600
500
400
300
200
100
0
−100
−200
−600
−400
−200
0
200
400
600
800
1000

# Problem

## Model

$$S(x,y)\frac{\partial s(x,y,t)}{\partial t} = \nabla \cdot (T(x,y)\nabla s(x,y,t)) + Q\delta(x)\delta(y) \quad \text{in } \Omega \qquad (1)$$

$$s(x,y,t) = 0 \quad \text{on } \partial\Omega \qquad (2)$$

$$s(x,y,0) = 0 \qquad (3)$$

where $s = s(x,y,t)$ is drawdown, $T = T(x,y))$ and $S = S(x,y)$ are spatially distributed transmissivities and storativities, and $Q$ is the constant pumping rate over the duration of pumping.

## Problem

- Carry out a pump test and observe drawdown in bores O1-O4.
- Estimate distributed log-transmissivity $\log T = \log T(x,y)$ and log-storativity $\log S = \log S(x,y)$

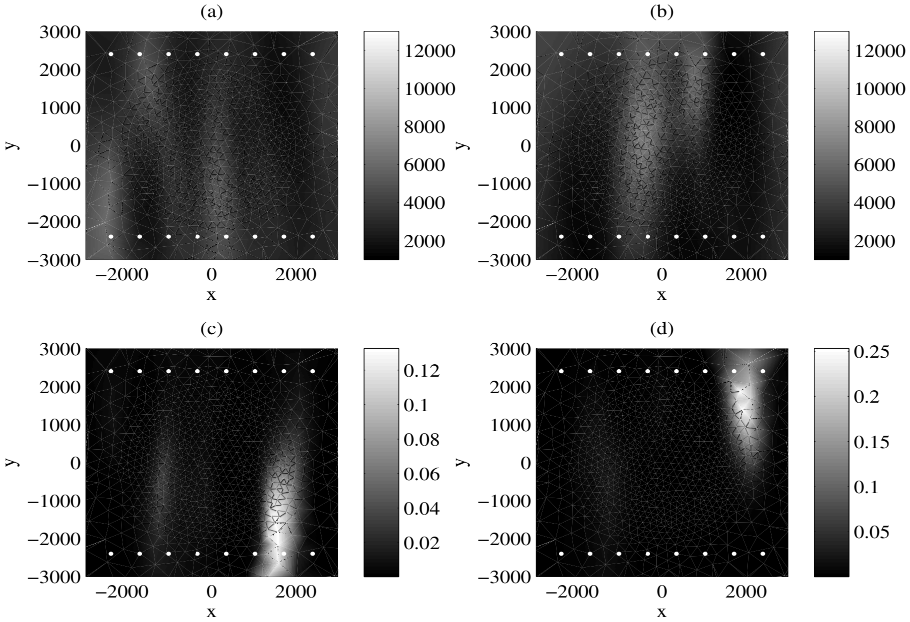
# Reflections

- What can we reasonably expect to reconstruct from the data from bores O1-O4?
- Presumably, there will be a zone where the data tells us something, outside the data will give little/no information.
- What is a sensible prior model?

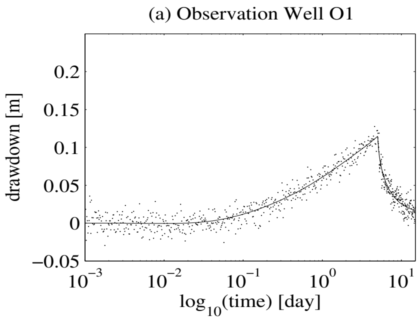
  -what structural information does it make sense to use?

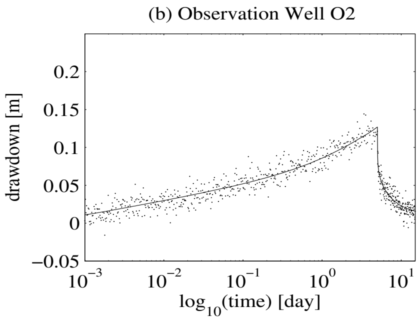
  -what scales can we expect capture?

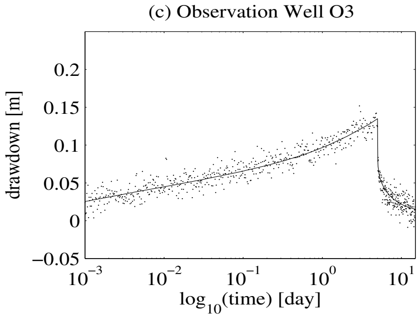
# Draws from prior

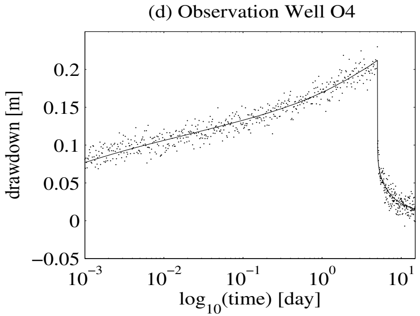


# Data

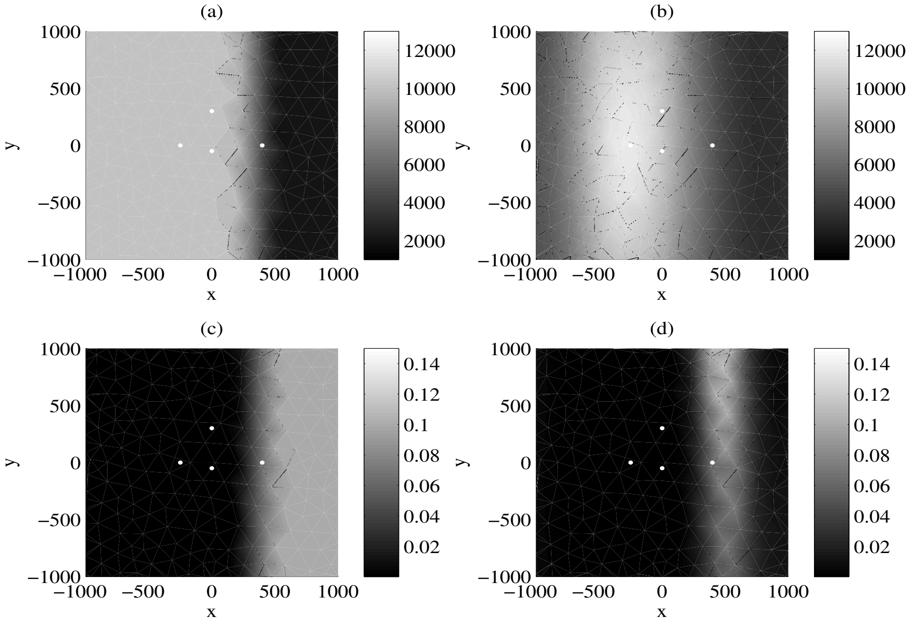

(a) Observation Well O1
drawdown [m]
0.2
0.15
0.1
0.05
0
−0.05
$10^{-3}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$\log_{10}$(time) [day]

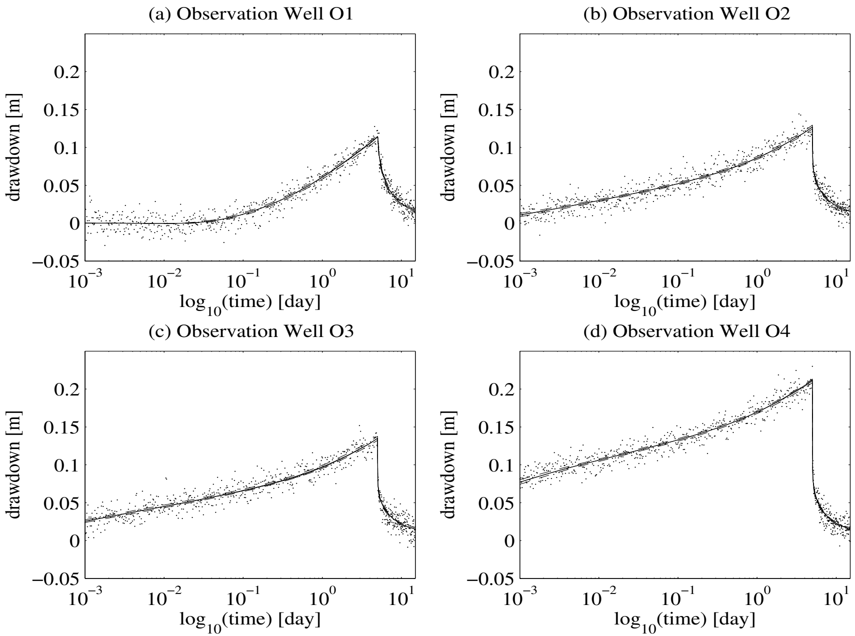

(b) Observation Well O2
drawdown [m]
0.2
0.15
0.1
0.05
0
−0.05
$10^{-3}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$\log_{10}$(time) [day]


(c) Observation Well O3
drawdown [m]
0.2
0.15
0.1
0.05
0
−0.05
$10^{-3}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$\log_{10}$(time) [day]

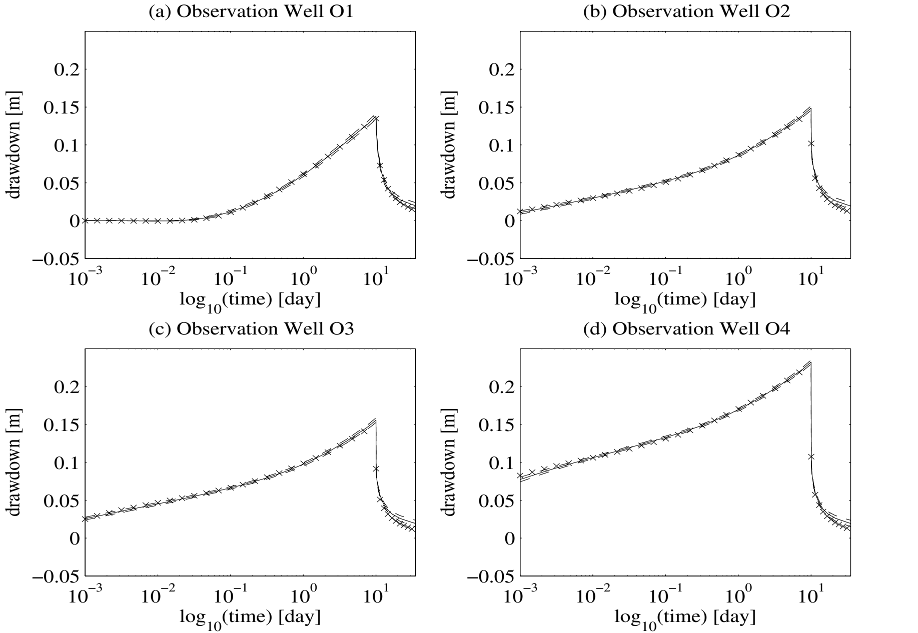

(d) Observation Well O4
drawdown [m]
0.2
0.15
0.1
0.05
0
−0.05
$10^{-3}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$\log_{10}$(time) [day]

# Reconstruction

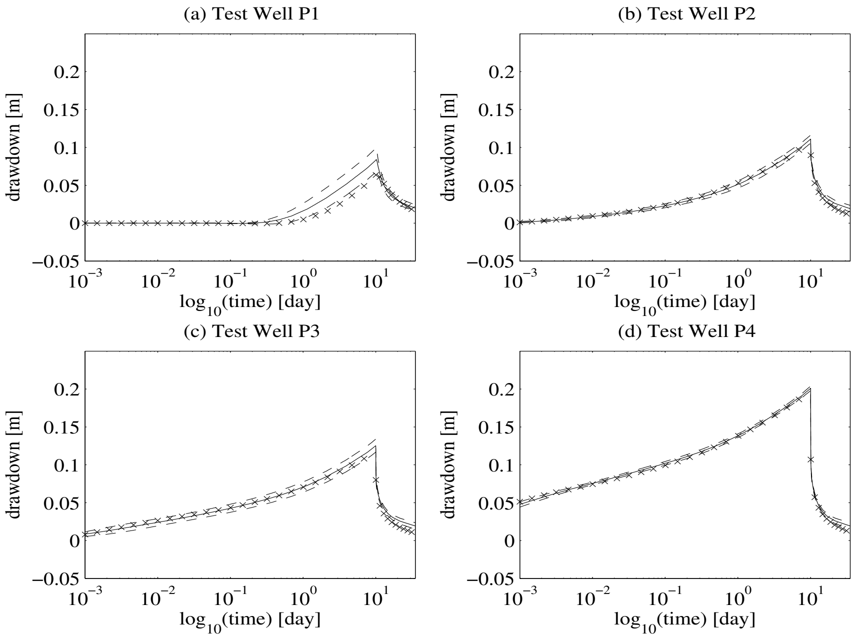



(a) Observation Well O1
(b) Observation Well O2
(c) Observation Well O3
(d) Observation Well O4
drawdown [m]
0.2
0.15
0.1
0.05
0
−0.05
$10^{-3}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$\log_{10}$(time) [day]

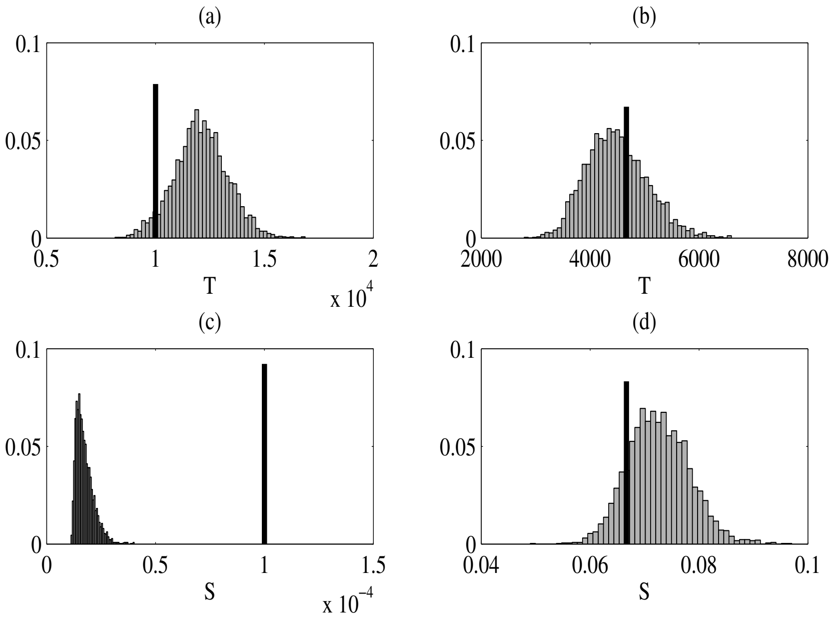
# Prediction at measurement points



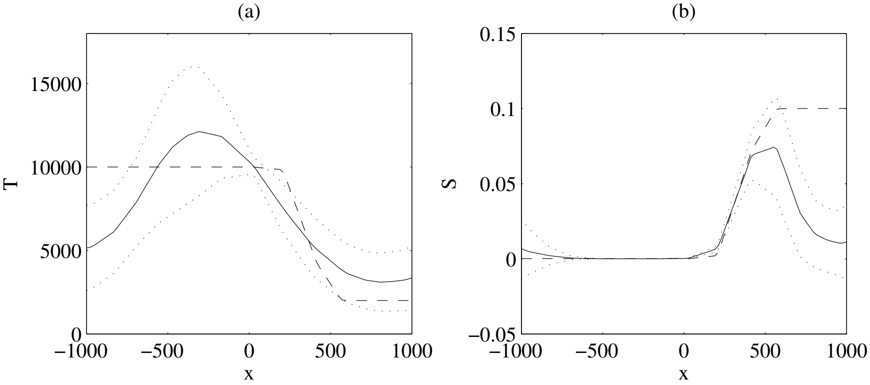
# Prediction at test points P1-P4



# Parameter reconstruction

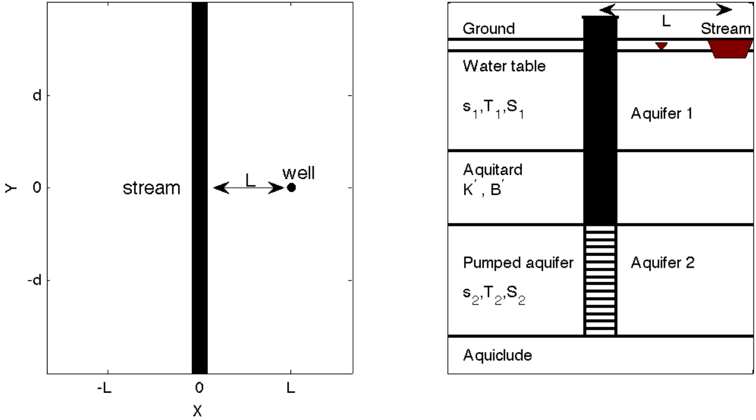


# Longitudinal cut through parameter reconstruction



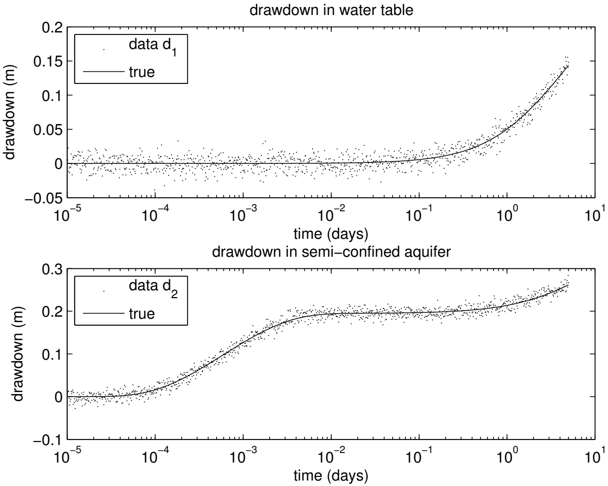
# Outline

# Scenario



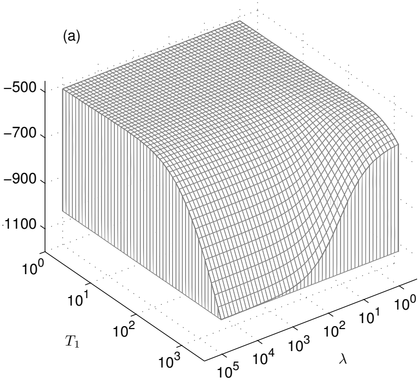
d
0
-d
Y
stream
L
well
-L
0
L
X
Ground
L
Stream
Water table
s1,T1,S1
Aquifer 1
Aquitard
K′, B′
Pumped aquifer
s2,T2,S2
Aquifer 2
Aquiclude

# Pump test data



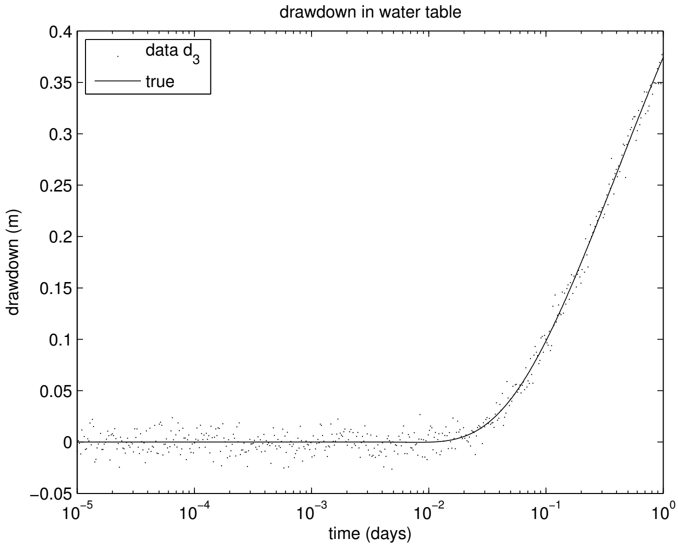

## Quantify stream depletion

# Unbounded uncertainty

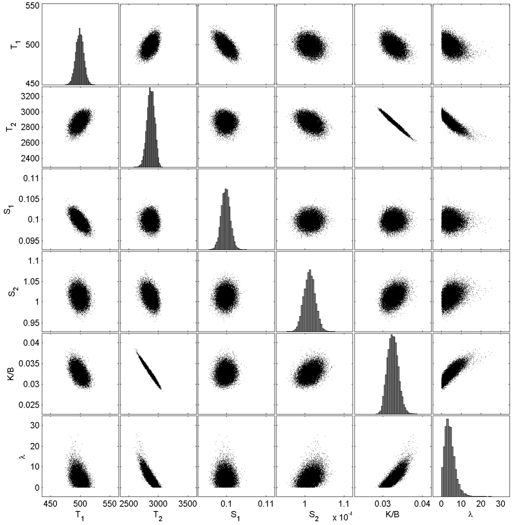



- Plots of log-likelihood as functions of $T_1$ and $\lambda$ (with all other parameters true values)
- (a) uses $d_2$, (b) $d_1$ and $d_2$
- Seems $\lambda$ somewhat ill-determined by data

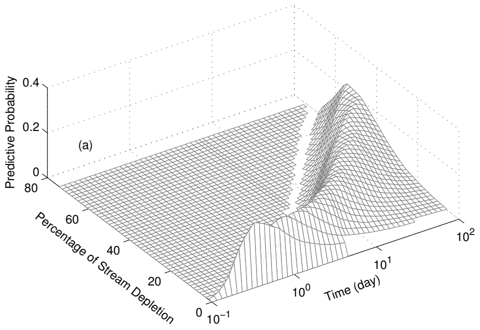
# Data for shallow pumping test

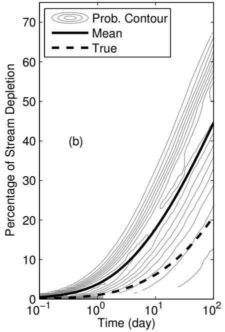


# Posterior parameter uncertainty using data $d_2$ and $d_3$



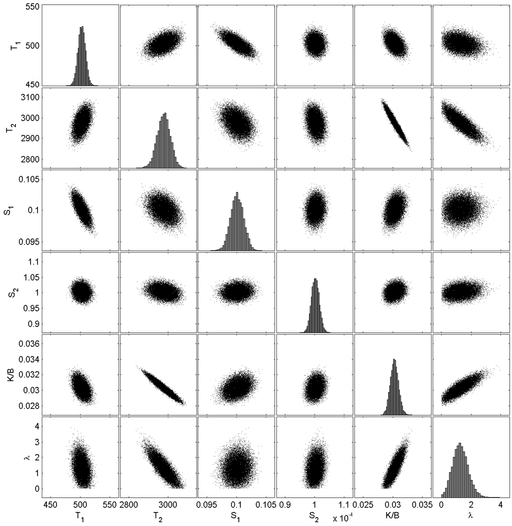
# Uncertainty in stream depletion using data $d_2$ and $d_3$



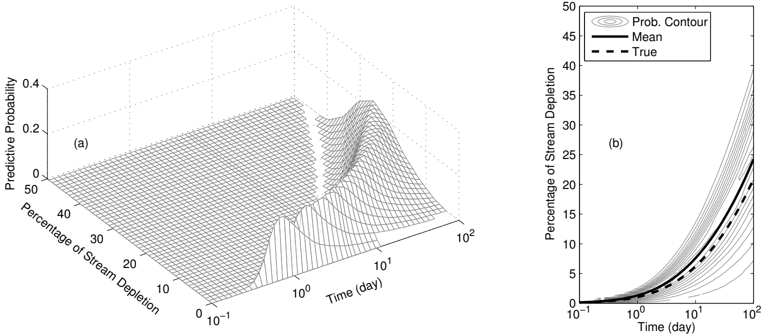




- At least stream depletion bounded
- Though would grossly overestimate

# Posterior parameter uncertainty using all data $d_1 - d_3$



# Uncertainty in stream depletion using all data $d_1 - d_3$




- More acceptable
- Mildly overestimates

# Outline

# Discussion: Key references


- Cui and Dudley Ward (2012), Uncertainty quantification for stream depletion tests, J.Hydr.Eng
- Cui, Dudley Ward & Kaipio (2012), Characterisation of aquifer parameters from pumping test data for a heterogenous aquifer, Under review, J.Hydr.Eng
- Dudley Ward & Kaipio (2012), Uncertainty, decision and control: an introduction, Preprint.
- Dudley Ward (2012), The Book, MS.

# Discussion: Summary

## Why do quantified uncertainty?

- Insufficient consideration of uncertainty leads to poor predictive reliability
- Good UQ essential to decision making, since quantified risks can be controlled
- But: bad UQ is unhelpful at best, positively misleading at worst

# Discussion: Summary

## Why do quantified uncertainty?

- Insufficient consideration of uncertainty leads to poor predictive reliability
- Good UQ essential to decision making, since quantified risks can be controlled
- But: bad UQ is unhelpful at best, positively misleading at worst

## Take home advice:

- Beware the pitfalls of Monte Carlo sampling

# Discussion: Summary

## Why do quantified uncertainty?

- Insufficient consideration of uncertainty leads to poor predictive reliability
- Good UQ essential to decision making, since quantified risks can be controlled
- But: bad UQ is unhelpful at best, positively misleading at worst

## Take home advice:

- Beware the pitfalls of Monte Carlo sampling

## To reflect on:

- Effective public communication of risk
- See `http://understandinguncertainty.org`