# Bayesian Calibration of Computer Model Ensembles

*Calibrating the Community Ice Sheet Model*

M. T. Pratola[1], S. Price[2], J. Gattiker[1] and D. Higdon[1]

[1]Statistical Sciences Group, Los Alamos National Laboratory

[2]COSIM, Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory

September 9, 2011

# Outline

- Introduction: Community Ice Sheet Model (CISM)
- Statistical Calibration
- Calibrating Computer Model Ensembles
- Early CISM results, future directions

# CISM

- Historically, ice sheets thought to respond slowly to short-term climate change
- However, recent observations indicate significant ice sheet volume changes as a result of decadal-scale climate forcing
- Potential changes in discharge from Greenland and/or Antarctic ice sheets are the largest unknown w.r.t. future sea-level rise
- CISM describes ice sheet evolution (velocities, thickness, temperature, etc.) assuming appropriate boundary and initial conditions and atmospheric and oceanic forcing (e.g., from CESM).

# Goals

- Our goal is to leverage a statistical model calibration framework to better understand and quantify uncertainties in ice-sheet evolution as simulated by CISM
- Here, we investigate idealized scenarios with uncertainties in:
  - Flow law exponent, $n$
  - Flow law rate factor activation energy, $Q$

# Experimental Setup

- Modified version of the standard “confined shelf” test case:
  - isothermal, rectangular shelf of uniform thickness
  - confined at upstream and lateral margins (zero flux bc)
  - open to the ocean at downstream margin (specified stress bc)
- Constant and steady surface mass balance
- All experiments evolve to approximate SS from $t=0$ to $t=1000$ yrs
- Calibrate to SS reached at $t=1000$ yrs

CISM Confined Shelf Example

# CISM Experiment

- Represent CISM as $\eta(\theta)$ where $\theta$ is some parameter vector of interest
- Ensemble of CISM runs $\{\chi_i = \eta(\theta_i)\}$ at different $\theta_i$'s to get ensemble of outputs
- Choose a "true" value $\theta_0$ to simulate a field observation. The observation, $y$, is constructed as $\eta(\theta_0)+$ error.
- Select 8 settings for $\theta = (n, Q) \in ([1.5, 4.0], [4e4, 8e4])$. $(n_0, Q_0) =?$

# Calibration Experiments

- Real world process too difficult/expensive/... to observe/model directly
- Computer model that describes the real world process given some parameters. Still expensive to run.
- Calibration: Combine computer model runs and limited field observations
  - to estimate computer model parameters
  - to construct predictions of real world process
  - and quantify sources of uncertainty
- Popular approaches - eg. Kennedy & O'Hagan (2001), Higdon et al. (2004), amongst others.

# Calibration Experiments

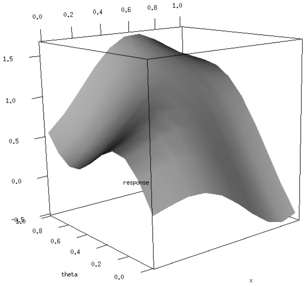
$$\text{Model: } y = \eta(\theta_0) + \epsilon \quad ; \quad \chi_i = \eta(\theta_{\mathbf{i}}) \quad ; \quad x_i = (\theta_i, \chi_i)$$

- We have computer model outputs $x_1, \ldots, x_m$ and field observations $y_1, \ldots, y_n$
- Interested in estimating the unknown parameters $\theta_0$ and the unknown state $\chi = \eta(\theta_0)$
- In a Bayesian framework, the posterior is

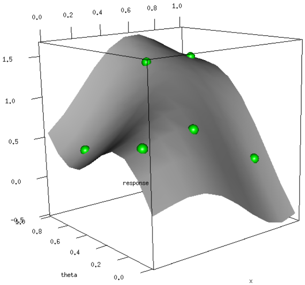
$$\theta_0, \sigma_\epsilon^2, \chi, \bullet | y_1, \ldots, y_n, x_1, \ldots, x_m$$

$$\propto L(y_1, \ldots, y_n, x_1, \ldots, x_m | \cdot) \pi(\theta_0, \sigma_\epsilon^2, \chi, \bullet)$$

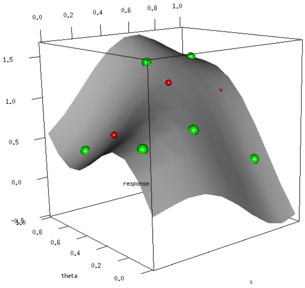
# The Idea (in pictures)



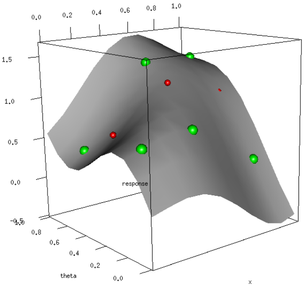
# The Idea (in pictures)

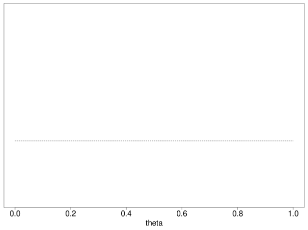


# The Idea (in pictures)

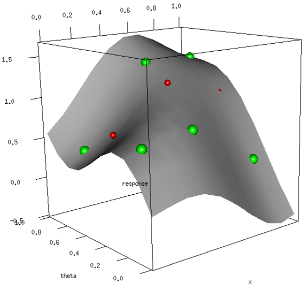


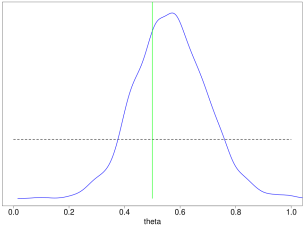
# The Idea (in pictures)



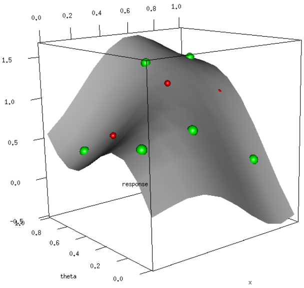



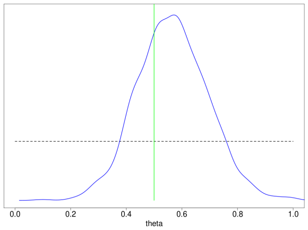
# The Idea (in pictures)

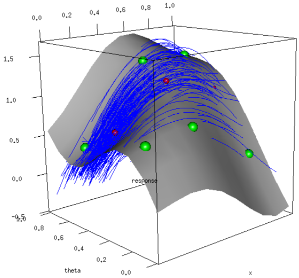


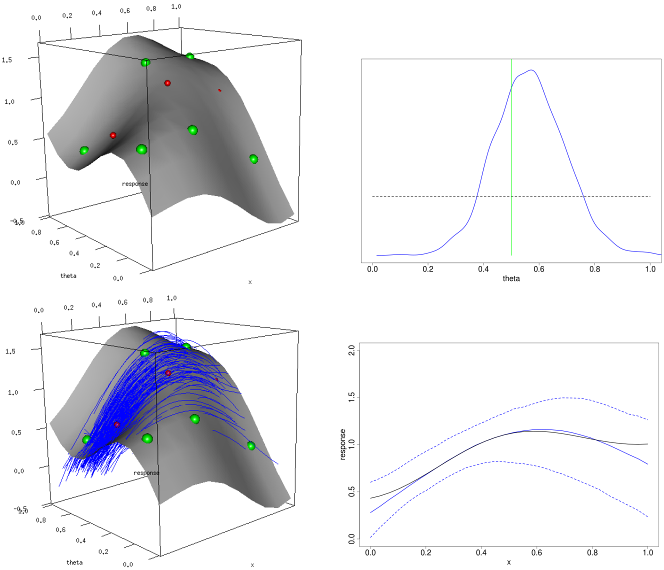
# The Idea (in pictures)



# The Idea (in pictures)



# Bayesian Calibration of Ensembles

- Model: $y = Hx + \epsilon$
- $H$ is an operator that maps the augmented state vector to the observations, in the simplest case, $Hx = \chi$
- Model observational errors as $\epsilon \sim n(0, \lambda_y^{-1} I)$
- Model $x$'s as i.i.d. draws from $n(\mu_x, \Sigma_x)$
- Conjugate priors:
  - $\pi(\lambda_y) \sim g(a, b)$
  - $\pi(\Sigma_x) \sim IW(\nu_0, \Lambda_0^{-1})$
  - $\pi(\mu_x | \Sigma_x) \sim n(\mu_0, \frac{\Sigma_x}{k_0})$

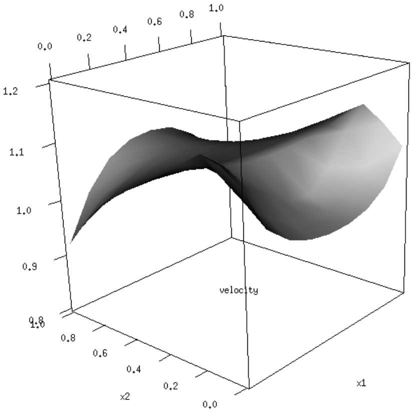
# Bayesian Calibration of Ensembles

- Can then derive the full conditionals,
  - $\pi(\theta_0|\cdot) \sim n(\mu_{\theta_0|\cdot}, \Sigma_{\theta_0|\cdot})$
  - $\pi(\chi|\cdot) \sim n(\mu_{\chi|\cdot}, \Sigma_{\chi|\cdot})$
  - $\pi(\mu_x, \Sigma_x|\cdot) \sim niw(\mu_m, k_m, \nu_m, \Lambda_m)$
  - $\pi(\lambda_y|\cdot) \sim g(\tilde{a}, \tilde{b})$
- sample the joint posterior using Gibbs.
- For convenience, we typically make the approximation $\pi(\mu_x, \Sigma_x|x, x_1, \ldots, x_m) \approx \pi(\mu_x, \Sigma_x|x_1, \ldots, x_m)$.

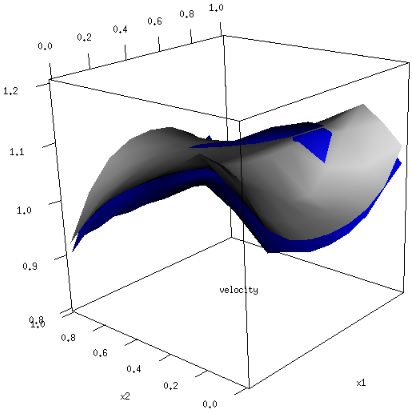
# Calibrating CISM

- Apply our approach to CISM
  - Have $m = 8$ model runs forming our ensemble (outputs normalized to $n(0,1)$)
  - Two parameters to estimate, $\theta = (n, Q)$
  - State vector $\chi$ represents the approximate SS spatial velocity field at the $t = 1000$ time step of the model
  - Simulate $n = 4$ field observations constructed from a model run at $\theta_0 = (3, 6e5)$ with added i.i.d. normal error
  - Use non-informative priors $(\mu_0 = 0, \Lambda_0 = I, \nu_0 = 10, k_0 = 1, a = 25, b = .25)$

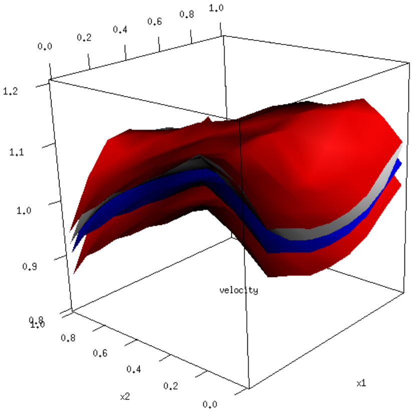
# CISM Results



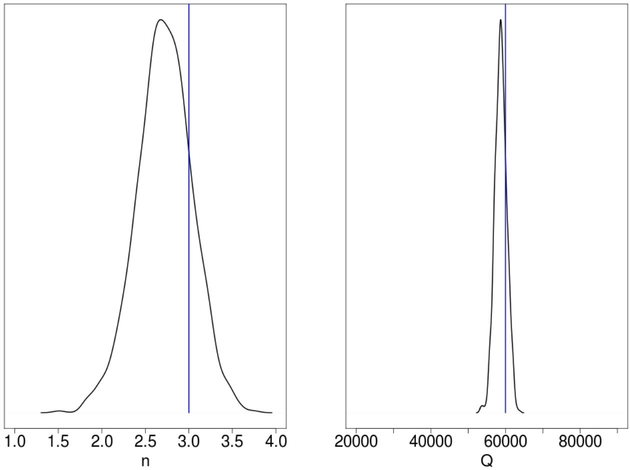
# CISM Results



# CISM Results



# CISM Results



# Connection to Ensemble Kalman Filter

- Take $\mu_x, \Sigma_x$ and $\lambda_y$ as fixed, known
- Similar to letting $k_0 \to 0, \Lambda_0 \to diag(0)$
- Then the update step of the EnsKF samples the posterior $p(x|y)$ using perturbed observations and a conditional simulation approach
- Only get $\#$ of ensemble member draws from the posterior (in our case 8)

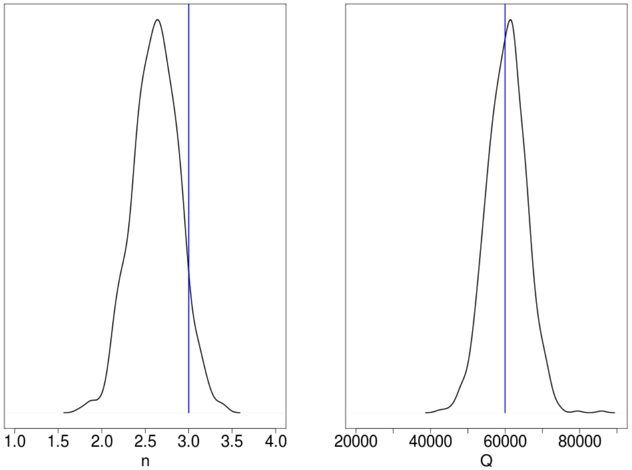
# Optimal Design

- For many of the problems we are investigating, we may be very restricted in our sampling of the real-world process
- Considering $\lambda_y, \mu_x, \Sigma_x$ known, Bayesian D-Optimal design selects design points $\xi$ maximizing $U(\xi)$ where

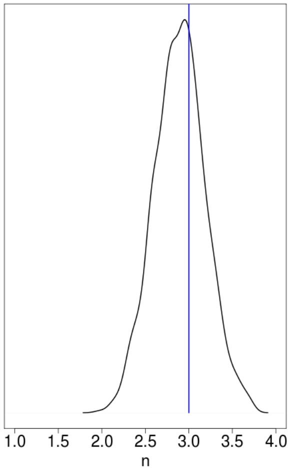
$$U(\xi) = \int log(f(\theta_0|y,\xi))f(y,\theta_0|\xi)d\theta_0 dy$$

- In our case, $U(\xi) = -\frac{1}{2} logdet(\Sigma_\theta - \frac{1}{n}\Sigma_{\theta,\chi}(\Sigma_\chi + \lambda_y^{-1} I)^{-1}\Sigma_{\chi,\theta})$
- $\xi$ is manifested through the operator $H$, since, for instance, $\Sigma_\chi = H\Sigma_x H^T$

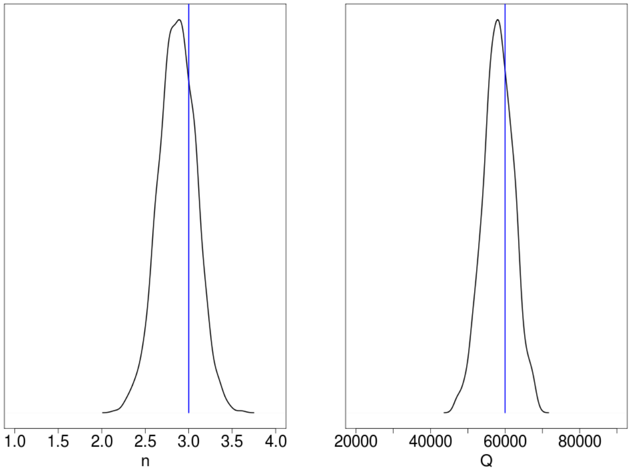
# 5 Optimal Points



# 10 Optimal Points



# 20 Optimal Points



# Conclusions & Future Directions

- Simpler approach to calibration than RSM
- Computationally efficient
- Related to EnsKF
- Optimal design, an important capability for our problem
- Extend to account for model discrepancy, “auxilary” parameters.