Bayesian Multilevel Model Calibration for Inversion of "Perfect" Data in the Presence of Epistemic and Aleatory Uncertainty

Joseph B. Nagel & Bruno Sudret

Chair of Risk, Safety & Uncertainty Quantification

MascotNum 2014, April 23-25, 2014

Outline

Motivating Example

NASA Langley Multidisciplinary UQ Challenge

- Originating from a realistic application from aerospace engineering
- Widely discipline-independent "blackbox" problem formulation

Five UQ problems

- **Uncertainty characterization**
- Sensitivity analysis
- Uncertainty propagation
- Extreme-case analysis
- Robust design

L. G. Crespo, S. P. Kenny, and D. P. Giesy. "The NASA Langley Multidisciplinary Uncertainty Quantification Challenge". In: 16th AIAA Non-Deterministic Approaches Conference (SciTech 2014), National Harbor, Maryland (USA). January 13-17, 2014

Uncertainty Characterization Subproblem

Generic problem statement

With responses of a given forward model, learn about unknown forward model inputs, that are subject to uncertainty and variability

- Estimating forward model parameters
- Identifying the distribution of variable inputs
- Under additional nuisance variables with known distribution

Sources of information

- $2 \times 25 = 50$ "perfect" observations
- Computational "blackbox" forward model
- "Prior" uncertainty model of its inputs

Bayesian interpretation & solution

Bayesian calibration of a suitably defined multilevel model

Numerical solution by means of Markov chain Monte Carlo

Section Outline

Bayesian Multilevel Modeling: Definition

Definition

A **hierarchical** or **multilevel model** is an "assembly of multiple submodels that are hierarchically interlinked"

Hierarchy

- Probabilistic relations
- Deterministic mappings

Submodels

1. **Forward model** (physics involved)
2. **Prior model** (input uncertainty)
3. **Residual model** (prediction error)

Forward Model: Physical Model

Forward model (physical model)

- Mathematical representation of a system under investigation
- Deterministic map: parameters $\mapsto$ data

$$\mathcal{M}\colon (\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{\zeta}, \boldsymbol{d}) \mapsto \tilde{\boldsymbol{y}} = \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{\zeta}, \boldsymbol{d})$$

Input uncertainty (w.r.t. experiments $i = 1, \ldots, n$)

- Experimental conditions: $\boldsymbol{d}$ $\rightarrow \boldsymbol{d}_i$
- Fixed yet unknown model parameters: $\boldsymbol{m}$ $\rightarrow \boldsymbol{m}$
- Inputs with well-known variability: $\boldsymbol{\zeta}$ $\rightarrow \boldsymbol{\zeta}_i$
- Inputs with unknown variability: $\boldsymbol{x}$ $\rightarrow \boldsymbol{x}_i$

Prior Model: Uncertainty & Variability

Unknown model parameter

epistemic uncertainty: $\boldsymbol{M} \sim \pi_{\boldsymbol{M}}(\boldsymbol{m})$ (parametric prior)

Well-known input variability

aleatory uncertainty: $\left\{ \begin{array}{l} \boldsymbol{Z}_i \sim f_{\boldsymbol{Z}}(\zeta_i; \boldsymbol{\theta}_{\boldsymbol{Z}}) \\ \text{with known } \boldsymbol{\theta}_{\boldsymbol{Z}} \end{array} \right\}$ (structural prior)

Unknown input variability

aleatory uncertainty: $(\boldsymbol{X}_i | \boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}_i | \boldsymbol{\theta}_{\boldsymbol{X}})$ (structural prior)

epistemic uncertainty: $\boldsymbol{\Theta}_{\boldsymbol{X}} \sim \pi_{\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{\theta}_{\boldsymbol{X}})$ (parametric prior)

Residual Model: Statistical Data Model

Residual model

- Observations are in some way "imperfect" (e.g. due to measurement errors, numerical approximations and forward model inadequacies)
- Representation as $\boldsymbol{y}_i = \tilde{\boldsymbol{y}}_i + \boldsymbol{\varepsilon}_i$ with $\boldsymbol{\varepsilon}_i$ a realization of $\boldsymbol{E}_i \sim f_{\boldsymbol{E}_i}(\boldsymbol{\varepsilon}_i)$

"Imperfect" data

Data are viewed as realizations $\boldsymbol{y}_i$ of random variables $(\boldsymbol{Y}_i | \boldsymbol{m}, \boldsymbol{x}_i, \boldsymbol{\zeta}_i)$ distributed according to $f(\boldsymbol{y}_i | \boldsymbol{m}, \boldsymbol{x}_i, \boldsymbol{\zeta}_i) = f_{\boldsymbol{E}_i}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}_i, \boldsymbol{\zeta}_i, \boldsymbol{d}_i)\big)$

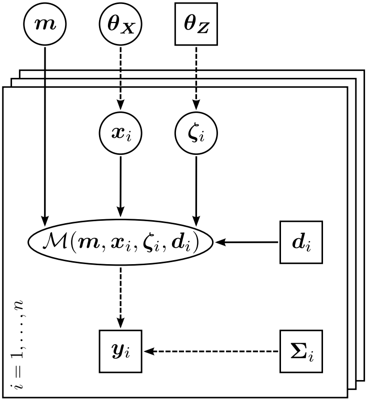
Bayesian Multilevel Model

Bayesian Multilevel Model

$$(\boldsymbol{Y}_i | \boldsymbol{m}, \boldsymbol{x}_i, \zeta_i) \sim f_{\boldsymbol{E}_i}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}_i, \zeta_i, \boldsymbol{d}_i)\big)$$
$$\boldsymbol{M} \sim \pi_{\boldsymbol{M}}(\boldsymbol{m})$$
$$\boldsymbol{Z}_i \sim f_{\boldsymbol{Z}}(\zeta_i; \boldsymbol{\theta}_{\boldsymbol{Z}})$$
$$(\boldsymbol{X}_i | \boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}_i | \boldsymbol{\theta}_{\boldsymbol{X}})$$
$$\boldsymbol{\Theta}_{\boldsymbol{X}} \sim \pi_{\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{\theta}_{\boldsymbol{X}})$$

- Unless denoted otherwise, random variables are (conditionally) independent
- This defines a joint probability model of the entirety of random quantities

Directed Acyclic Graph



Bayesian Multilevel Model Calibration: Joint Inference

Unique approach to inference

Formulate a joint probability density of all unknowns $\big(\boldsymbol{m}, \langle \boldsymbol{x}_i \rangle, \langle \boldsymbol{\zeta}_i \rangle, \boldsymbol{\theta}_{\boldsymbol{X}}\big)$ conditioned on all knowns $\big(\langle \boldsymbol{y}_i \rangle, \boldsymbol{\theta}_{\boldsymbol{Z}}\big)$. Then integrate out nuisance.

$$\langle \boldsymbol{q}_i \rangle \equiv \langle \boldsymbol{q}_i \rangle_{1 \leq i \leq n} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n)$$

Joint prior

$$\pi\big(\boldsymbol{m}, \langle \boldsymbol{x}_i \rangle, \langle \boldsymbol{\zeta}_i \rangle, \boldsymbol{\theta}_{\boldsymbol{X}}\big) = \left(\prod_{i=1}^{n} f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}_i | \boldsymbol{\theta}_{\boldsymbol{X}}) \right) \left(\prod_{i=1}^{n} f_{\boldsymbol{Z}}(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_{\boldsymbol{Z}}) \right) \pi_{\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{\theta}_{\boldsymbol{X}})\, \pi_{\boldsymbol{M}}(\boldsymbol{m})$$

Joint posterior

$$\pi\big(\boldsymbol{m}, \langle \boldsymbol{x}_i \rangle, \langle \boldsymbol{\zeta}_i \rangle, \boldsymbol{\theta}_{\boldsymbol{X}} \,|\, \langle \boldsymbol{y}_i \rangle\big) \propto \left(\prod_{i=1}^{n} f_{\boldsymbol{E}_i}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}_i, \boldsymbol{\zeta}_i, \boldsymbol{d}_i)\big) \right) \pi\big(\boldsymbol{m}, \langle \boldsymbol{x}_i \rangle, \langle \boldsymbol{\zeta}_i \rangle, \boldsymbol{\theta}_{\boldsymbol{X}}\big)$$

"Perfect" Data Model

"Perfect" data

Data, i.e. forward model responses $\tilde{\boldsymbol{y}}_i = \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}_i, \boldsymbol{\zeta}_i, \boldsymbol{d}_i)$, are viewed as realizations of random variables $(\tilde{\boldsymbol{Y}}_i | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$ with $f_{\tilde{\boldsymbol{Y}}_i}(\tilde{\boldsymbol{y}}_i | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$

Bayesian model

$$(\tilde{\boldsymbol{Y}}_i | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\tilde{\boldsymbol{Y}}_i}(\tilde{\boldsymbol{y}}_i | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$$
$$(\boldsymbol{M}, \boldsymbol{\Theta}_{\boldsymbol{X}}) \sim \pi(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}) = \pi_{\boldsymbol{M}}(\boldsymbol{m})\, \pi_{\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{\theta}_{\boldsymbol{X}})$$

Likelihood function

$$\mathcal{L}\big(\langle \tilde{\boldsymbol{y}}_i \rangle | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}\big) = \prod_{i=1}^{n} f_{\tilde{\boldsymbol{Y}}_i}(\tilde{\boldsymbol{y}}_i | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$$

Posterior density

$$\pi\big(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}} | \langle \tilde{\boldsymbol{y}}_i \rangle\big) \propto \mathcal{L}\big(\langle \tilde{\boldsymbol{y}}_i \rangle | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}\big)\, \pi(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$$

Forward Uncertainty Propagation: Push-Forward

Formulate a likelihood function: $\mathcal{L}_{\text{data}}(\text{QoI}) = \mathcal{P}(\text{data}|\text{QoI})$

- Given the forward model $\mathcal{M}\colon (\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{\zeta}, \boldsymbol{d}) \mapsto \tilde{\boldsymbol{y}} = \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{\zeta}, \boldsymbol{d})$
- For fixed $(\boldsymbol{m}, \boldsymbol{d}_i)$ consider $\mathcal{M}_{\boldsymbol{m},\boldsymbol{d}_i}\colon (\boldsymbol{x}, \boldsymbol{\zeta}) \mapsto \tilde{\boldsymbol{y}}_i = \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{\zeta}, \boldsymbol{d}_i)$
- Propagate **aleatory** input uncertainties $(\boldsymbol{X}_i|\boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}_i|\boldsymbol{\theta}_{\boldsymbol{X}})$ and $\boldsymbol{Z}_i \sim f_{\boldsymbol{Z}}(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_{\boldsymbol{Z}})$ as determined by hyperparameters $(\boldsymbol{\theta}_{\boldsymbol{X}}, \boldsymbol{\theta}_{\boldsymbol{Z}})$

$$\left.\begin{array}{c} (\boldsymbol{X}_i|\boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}_i|\boldsymbol{\theta}_{\boldsymbol{X}}) \\ \boldsymbol{Z}_i \sim f_{\boldsymbol{Z}}(\boldsymbol{\zeta}_i; \boldsymbol{\theta}_{\boldsymbol{Z}}) \end{array}\right\} \tilde{\boldsymbol{Y}}_i = \mathcal{M}(\boldsymbol{m}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{d}_i)$$

$$\rightsquigarrow \quad (\tilde{\boldsymbol{Y}}_i|\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\tilde{\boldsymbol{Y}}_i}(\tilde{\boldsymbol{y}}_i|\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$$

Likelihood function: $\mathcal{L}\big(\langle\tilde{\boldsymbol{y}}_i\rangle|\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}\big) = \prod_{i=1}^{n} f_{\tilde{\boldsymbol{Y}}_i}(\tilde{\boldsymbol{y}}_i|\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$

Likelihood Estimation

KDE-based evaluation of the likelihood

$$\text{Sample} \left\{ \begin{array}{l} \boldsymbol{x}^{(k)} \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}^{(k)}|\boldsymbol{\theta}_{\boldsymbol{X}}) \\ \boldsymbol{\zeta}^{(k)} \sim f_{\boldsymbol{Z}}(\boldsymbol{\zeta}^{(k)};\boldsymbol{\theta}_{\boldsymbol{Z}}) \end{array} \right\} \text{ for } k = 1, \ldots, K$$

$$\text{Compute } \tilde{\boldsymbol{v}}_i^{(k)} = \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}^{(k)}, \boldsymbol{\zeta}^{(k)}, \boldsymbol{d}_i)$$

$$\hat{\mathcal{L}}_{\mathrm{KS}}(\langle \tilde{\boldsymbol{y}}_i \rangle | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}) = \prod_{i=1}^{n} \left(\frac{1}{K} \sum_{k=1}^{K} \mathcal{K}_{\boldsymbol{H}} \left(\tilde{\boldsymbol{y}}_i - \tilde{\boldsymbol{v}}_i^{(k)} \right) \right)$$

- Free algorithmic parameters have to be tuned: number of samples K and the kernel bandwidth $\boldsymbol{H}$ $\rightarrow$ **"optimal" parameter tuning**?
- However, the likelihood is estimated in every MCMC iteration
- The optimal tuning of free parameters has to be assessed with respect to the **posterior fidelity**

MCMC & Likelihood Estimation

Bayesian inference by MCMC

Construct a MC whose invariant distribution equals the posterior

- Metropolis-Hastings correction: $\alpha = \min\left(1, \frac{\pi_1(\boldsymbol{q}^{(\star)})\, P(\boldsymbol{q}^{(t)}\,|\,\boldsymbol{q}^{(\star)})}{\pi_1(\boldsymbol{q}^{(t)})\, P(\boldsymbol{q}^{(\star)}\,|\,\boldsymbol{q}^{(t)})}\right)$
- **Statistical likelihood estimations induce modifications on the level of the Markov chain transition kernel**
- This may alter the equilibrium distribution of the Markov chain ☠

Key concept: posterior fidelity

We define posterior fidelity as the degree as to which the induced equilibrium distribution is in congruence with the true posterior

- Posterior fidelity depends on a complex interplay between the estimator $\hat{\mathcal{L}}$, the true posterior π_1 and the proposal distribution P

Posterior Fidelity & Parameter Tuning

- We do not have a means to define "optimal" parameter tunings
- Setting K and $\boldsymbol{H}$ is trading-off the fidelity of the posterior against the ease of its computation (**fidelity vs. feasibility**)
- An "appropriate" decision is made when the likelihood ratio can be estimated "sufficiently accurate"

$$\underline{\text{likelihood ratio:}} \quad \frac{\hat{\mathcal{L}}(\boldsymbol{q}^{(\star)})}{\hat{\mathcal{L}}(\boldsymbol{q}^{(t)})} \approx \frac{\mathcal{L}(\boldsymbol{q}^{(\star)})}{\mathcal{L}(\boldsymbol{q}^{(t)})}$$

- High posterior fidelity is achieved if "appropriate" decisions are being frequently made over the course of the Markov chain

$\rightarrow$ **Heuristics and plausibility checks**

Section Outline

The NASA Langley UQ Challenge

Problem statement

Given forward model outputs, learn about unknown forward model inputs

Available information

- The forward model $\mathcal{M} \equiv h_1$ is distributed as MATLAB p-code file
- Data $\langle \tilde{y}_i \rangle_{1 \leq i \leq 50}$ with $\tilde{y}_i = h_1(p_{1,i}, p_2, p_{3,i}, p_{4,i}, p_{5,i})$
- Information about the forward model inputs $(p_{1,i}, p_2, p_{3,i}, p_{4,i}, p_{5,i})$

Types of forward model inputs

- Unknown forward model parameter to be estimated: p_2
- Inputs whose variability is of inferential interest: p_1, p_4, p_5
- Additional nuisance variables with known distribution: p_3

Category I & II: Aleatory & Epistemic Uncertainty

Fixed yet unknown forward model parameter: p_2

- The model parameter $\boldsymbol{m} \equiv p_2$ is an unknown constant
- Its value is known to be contained in an epistemic interval $\Delta = [0, 1]$
- **Parametric prior**: $P_2 \sim \pi_2(p_2) = \mathcal{U}(p_2 | 0, 1)$

Forward model input with well-known variability: p_3

- Experiment-specific realizations $\boldsymbol{\zeta}_i \equiv p_{3,i}$ for $i = 1, \ldots, n$ were sampled from a uniform population distribution $\mathcal{U}(p_{3,i} | a_3, b_3)$
- Its hyperparameters $(a_3, b_3) = (0, 1)$ are well-known
- **Structural prior**: $(P_{3,i} | \boldsymbol{\theta}_3) \sim f_3(p_{3,i} | \boldsymbol{\theta}_3) = \mathcal{U}(p_{3,i} | 0, 1)$

Category III: Mixed Uncertainty: Correlated Gaussian

Variable forward model inputs: p_4, p_5

- Possibly correlated Gaussian variables: (p_4, p_5)

$$\big((P_{4,i}, P_{5,i})|\boldsymbol{\theta}_{45}\big) \sim f_{45}\big((p_{4,i}, p_{5,i})|\boldsymbol{\theta}_{45}\big) = \mathcal{N}\big((p_{4,i}, p_{5,i})|\boldsymbol{\mu}_{45}, \boldsymbol{\Sigma}_{45}\big)$$

$$\boldsymbol{\mu}_{45} = \begin{pmatrix} \mu_4 \\ \mu_5 \end{pmatrix}, \quad \boldsymbol{\Sigma}_{45} = \begin{pmatrix} \sigma_4^2 & \rho_{45}\,\sigma_4\,\sigma_5 \\ \rho_{45}\,\sigma_4\,\sigma_5 & \sigma_5^2 \end{pmatrix}$$

- Hyperparameters $\boldsymbol{\theta}_{45} \equiv (\mu_4, \sigma_4^2, \mu_5, \sigma_5^2, \rho_{45})$
- It is known that $-5 \leq \mu_j \leq 5$, ${}^1/_{400} \leq \sigma_j^2 \leq 4$ and $|\rho_{45}| \leq 1$

$$\left.\begin{aligned} \pi(\mu_j) &= \mathcal{U}(-5, 5), \\ \pi(\sigma_j^2) &= \mathcal{U}({}^1/_{400}, 4), \\ \pi(\rho_{45}) &= \mathcal{U}(-1, 1), \end{aligned}\right\} \; \pi_{45}(\boldsymbol{\theta}_{45}) = \left(\prod_{j=4}^{5} \pi(\mu_j)\,\pi(\sigma_j^2)\right) \pi(\rho_{45})$$

Category III: Mixed Uncertainty: Unimodal Beta

Variable forward model inputs: p_1

- Variables following a unimodal beta distribution: p_1

$$(P_{1,i}|\boldsymbol{\theta}_1) \sim f_1(p_{1,i}|\boldsymbol{\theta}_1) = \mathrm{Beta}(p_{1,i}|\mu_1, \sigma_1^2)$$

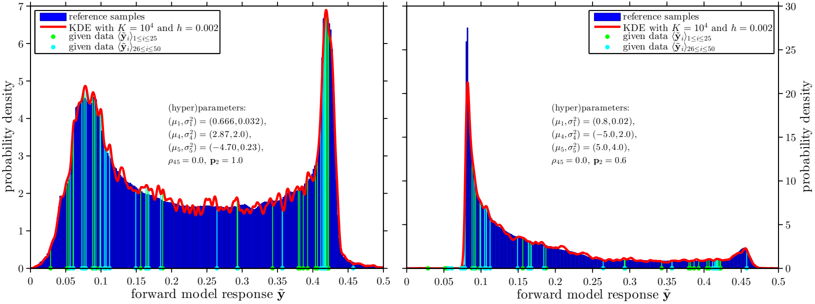
- Statistical $\boldsymbol{\theta}_1 \equiv (\mu_1, \sigma_1^2)$ or shape hyperparameters $\boldsymbol{\theta}_1 \equiv (\alpha_1, \beta_1)$
- It is requested that $3/5 \leq \mu_1 \leq 4/5$ and $1/50 \leq \sigma_1^2 \leq 1/25$
- Moreover the unimodality translates into $\alpha_1, \beta_1 > 1$
- Therefore one can state a uniform hyperprior on $\boldsymbol{\theta}_1 \equiv (\mu_1, \sigma_1^2)$

$$\boldsymbol{\Theta}_1 \sim \pi_1(\boldsymbol{\theta}_1) = \mathcal{U}(\boldsymbol{\theta}_1|\boldsymbol{S}), \text{ with}$$
$$\boldsymbol{S} = \left\{(\mu_1, \sigma_1^2) \in \mathbb{R}^2 \,\middle|\, 3/5 \leq \mu_1 \leq 4/5,\ 1/50 \leq \sigma_1^2 \leq 1/25,\ \alpha_1, \beta_1 > 1\right\}$$

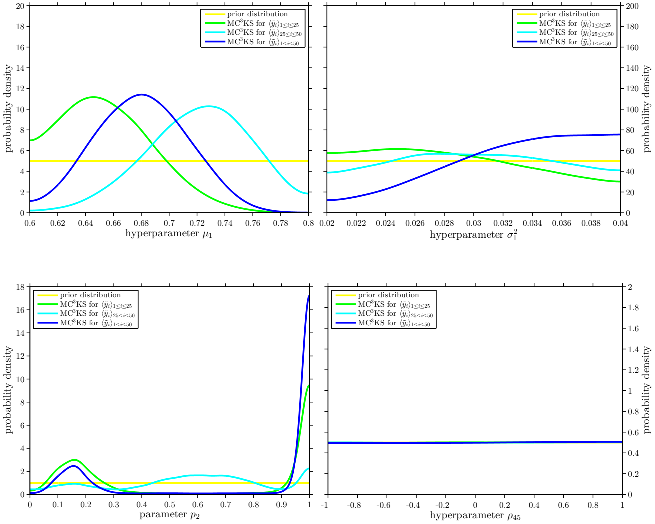
Estimation of the Target Density $f_{\tilde{Y}_i}\left(\tilde{y}_i \,|\, p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45}\right)$

Tuning of free algorithmic parameters

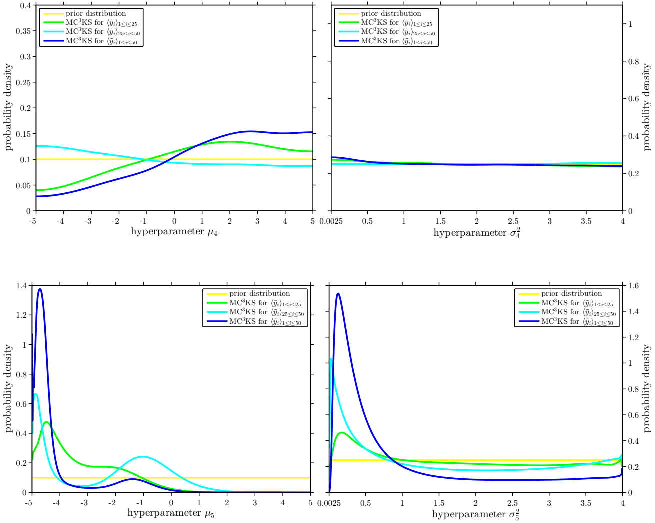
- Resource limitations bound the number of forward model runs K
- Cascade of runs with $K = 10^4$ and decreasing bandwidths h
- Observations: Initial shrinkage, eventual breakdown, in between the distribution is rather stable with respect to changes in h



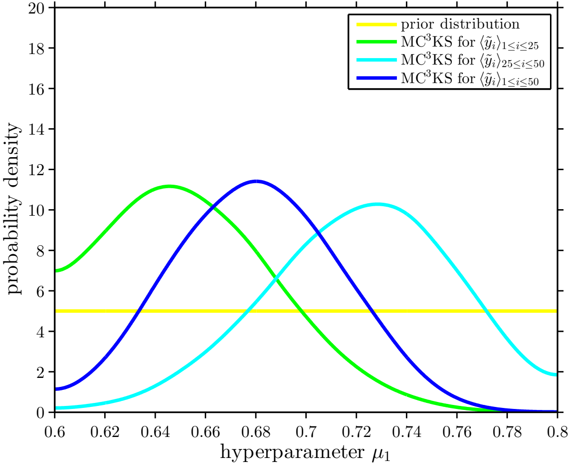
Marginal Posteriors of μ_1, σ_1^2, p_2 and ρ_{45}



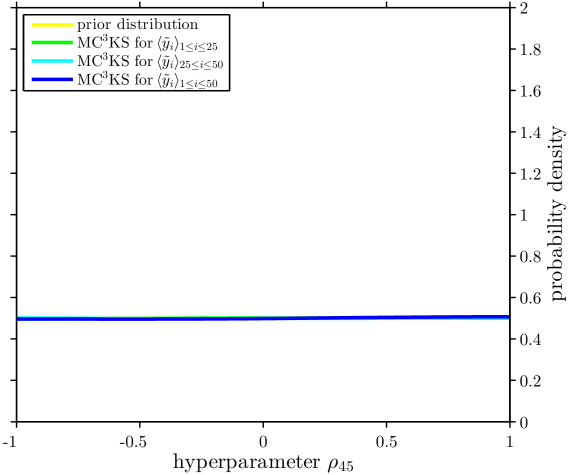
Marginal Posteriors of μ_4, σ_4^2, μ_5 and σ_5^2



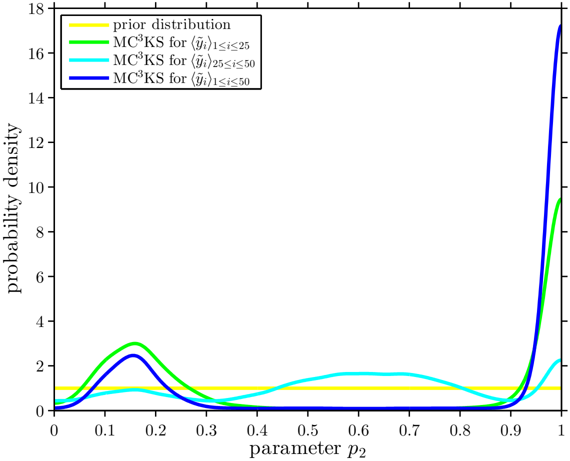
Posterior Marginal of μ_1



Posterior Marginal of ρ_{45}



Posterior Marginal of p_2



Section Outline

Partial Data Augmentation

Data Augmentation

- Traditionally data augmentation is a powerful MCMC technique that aims at improving numerical efficiency of posterior sampling
- Here we will try to exploit partial data augmentation for gaining posterior fidelity
- **Partial data augmentation**: explicitly infer $\langle p_{1,i} \rangle$ as auxiliary variables, then integrate them out as nuisance
- Presume that $\pi(\langle p_{1,i} \rangle, p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45} \,|\, \langle \tilde{y}_i \rangle)$ is "easier" to sample than sampling from $\pi(p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45} \,|\, \langle \tilde{y}_i \rangle)$
- Partial data augmentation trades off dimensionality and fidelity

Partial Data Augmentation

Parametric & structural prior

$$\pi\big(\langle p_{1,i}\rangle, p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45}\big) = \left(\prod_{i=1}^{n} f_1\big(p_{1,i}|\boldsymbol{\theta}_1\big)\right) \pi_1\big(\boldsymbol{\theta}_1\big)\, \pi_2\big(p_2\big)\, \pi_{45}\big(\boldsymbol{\theta}_{45}\big)$$

Augmented likelihood

$$\mathcal{L}\big(\langle \tilde{y}_i\rangle | \langle p_{1,i}\rangle, p_2, \cancel{\boldsymbol{\theta}_1}, \boldsymbol{\theta}_{45}\big) = \prod_{i=1}^{n} f\big(\tilde{y}_i | p_{1,i}, p_2, \boldsymbol{\theta}_{45}\big)$$

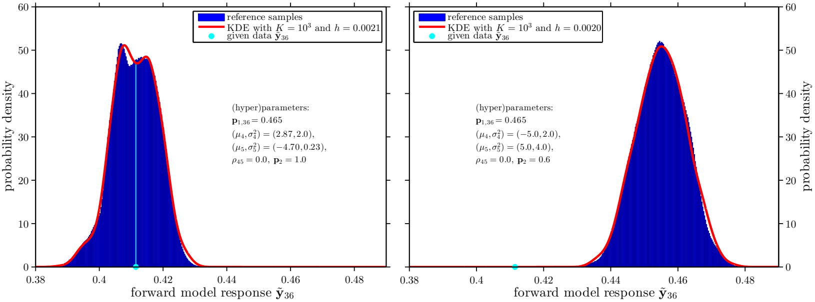
Augmented posterior

$$\pi\big(\langle p_{1,i}\rangle, p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45} | \langle \tilde{y}_i\rangle\big) \propto \mathcal{L}\big(\langle \tilde{y}_i\rangle | \langle p_{1,i}\rangle, p_2, \boldsymbol{\theta}_{45}\big)\, \pi\big(\langle p_{1,i}\rangle, p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45}\big).$$

Integrate out nuisance

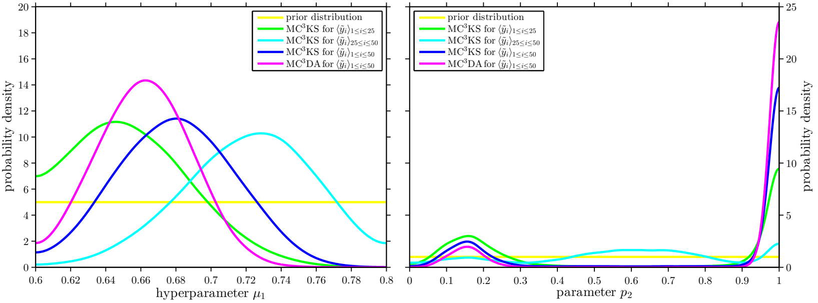
$$\pi\big(p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45} | \langle \tilde{y}_i\rangle\big) = \int \cdots \int_{\mathcal{D}_{p_1}^n} \pi\big(\langle p_{1,i}\rangle, p_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_{45} | \langle \tilde{y}_i\rangle\big)\, \mathrm{d}\langle p_{1,i}\rangle$$

Estimation of the Target Density $f(\tilde{y}_i \,|\, p_{1,i}, p_2, \boldsymbol{\theta}_{45})$

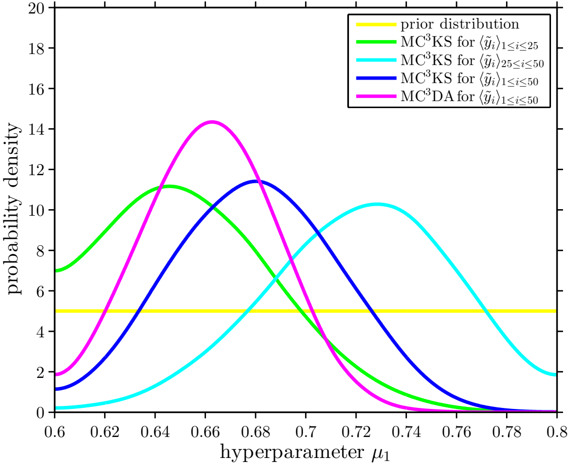


- The target density is “easier” to estimate and resembles a Gaussian
- Automatic bandwidth selection via Silverman’s rule of thumb
- Allows for more adequate likelihood estimations and therefore promises a boost in posterior fidelity

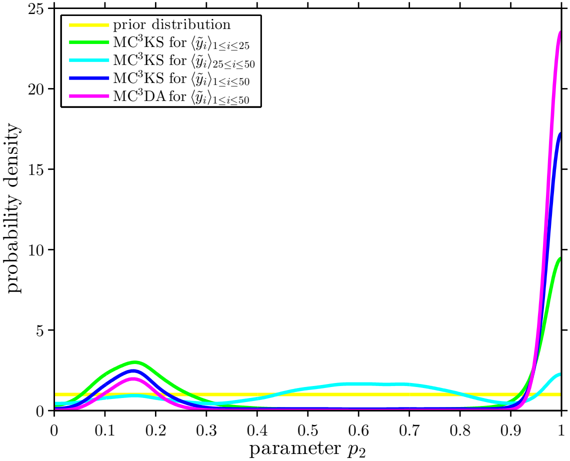
Marginal Posteriors of μ_1 and p_2



Posterior Marginal of μ_1



Posterior Marginal of p_2



Conclusions

Conclusions

- Bayesian multilevel modeling establishes a convenient framework for solving complex inverse problems under uncertainty
- “Perfect” data model had to be devised (intractable likelihood)
- Numerical complexity has been demonstrated: Multilevel model calibration more complex than parameter estimation
- Motivates metamodeling (PCE,...) when involving computationally more expensive forward models (FEM,...)
- Posterior fidelity has been introduced as an important concept
- Data augmentation has been used to gain posterior fidelity