Computer Model Calibration: Bayesian Methods for Combining Simulations and Experiments for Inference and Prediction

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

Example

- The Lyon-Fedder-Mobary (LFM) model simulates the interaction of solar wind plasma in the magnetosphere

- We see this as the Aurora Borealis

- Have a computer model that attempts to capture the main features of this phenomenon

- Outputs are large 3-dimensional space-time fields or extracted features thereof

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

Example

- Field observations available from the Polar Ultraviolet Imager satellite

- Computer model is assumed to capture all salient features of the solar wind interactions, up to random error

- However, the inputs ($\theta = (\alpha, \beta, R)$) are not known

- Scientific problem: estimate $\theta = (\alpha, \beta, R)$

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


4 17 30

Basic inverse problem – Statistical formulation

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

$$y_f(\theta) = \eta(\theta) + \epsilon$$

- Where,
  - y_f system response
  - y_s simulator response at input t
  - θ calibration parameters
  - ε random error

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

Problem is to estimate the calibration parameters

Solutions

- Many different solutions to such problems… can you think of one?
- Suppose the computer model is fast?
- Suppose the computer model is slow?

Model calibration – Statistical formulation

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

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

- 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 – computer model and field observations

Problem is to estimate the calibration parameters

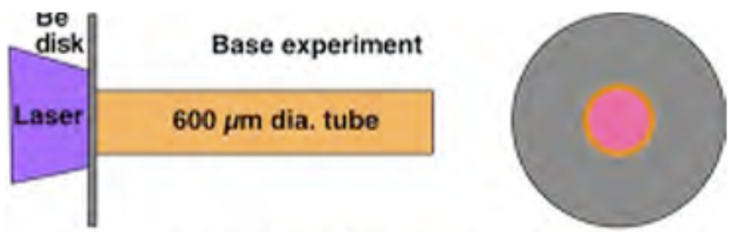
Solutions

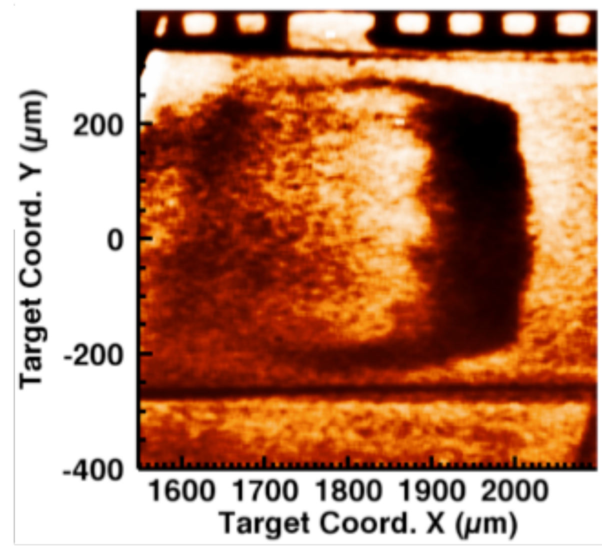
- Many different solutions to such problems… can you think of one?
- Suppose the computer model is fast?
- Suppose the computer model is slow?

Example: Radiative shock experiment

A conceptually simple experiment

– Launch a thin Be plasma down a shock tube at ~ 200 km/s

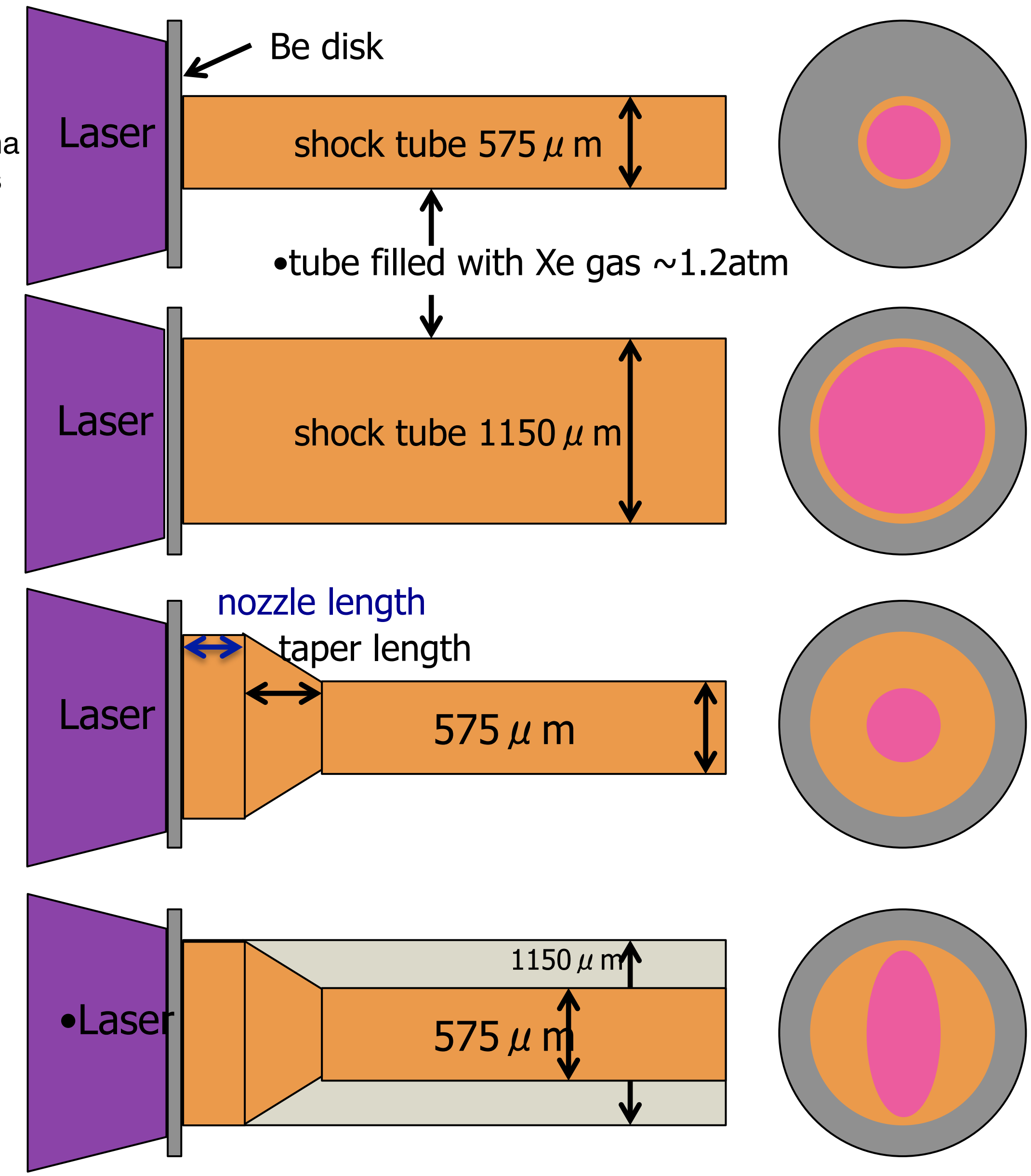




A radiative shock is a wave in which both hydrodynamic and radiation transport physics play a significant role in the shock's propagation and structure

Application

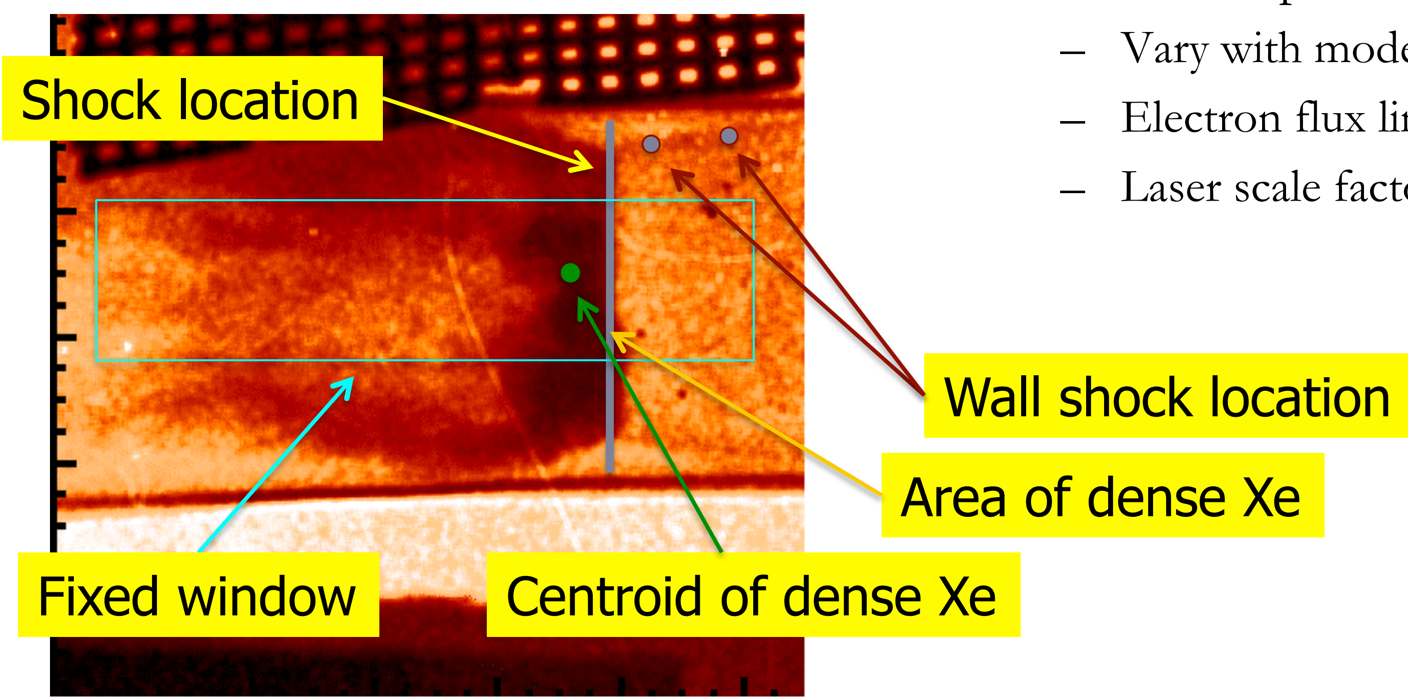
- Initial experiments:
  - 1 ns, 3.8 kJ laser irradiates Be disk➔ plasma down Xe gas filled shock tube at ~ 200 km/s
  - Circular tube; diam = 575µm
  - Timing 13-14 ns
- Additional experiments
  - Laser energy ~ 3.8kJ
  - Circular tube; diam = 575, 1150µm
  - Timing 13,20,26 ns
- Nozzle experiments
  - Laser energy ~ 3.8kJ
  - nozzle length = taper length = 500µm
  - Circular tube; diam = 575µm
  - Timing 26 ns
- Extrapolation experiments (5th year)
  - Laser energy ~ 3.8kJ
  - **Elliptical** tube; diam = 575-1150µm
  - Aspect ratio = 2
  - Includes nozzle in shock tube
  - Timing 26 ns



Have several outputs & inputs

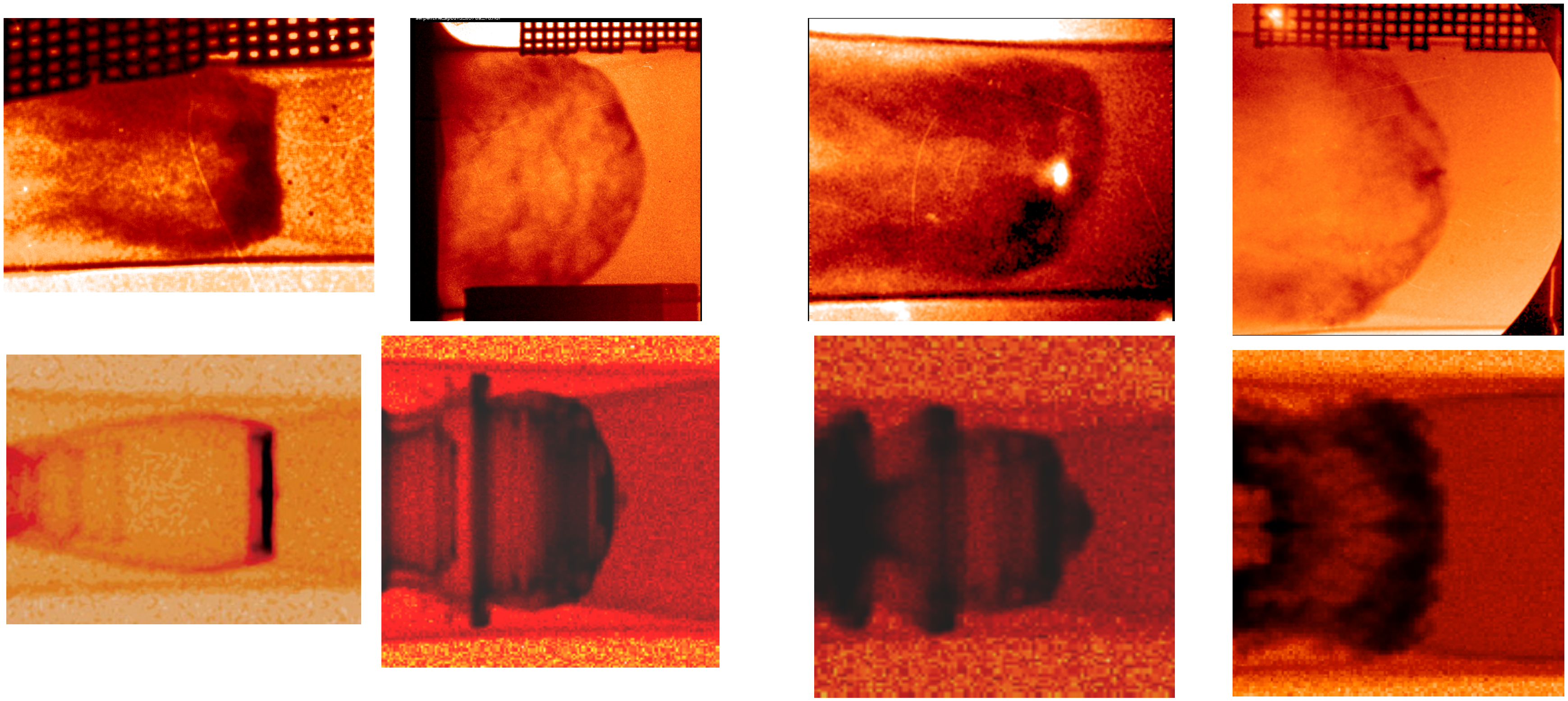
- Outputs (y)
  - Shock location
  - Shock breakout time
  - Wall shock location
  - Axial centroid of Xe
  - Area of dense Xe

- Inputs (x)
  - Observation time
  - Laser energy
  - Be disk thickness
  - Xe gas pressure
  - Tube geometry
- Calibration parameters (θ)
  - Vary with model
  - Electron flux limiter
  - Laser scale factor



We can measure and we can compute

600 μm tube | 1200 μm tube | Circular nozzle | Elliptical nozzle



Goal is to predict elliptical tube quantities of interest and uncertainty, without using any data from elliptical tube experiments

- Shocks at 13 ns

Data

- Have observations from 1-D CRASH model and experiments

- Experiment data:
  - 9 experiments
  - experiment variables: Be thickness, Laser Energy, Xe pressure and Time
  - responses: Shock location

- 1-D CRASH Simulations
  - 320 simulations, varied over 8 inputs
  - experiment variables: Be thickness, Laser Energy, Xe pressure and Time
  - calibration parameters: Be Gamma, Be OSF, Xe Gamma, Xe OSF
  - response: Shock location

Use observations and simulations for prediction in real world

- Have simulations from the CRASH code
- Have observations from laboratory experiments
- Want to combine these sources of data to make prediction of the real-world process ... Also have to estimate the calibration parameters

- Additional complicating factor, the computer model is not an exact representation of the mean of the physical system… darn!

- **Approach:** Model calibration (Kennedy and O'Hagan, 2001; Higdon et al., 2004)

Model calibration – Statistical formulation

$$
\begin{aligned}
y_s(x,t) &= \eta(x,t) \\
y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon
\end{aligned}
$$

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

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

Model calibration – Statistical formulation

$$y_s(x,t) = \boxed{\eta(x,t)}$$
$$y_f(x,\theta) = \boxed{\eta(x,\theta)} + \delta(x) + \varepsilon$$

Shared signal

Model calibration – Statistical formulation

$$
\begin{aligned}
y_s(x,t) &= \eta(x,t) \\
y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon
\end{aligned}
$$

Discrepancy

Model calibration – Statistical formulation

$$
\begin{aligned}
y_s(x,t) &= \eta(x,t) \\
y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon
\end{aligned}
$$

Gaussian process models

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

Back to Gaussian process models

- $y(x_i) = \mu + z(x_i)$

- $z(x_i) \sim N(0, \sigma_z^2)$
- $cor\left((z(x_i), z(x_j)\right) = e^{-\sum_{k=1}^{d} \theta_k (x_{ik} - x_{jk})^2}$
- $z(x) \sim N(0_n, \sigma_z^2 R)$
- $y(x) \sim N(\mu 1_n, \sigma_z^2 R)$

Gaussian process model for the emulator in this setting

- $y_s(x_i, t_i) = \mu + z(x_i, t_i)$

- $z(x_i, t_i) \sim N(0, \sigma_z^2)$
- $cor\left((z(x_i), z(x_j)\right) = e^{-\sum_{k=1}^{d} \theta_k (x_{ik} - x_{jk})^2 - \sum_{k=1}^{d'} \omega_k (t_{ik} - t_{jk})^2}$
- $z(x, t) \sim N(0_n, \sigma_z^2 R)$
- $y(x, t) \sim N(\mu 1_n, \sigma_z^2 R)$

Have Gaussian process model for the discrepancy

- $\delta(x_i) \sim N(0, \sigma_\delta^2)$
- $cor\left((\delta(x_i), \delta(x_j)\right) = e^{-\sum_{k=1}^{d} \gamma_k (x_{ik} - x_{jk})^2}$
- $\delta(x) \sim N(0_m, \sigma_\delta^2 R_\delta) = N(0_m, \Sigma_\delta)$

Hierarchical model is used to combine simulations and observations

- View computational model as a draw of a random process (like before)
- Denote vectors of simulation trials as and field measurements as y_s & y_f
- Suppose that these are *n* and *m*-vectors respectively

$\longrightarrow$

- Can combine sources of information using a single GP

$$\begin{aligned} y_s(x,t) &= \eta(x,t) \\ y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon \end{aligned}$$

$$y = \begin{pmatrix} y_s \\ \\ y_f \end{pmatrix} \sim N\left(\mu,\ \Sigma_\eta + \Sigma_\delta + \Sigma_\varepsilon\right)$$

Hierarchical model is used to combine simulations and observations

- View computational model as a draw of a random process
- Denote vectors of simulation trials as and field measurements as y_s & y_f
- Suppose that these are *n* and *m*-vectors respectively

- Can combine sources of information using a single GP

$$
\begin{aligned}
y_s(x,t) &= \eta(x,t) \\
y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon
\end{aligned}
$$

$$
y = \begin{pmatrix} y_s \\ \\ y_f \end{pmatrix} \sim N\left(\mu, \Sigma_\eta + \Sigma_\delta + \Sigma_\varepsilon\right)
$$

Contains correlations between all sources of data via the joint signal in the computer model

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

Hierarchical model is used to combine simulations and observations

- View computational model as a draw of a random process
- Denote vectors of simulation trials as and field measurements as y_s & y_f
- Suppose that these are *n* and *m*-vectors respectively

- Can combine sources of information using a single GP

$$\begin{aligned} y_s(x,t) &= \eta(x,t) \\ y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon \end{aligned}$$

$$y = \begin{pmatrix} y_s \\ y_f \end{pmatrix} \sim N\left(\mu, \Sigma_\eta + \Sigma_\delta + \Sigma_\varepsilon\right)$$

Problem… correlations involving field trials have $t = \theta$

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

Hierarchical model is used to combine simulations and observations

- View computational model as a draw of a random process
- Denote vectors of simulation trials as and field measurements as
  y_s & y_f
- Suppose that these are *n* and *m*-vectors respectively

- Can combine sources of information using a single GP

$$\begin{aligned} y_s(x,t) &= \eta(x,t) \\ y_f(x,\theta) &= \eta(x,\theta) + \delta(x) + \varepsilon \end{aligned}$$

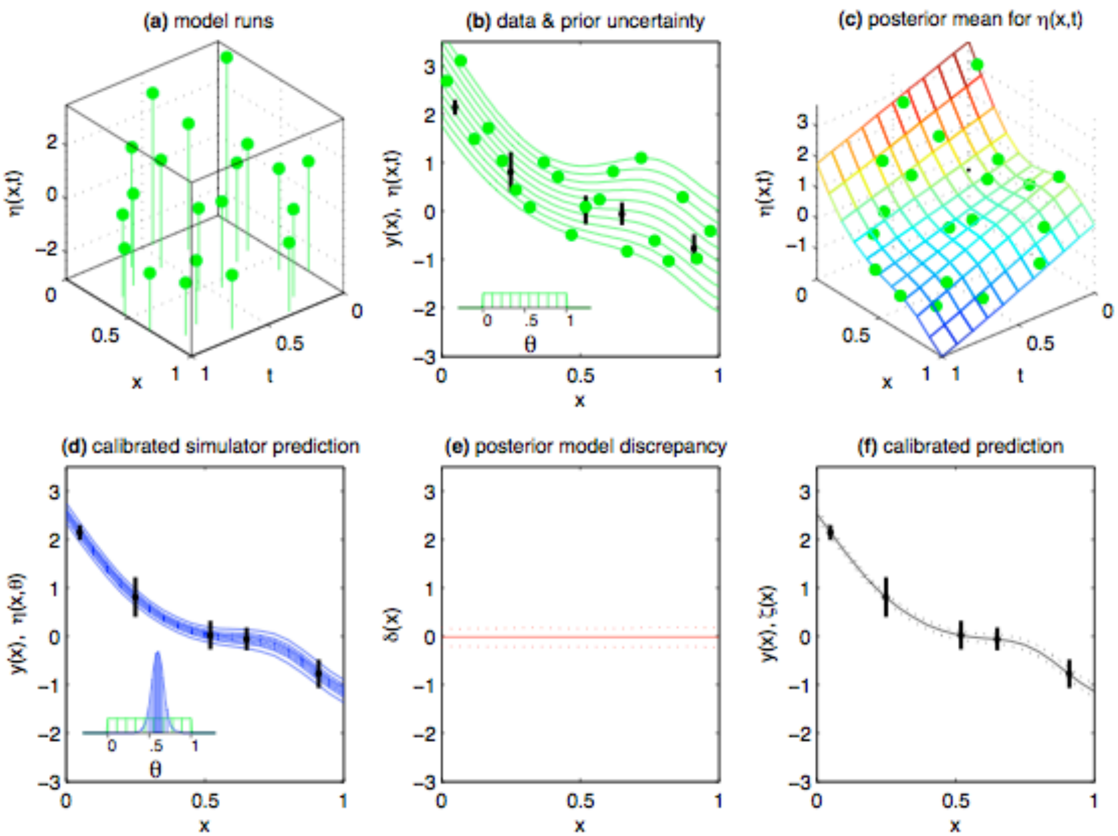
$$y = \begin{pmatrix} y_s \\ \\ y_f \end{pmatrix} \sim N\left(\mu,\ \Sigma_\eta + \Sigma_\delta + \Sigma_\varepsilon\right)$$

Only operates on the field data

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

Calibration idea: No discrepancy model

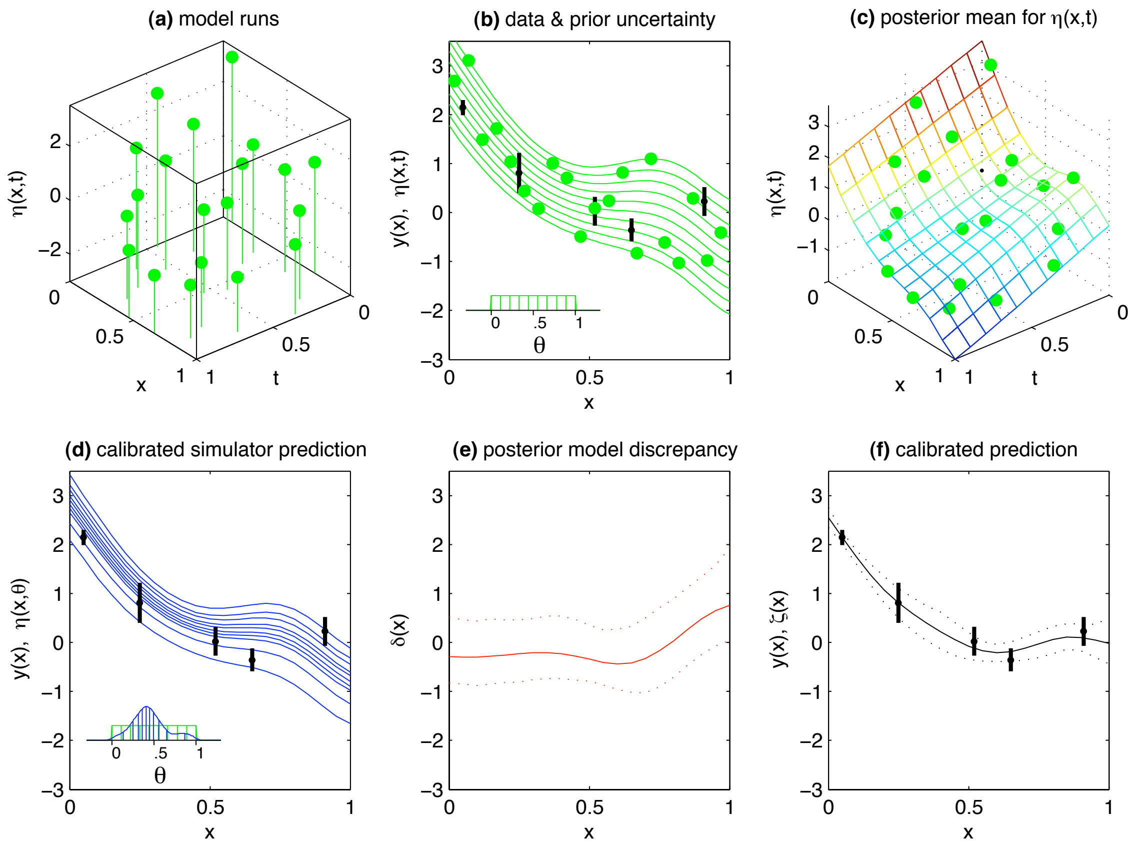
(Plot taken shamelessly from Dave Higdon)



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

Calibration idea: Discrepancy model

(Plot taken shamelessly from Dave Higdon)



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

Comments

- Interpretation of unknown constants can change
- Without discrepancy: estimating unknown physical constants
- With discrepancy: selecting tuning constants that best fit the model

Estimation

- What parameters do we have to estimate?

Estimation

- Can apply Bayes' Rule:

$$[A|B] = \frac{[B|A][A]}{[B]}$$

Ignoring statistical model parameters for the moment…

- In the unbiased case for a fast simulator,

$$[\theta|\eta, y_f] \propto [y_f|\eta, \theta][\eta|\theta][\theta]$$

- For the simulators we have been considering, use an emulator in place of computer model… basic idea

$$[\eta|\theta] \to [\hat{\eta}|\theta, y_s]$$

$$[\theta|\hat{\eta}, y_f] \propto [y_f|\hat{\eta}, \theta][\hat{\eta}|\theta, y_s][\theta]$$

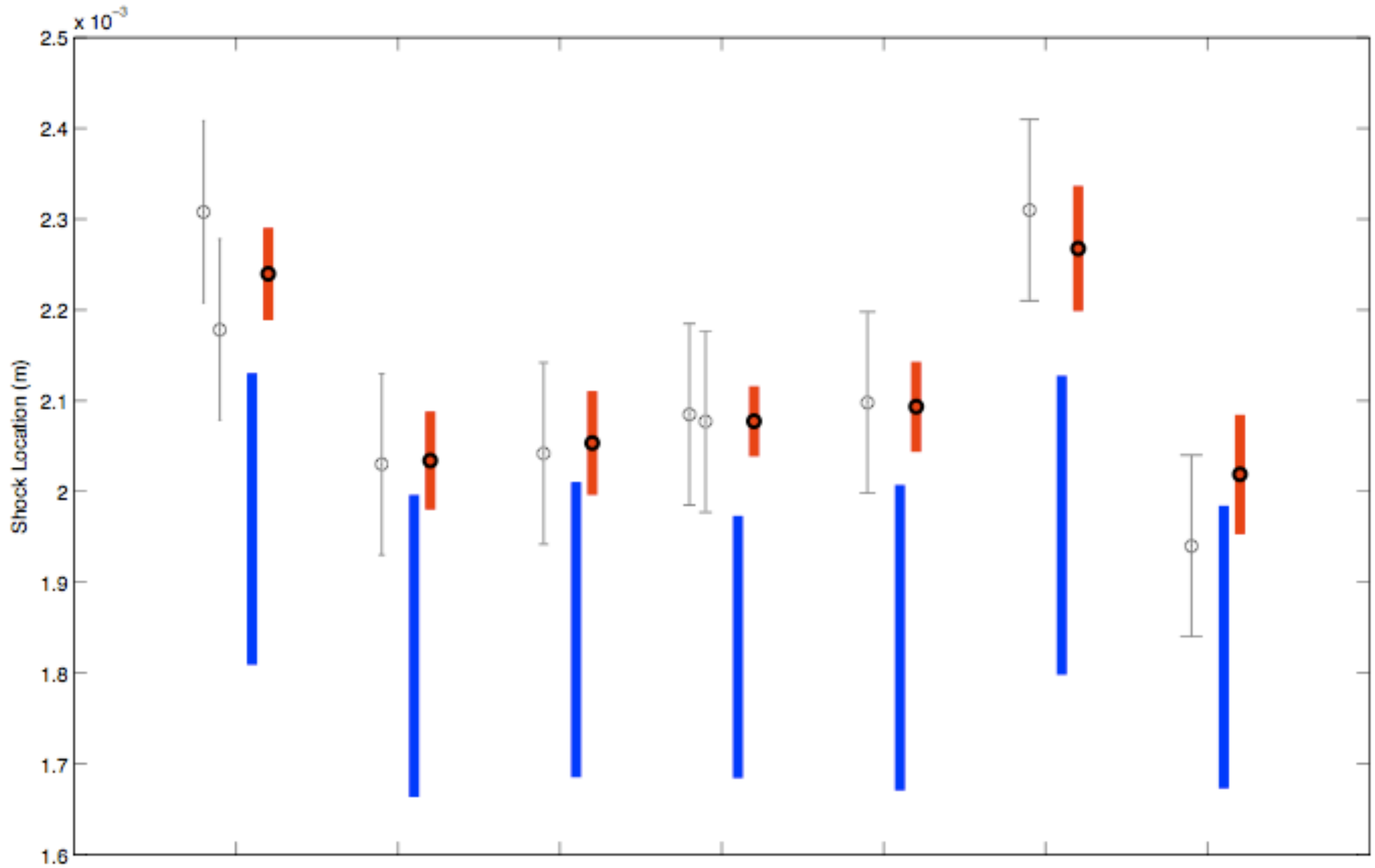
Estimation

- How could we put this together for the CRASH problem?

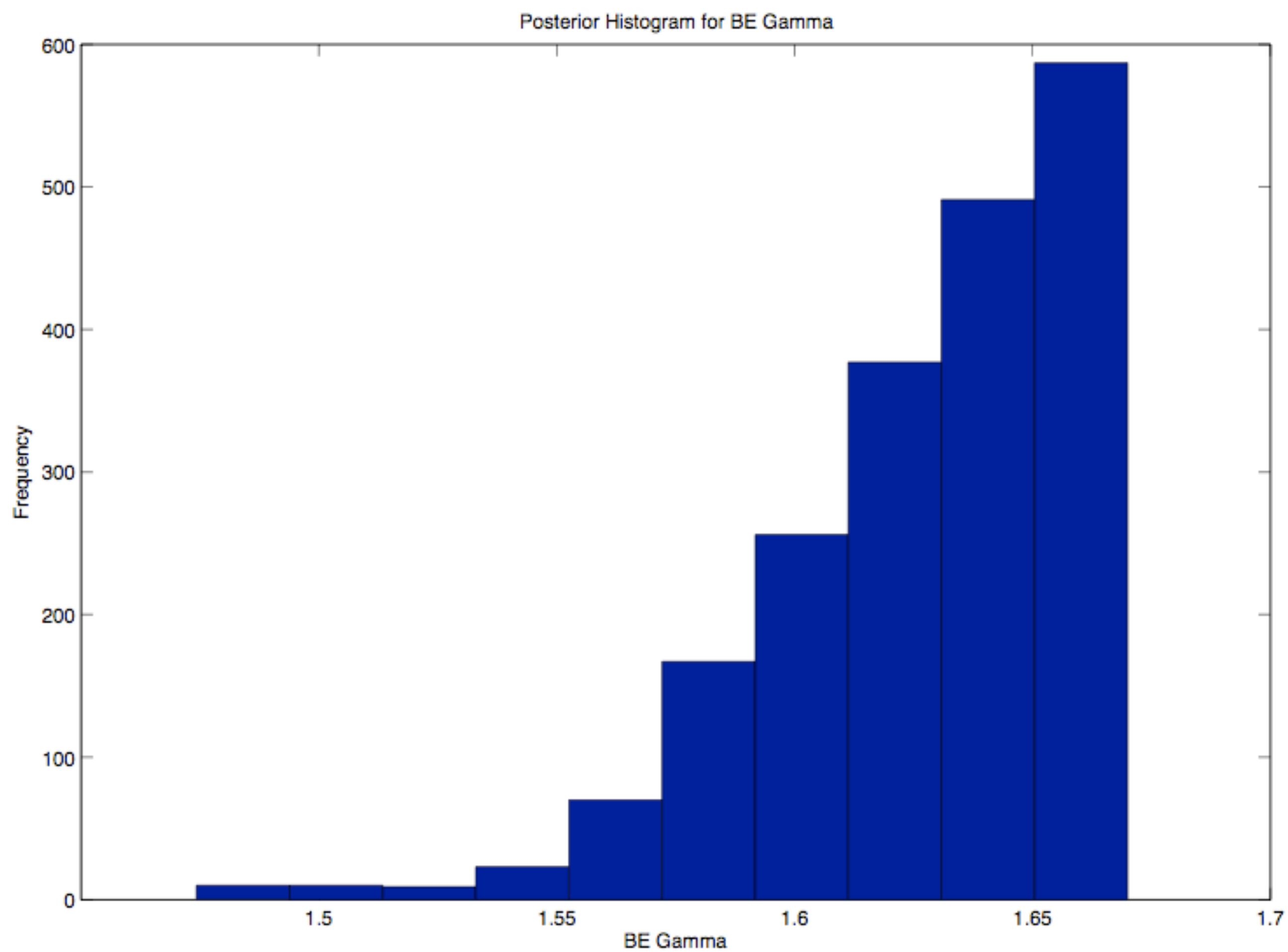
Back to CRASH

- Inverted-gamma priors for variance components
- Gamma priors for the correlation parameters
- Log-normal priors for the calibration parameters
- Samples from the posterior are generated through MCMC… anyone know how to do this?

CRASH predictions

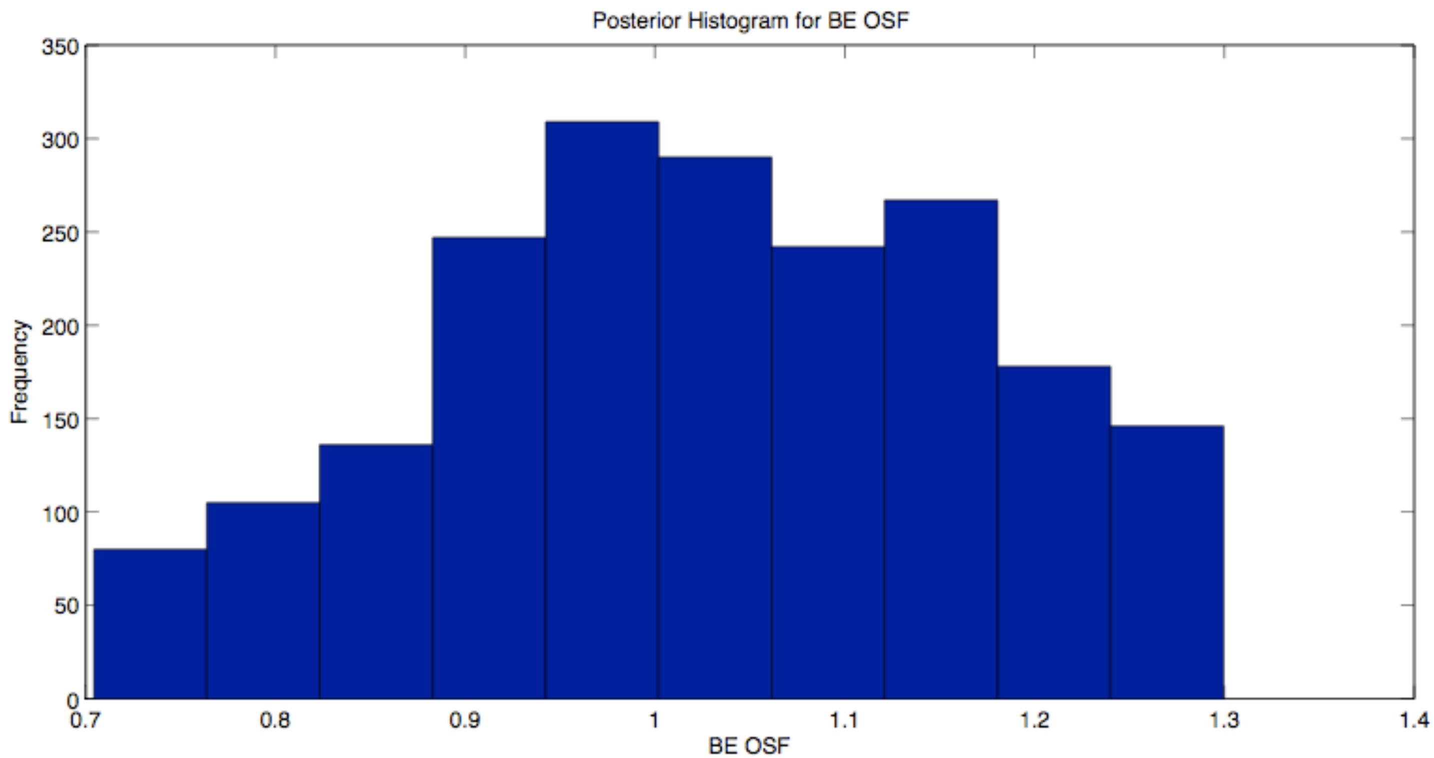


Calibration - BE Gamma

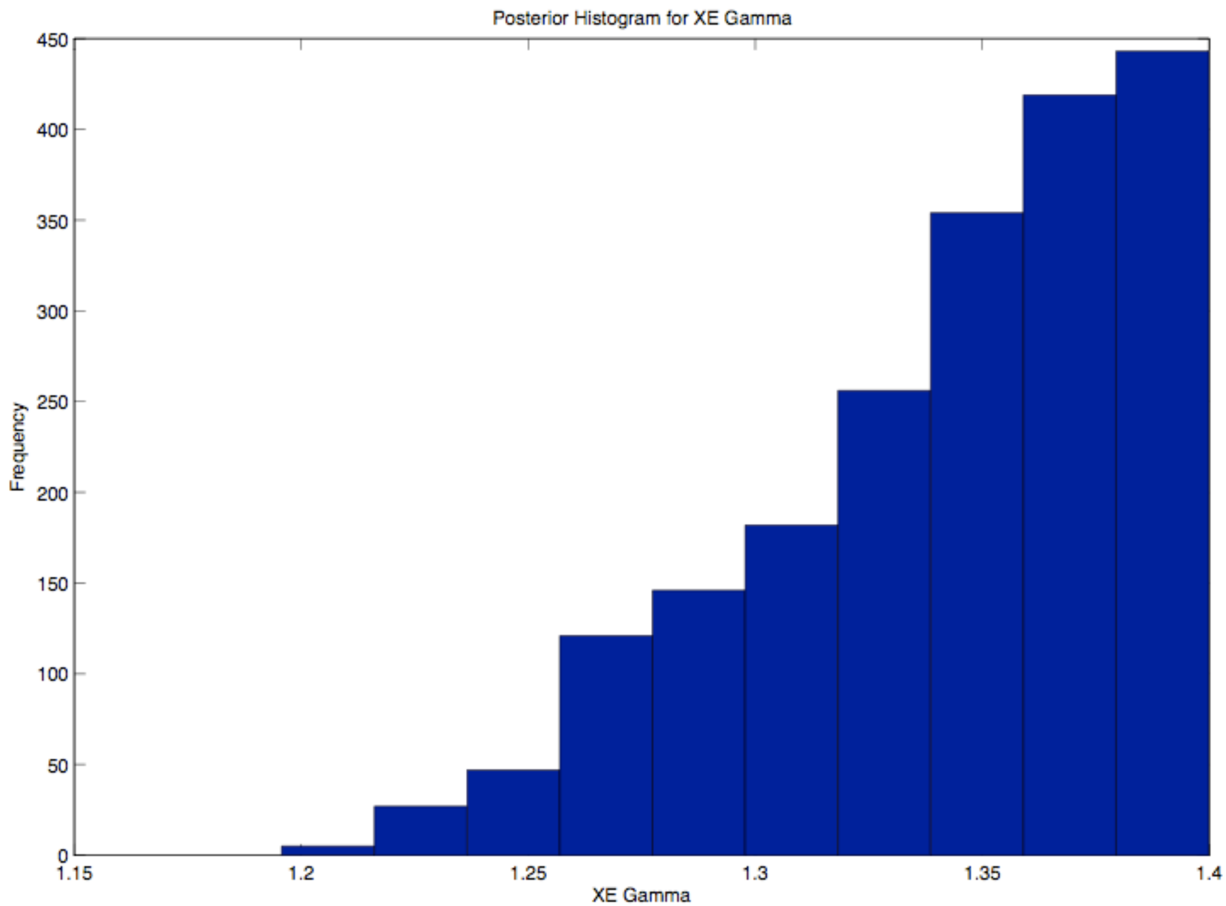


SFU
Department of Statistics and Actuarial Science
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Calibration - BE OSF

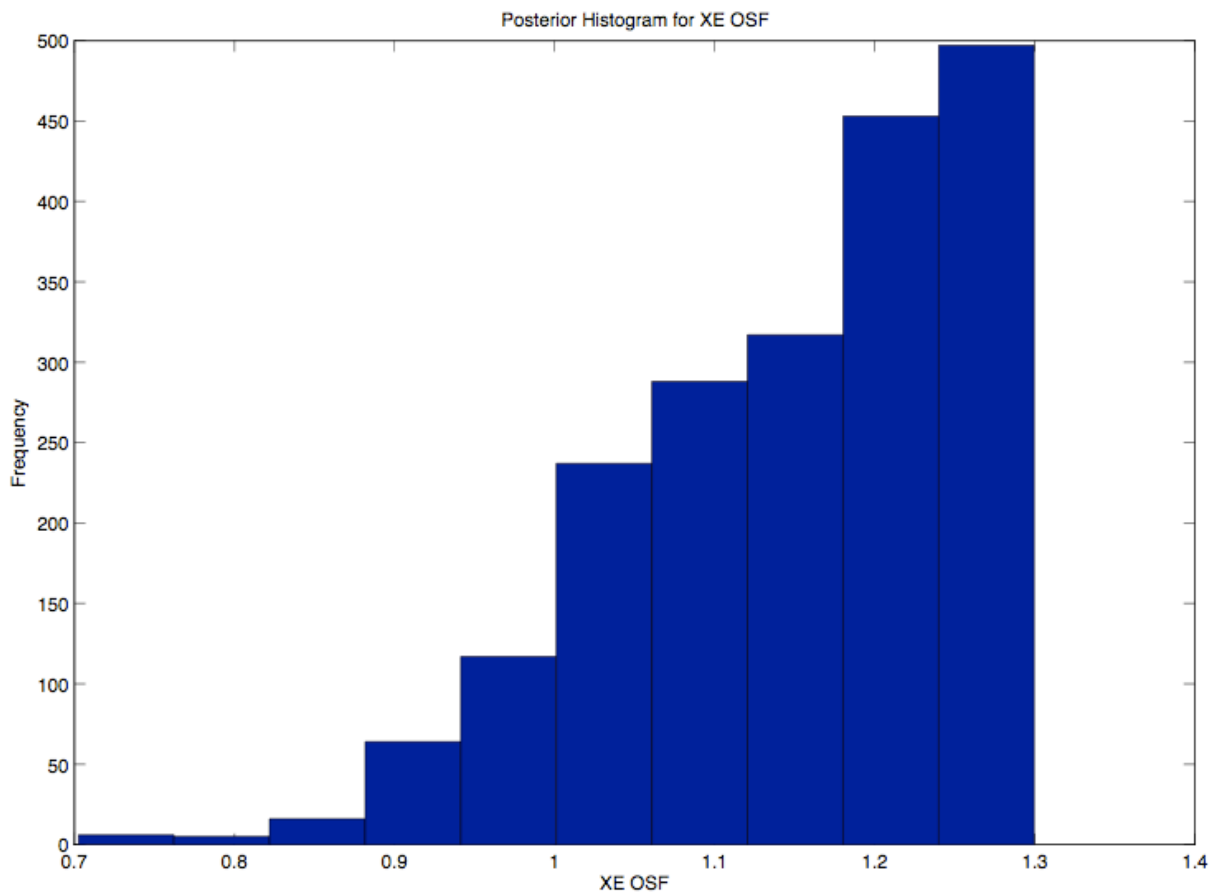


Calibration - XE Gamma



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

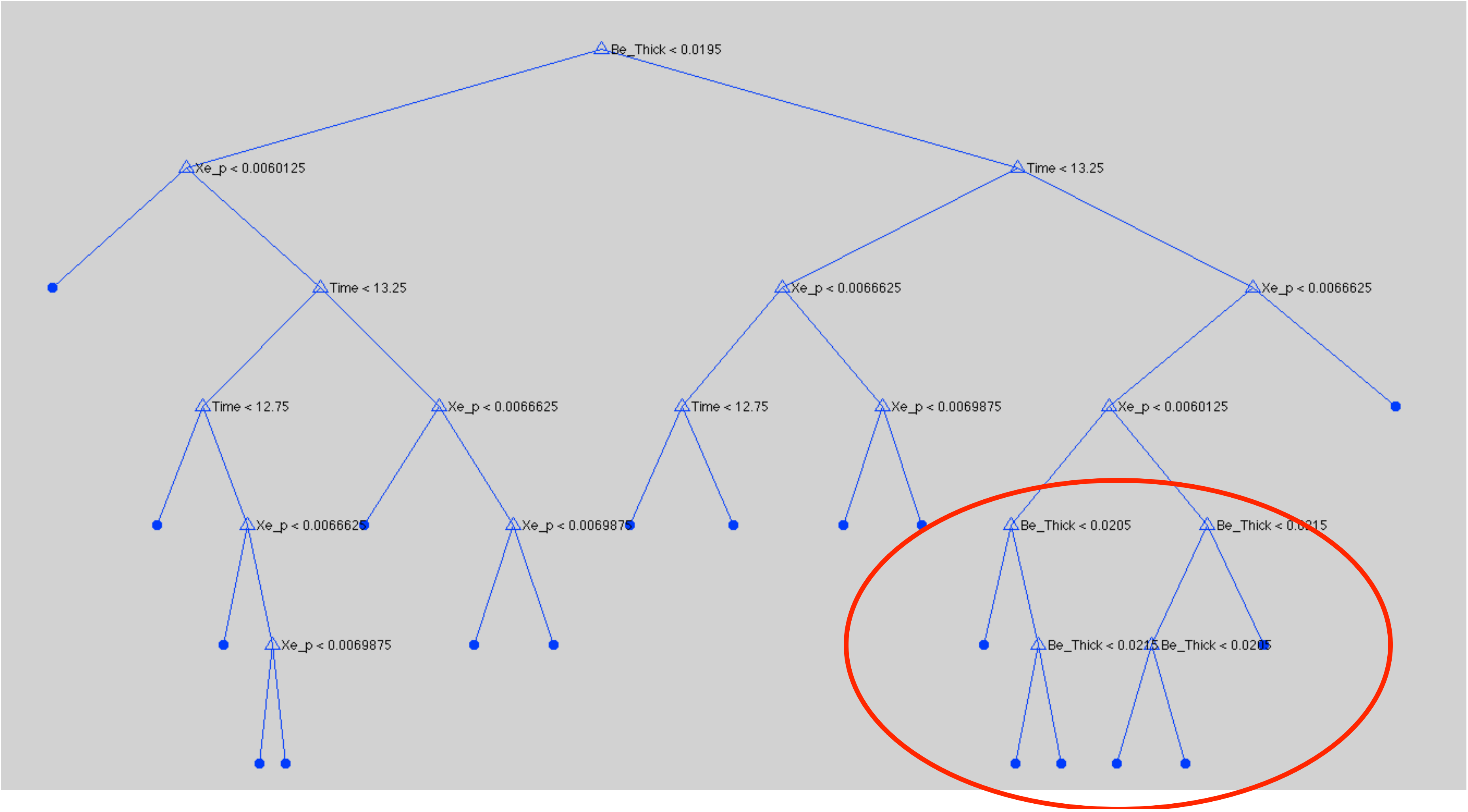
Calibration - XE OSF



Exploring the discrepancy

- The CRASH project was an ongoing endeavor
- Computer codes were being developed
- Idea is to use the statistical model to inform code development
- Predicted the discrepancy over a 4-d grid
- Used posterior mean surface

Discrepancy



- Reasons for the discrepancy
  - It was interesting that the discrepancy is always positive
  - Incorrect radial loss of energy results in Xe that is too hot in front of the shock, and that systematically messes up the shock speed
  - The region where the discrepancy is highest is thought to be the region where Xe pressure is largely going to impact
  - Why might this be a crazy approach to explore the discrepancy?