Calibration of computer models

Bayesian Calibration

Pierre Barbillon

Fall 2023, école ETICS

Outline

Field data

- Field data provided by physical experiments:

$$\mathbf{y}^e = y^e(\mathbf{x}_1^e), \ldots, y^e(\mathbf{x}_{n_e}^e)\,,$$

- n_e is small, $\mathbf{x}_1, \ldots \mathbf{x}_{n_e} \in \mathcal{X}$ hard to set, sometimes uncontrollable, included in a small domain...

- Model:

$$y^e(\mathbf{x}_i^e) = \zeta(\mathbf{x}_i^e) + \epsilon(\mathbf{x}_i^e)\,,$$

  where
  - $\zeta(\cdot)$ real physical process (unknown),
  - $\epsilon(\mathbf{x}_i^e)$ often assumed i.i.d. $\mathcal{N}(0, \sigma^2)$,
  - σ^2 sometimes treated as known...

Relationship between the simulator and the data

for $i = 1, \ldots, n_e$,

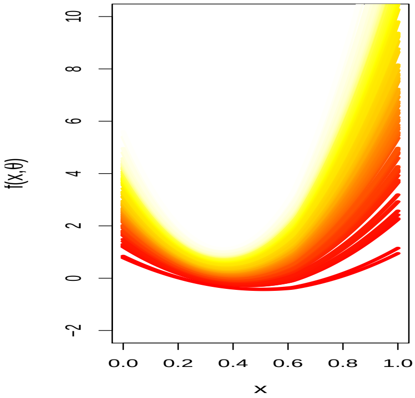
- if the simulator sufficiently represents the physical system:

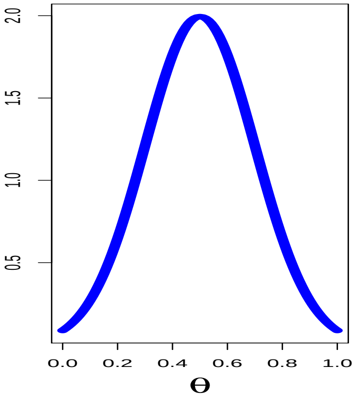
$$y_i^e = f(\mathbf{x}_i^e, \boldsymbol{\theta}^*) + \epsilon(\mathbf{x}_i^e)\,, \tag{1}$$

  i.e. for the unknown value $\boldsymbol{\theta} = \boldsymbol{\theta}^*$: $f(\mathbf{x}, \theta^*) = \zeta(\mathbf{x})$ for any $\mathbf{x} \in \mathcal{X}$,

A calibration example

Prior:
prior distribution on unknown $\boldsymbol{\theta}$: $\pi(\cdot)$
from expert judgment, past experiments...
Possible choice $\pi(\boldsymbol{\theta}) = \mathcal{N}(\theta_0, \sigma_0^2) = \mathcal{N}(0.5, 0.04)$.





A calibration example

Data:
Couples $(\mathbf{x}_1^e, y_1^e), \ldots, (\mathbf{x}_{n_e}^e, y_{n_e}^e)$ from physical experiments.

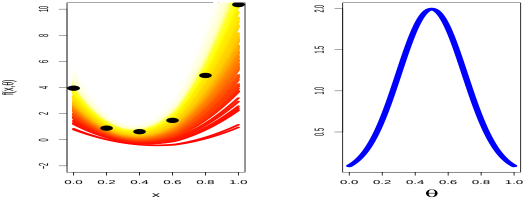
Posterior distribution:

$$
\begin{aligned}
\pi(\boldsymbol{\theta}|\mathbf{y}^e) &\propto \mathcal{L}(\boldsymbol{\theta}|\mathbf{y}^e) \cdot \pi(\boldsymbol{\theta}) \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y(\mathbf{x}_i) - f(\mathbf{x}_i, \boldsymbol{\theta}))^2 - \frac{1}{2\sigma_0^2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^2\right)
\end{aligned}
$$

- Analytical posterior if $\boldsymbol{\theta} \mapsto f(\mathbf{x}, \boldsymbol{\theta})$ is a linear map,
- Otherwise MH sampling to simulate according to the posterior distribution. [Robert et al., 1999]

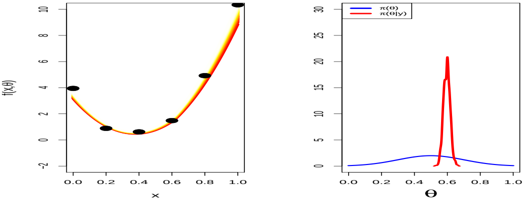
A calibration example

Prior with data:



⇓ Metropolis-Hastings algorithm ⇓

Posterior on θ:



More details on the MH algorithm

Initialisation:
$\boldsymbol{\theta}^0$ chosen.

Update:
iterations $t = 1, \ldots,$

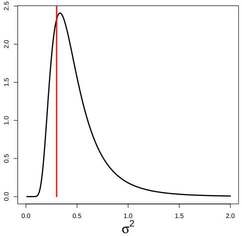
1. Proposal: $\tilde{\boldsymbol{\theta}}^{t+1} = \boldsymbol{\theta}^t + \mathcal{N}(0, \tau^2)$.
2. Compute
$$\alpha(\boldsymbol{\theta}^t, \tilde{\boldsymbol{\theta}}^{t+1}) = \frac{\pi(\tilde{\boldsymbol{\theta}}^{t+1}|\mathbf{y}^e)}{\pi(\boldsymbol{\theta}^t|\mathbf{y}^e)}$$
3. Acceptation:
$$\boldsymbol{\theta}^{t+1} = \begin{cases} \tilde{\boldsymbol{\theta}}^{t+1} & \text{with probability } \alpha(\boldsymbol{\theta}^t, \tilde{\boldsymbol{\theta}}^{t+1}) \\ \boldsymbol{\theta}^t & \text{otherwise.} \end{cases}$$

Note that the ratio $\alpha(\boldsymbol{\theta}^t, \tilde{\boldsymbol{\theta}}^{t+1})$ needs several computations of $f(\mathbf{x}, \boldsymbol{\theta})$ at each step since

$$\pi(\boldsymbol{\theta}|\mathbf{y}^e) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i^e - f(\mathbf{x}_i^e, \boldsymbol{\theta}))^2 - \frac{1}{2\sigma_0^2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^2\right).$$

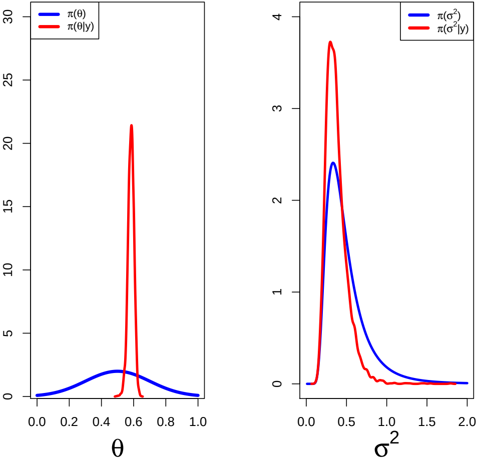
Unknown σ^2

- prior distribution on σ^2: $\pi(\sigma^2) = \mathcal{IG}(5, 2)$



- Gibbs algorithm to simulate couples $(\boldsymbol{\theta}, \sigma^2)$ from $\pi(\boldsymbol{\theta}, \sigma^2|\mathbf{y}^e)$. Iterate :
  1. MH algorithm to simulate $\boldsymbol{\theta}_t$ from $\pi(\cdot|\mathbf{y}^e, \sigma^2_{t-1})$,
  2. conditional simulation of σ^2_t from $\pi(\cdot|\mathbf{y}^e, \boldsymbol{\theta}_t)$.

Posterior distributions



Comparison

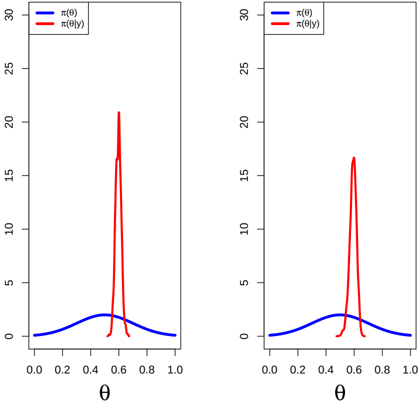


Figure: known σ^2 vs unknown σ^2

Outline

Relationship between the simulator and the data

for $i = 1, \ldots, n$,

- if the simulator represents sufficiently well the physical system:

$$y_i^e = f(\mathbf{x}_i^e, \boldsymbol{\theta}^*) + \epsilon(\mathbf{x}_i^e)\,,$$

  i.e. for the unknown value $\boldsymbol{\theta} = \boldsymbol{\theta}^*$: $f(\mathbf{x}, \theta^*) = \zeta(\mathbf{x})$ for any $\mathbf{x} \in \mathcal{X}$,

- if the field observations are inconsistent with the simulations (irreducible model discrepancy):

$$y_i^e = f(\mathbf{x}_i, \boldsymbol{\theta}^*) + \delta(\mathbf{x}_i) + \epsilon(\mathbf{x}_i)\,.$$

  $\delta(\cdot)$ models the difference between the simulator and the physical system:

$$\delta(\mathbf{x}) = \zeta(\mathbf{x}) - f(\mathbf{x}, \theta^*)\,.$$

Limited computational budget:

Limited number M of runs of the simulator.

Ref.: [Kennedy and O'Hagan, 2001, Higdon et al., 2004]

Statistical models

$$
\begin{aligned}
\mathcal{M}_0 &: \forall i \in [\![1,\dots,n_e]\!] \quad y_i^e = \zeta(\mathbf{x}_i^e) + \epsilon_i \\
\mathcal{M}_1 &: \forall i \in [\![1,\dots,n_e]\!] \quad y_i^e = f_c(\mathbf{x}_i^e, \boldsymbol{\theta}) + \epsilon_i, \\
\mathcal{M}_2 &: \forall i \in [\![1,\dots,n_e]\!] \quad y_i^e = F(\mathbf{x}_i^e, \boldsymbol{\theta}) + \epsilon_i, \\
\mathcal{M}_3 &: \forall i \in [\![1,\dots,n_e]\!] \quad y_i^e = f_c(\mathbf{x}_i^e, \boldsymbol{\theta}) + \delta(\mathbf{x}_i^e) + \epsilon_i, \\
\mathcal{M}_4 &: \forall i \in [\![1,\dots,n_e]\!] \quad y_i^e = F(\mathbf{x}_i^e, \boldsymbol{\theta}) + \delta(\mathbf{x}_i^e) + \epsilon_i.
\end{aligned}
$$

where

- $\epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma_{err}^2)$,
- $F(\bullet,\bullet) \sim \mathcal{GP}\Big(m_S(\bullet,\bullet), c_S\{(\bullet,\bullet),(\bullet,\bullet)\}\Big)$, on $\mathbb{X} \times \Theta$
- $\delta(\bullet) \sim \mathcal{GP}\Big(\boldsymbol{m}_\delta(\bullet), c_\delta(\bullet,\bullet)\Big)$ on $\mathbb{X}$.

[Carmassi et al., 2019]

Likelihood for $\mathcal{M}_1$ and $\mathcal{M}_3$

$$\mathcal{L}^F(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta; \mathbf{y}^e, \mathbf{X}^e) = \frac{1}{(2\pi)^{n_e/2}|\boldsymbol{V}_{exp}(\mathbf{X}^e)|^{1/2}} \exp\left\{ -\frac{1}{2}\left(\mathbf{y}^e - \boldsymbol{m}_{exp}(\mathbf{X}^e, \boldsymbol{\theta})\right)^T \boldsymbol{V}_{exp}(\mathbf{X}^e)^{-1} \left(\mathbf{y}^e - \boldsymbol{m}_{exp}(\mathbf{X}^e, \boldsymbol{\theta})\right)\right\}.$$

$$\mathbb{E}[\mathbf{y}^e|\boldsymbol{\theta}, \boldsymbol{\beta}_\delta; \mathbf{X}^e] = \boldsymbol{m}_{exp}^{\boldsymbol{\beta}_\delta}(\mathbf{X}^e, \boldsymbol{\theta}) = \boldsymbol{m}_{exp}(\mathbf{X}^e, \boldsymbol{\theta}) = f_c(\mathbf{X}^e, \boldsymbol{\theta}) + \boldsymbol{H}_\delta(\mathbf{X}^e)\boldsymbol{\beta}_\delta.$$

Then, the expression of the variance is given by

$$\mathbb{V}ar[\mathbf{y}^e|\Phi_\delta; \mathbf{X}^e] = \boldsymbol{V}_{exp}^{\Phi_\delta, \sigma_{err}^2}(\mathbf{X}^e) = \boldsymbol{V}_{exp}(\mathbf{X}^e) = \Sigma_\delta(\mathbf{X}^e) + \sigma_{err}^2 \boldsymbol{I}_{n_e},$$

with $\forall (i,j) \in [\![1, \ldots, n]\!]^2 : (\Sigma_\delta(\mathbf{X}^e))_{i,j} = (\Sigma_\delta^{\Phi_\delta}(\mathbf{X}^e))_{i,j} = \sigma_\delta^2 c_\delta(\{\boldsymbol{x}_i, \boldsymbol{x}_j\})$.

For $\mathcal{M}_1$

$\boldsymbol{m}_{exp}(\mathbf{X}^e, \boldsymbol{\theta}) = f_c(\mathbf{X}^e, \boldsymbol{\theta})$ and $\boldsymbol{V}_{exp}(\mathbf{X}^e) = \sigma_{err}^2 \boldsymbol{I}_{n_e}$.

When the code is slow

Data:

1. DoNE: Design of Numerical Experiments: $D^c = \{(\mathbf{x}_1, \boldsymbol{\theta}_1), \dots, (\mathbf{x}_N, \boldsymbol{\theta}_N)\}$ with corresponding evaluations of the computer model (time-consuming):

$$\mathbf{y}^c = f(D^c) = \{f(\mathbf{x}_1, \boldsymbol{\theta}_1), \dots, f(\mathbf{x}_N, \boldsymbol{\theta}_N)\}.$$

2. DoFE: Design of Field Experiments: $\mathbf{X}^e = \{\mathbf{x}_1^e, \dots, \mathbf{x}_{n_e}^e\}$ with corresponding noisy observation of ζ:

$$\mathbf{y}^e = \{y_1^e = \zeta(\mathbf{x}_1^e) + \epsilon_1, \dots, y_{n_e}^e = \zeta(\mathbf{x}_{n_e}^e) + \epsilon_{n_e}\}.$$

Model: $\forall 1 \leq i \leq n_e, \quad y_i^e = f(\mathbf{x}_i^e, \boldsymbol{\theta}) + \delta(\mathbf{x}_i^e) + \epsilon_i$ where:

- f is emulated via a GP Emulator [Sacks et al., 1989] : $f \sim \mathcal{GP}(m_S(\cdot), c_S(\cdot,\cdot))$, $f|f(D^c) \sim \mathcal{GP}$ is the emulator/surrogate/metamodel,
- δ the discrepancy modeled as a GP: $\delta \sim \mathcal{GP}(H_\delta(\cdot)\boldsymbol{\beta}_\delta, \sigma_\delta^2 C_\delta(\cdot,\cdot))$,
- $\epsilon \overset{iid}{\sim} \mathcal{N}(0, \sigma_{err}^2)$ are measurement errors.

Likelihood for $\mathcal{M}_2$ and $\mathcal{M}_4$

$\Phi = (\sigma_c^2, \phi_c, \sigma_\delta^2, \phi_\delta)$, and $\boldsymbol{\beta} = (\boldsymbol{\beta}_S, \boldsymbol{\beta}_\delta)$
Full Likelihood written for $\mathbf{y} = (\mathbf{y}^e, \mathbf{y}^c)$:

$$\begin{aligned}
&\mathcal{L}^F(\boldsymbol{\theta}, \boldsymbol{\beta}, \Phi, \sigma_{err}^2; \boldsymbol{y}, \mathbf{X}^e, D^c) \\
=&\frac{1}{(2\pi)^{(n_e+N)/2} |\boldsymbol{V}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)|^{1/2}} \\
&\exp\left\{ -\frac{1}{2}\Big(\boldsymbol{y} - \boldsymbol{m}_{\boldsymbol{y}}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)\Big)^T \boldsymbol{V}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)^{-1} \Big(\boldsymbol{y} - \boldsymbol{m}_{\boldsymbol{y}}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)\Big) \right\}.
\end{aligned}$$

with

$$\begin{aligned}
\mathbb{E}[\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\beta}; \mathbf{X}^e, D^c] = \boldsymbol{m}_{\boldsymbol{y}}^{\boldsymbol{\beta}}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) =& \boldsymbol{m}_{\boldsymbol{y}}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) = \boldsymbol{H}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)\boldsymbol{\beta} \\
&= \begin{pmatrix} \boldsymbol{H}_S(\mathbf{X}^e, \boldsymbol{\theta}) & \boldsymbol{H}_\delta(\mathbf{X}^e) \\ \boldsymbol{H}_S(D^c) & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_S \\ \boldsymbol{\beta}_\delta \end{pmatrix}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}ar[\boldsymbol{y}|\boldsymbol{\theta}, \Phi, \sigma_{err}^2; \mathbf{X}^e, D^c] =& \boldsymbol{V}^{\Phi, \sigma_{err}^2}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) = \boldsymbol{V}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) \\
=& \begin{pmatrix} \Sigma_{exp,exp}(\mathbf{X}^e, \boldsymbol{\theta}) + \Sigma_\delta(\mathbf{X}^e) + \sigma_{err}^2 \boldsymbol{I}_{n_e} & \Sigma_{exp,c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) \\ \Sigma_{exp,c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)^T & \Sigma_{c,c}(D^c) \end{pmatrix}
\end{aligned}$$

- $\forall (i,j) \in [\![1,\dots,n_e]\!]^2 : (\Sigma_{exp,exp}(\mathbf{X}^e,\boldsymbol{\theta}))_{i,j} = c_S\{(\mathbf{x}_i^e,\boldsymbol{\theta}),(\mathbf{x}_j^e,\boldsymbol{\theta})\}$,
- $\forall (i,j) \in [\![1,\dots,n_e]\!] \times [\![1,\dots,N]\!] : (\Sigma_{exp,c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c))_{i,j} = c_S\{(\mathbf{x}_i^e,\boldsymbol{\theta}),(\mathbf{x}_j,\boldsymbol{\theta}_j)\}$,
- $\forall (i,j) \in [\![1,\dots,n_e]\!]^2 : (\Sigma_\delta(\mathbf{X}^e))_{i,j} = c_\delta\{(\mathbf{x}_i^e,\mathbf{x}_j^e)\}$,
- $\forall (i,j) \in [\![1,\dots,N]\!]^2 : (\Sigma_{c,c}(D^c))_{i,j} = c_S\{(\mathbf{x}_i,\boldsymbol{\theta}_i),(\mathbf{x}_j,\boldsymbol{\theta}_j)\}$.

For $\mathcal{M}_2$

Mean

$$\mathbb{E}[\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\beta}_S; \mathbf{X}^e, D^c] = \boldsymbol{m}_{\boldsymbol{y}}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) = \boldsymbol{H}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)\boldsymbol{\beta}_S = \begin{pmatrix} \boldsymbol{H}_S(\mathbf{X}^e, \boldsymbol{\theta}) \\ \boldsymbol{H}_S(D^c) \end{pmatrix} \boldsymbol{\beta}_S$$

and the covariance

$$\mathbb{V}ar[\boldsymbol{y}|\boldsymbol{\theta}, \Phi, \sigma_{err}^2; \mathbf{X}^e, D^c] = \boldsymbol{V}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) = \begin{pmatrix} \Sigma_{exp,exp}(\mathbf{X}^e, \boldsymbol{\theta}) + \sigma_{err}^2 \boldsymbol{I}_{n_e} & \Sigma_{exp,c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c) \\ \Sigma_{exp,c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)^T & \Sigma_{c,c}(D^c) \end{pmatrix}$$

Modularization

Advocated in [Liu et al., 2009].

- From $\mathbf{y}^c$, compute the GP emulator from the partial Likelihood $\mathcal{L}^M(\boldsymbol{\beta}_S, \Phi_S; \mathbf{y}^c, D^c)$,
- Plug the GP emulator in the conditional Likelihood of $\mathbf{y}^e$: $\mathcal{L}^C(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta; \boldsymbol{\beta}_S, \Phi_S, \mathbf{y}^e|\mathbf{y}^c, \mathbf{X}^e, D^c)$.

[Gramacy, 2020] “Modularization or Compartmentalization is an engineering practice such that Components should perform robustly in isolation, irrespective of their anticipated role in a larger system.”

MLE estimates

MLE for $\boldsymbol{\beta}_S, \Phi_S$ from the partial likelihood:

$$\mathcal{L}^M(\boldsymbol{\beta}_S, \Phi_S; \mathbf{y}^c, D^c)$$
$$= \frac{1}{(2\pi)^{N/2}|\boldsymbol{V}_c(D^c)|^{1/2}} \exp\left\{ -\frac{1}{2}\Big(\mathbf{y}^c - \boldsymbol{m}_c(D^c)\Big)^T \boldsymbol{V}_c(D^c)^{-1}\Big(\mathbf{y}^c - \boldsymbol{m}_c(D^c)\Big)\right\}.$$

$$\mathbb{V}ar[\mathbf{y}^c|\Phi_S; D^c] = \boldsymbol{V}_c^{\Phi_S}(D^c) = \boldsymbol{V}_c(D^c) = \Sigma_{c,c}(D^c),$$

$$\mathbb{E}[\mathbf{y}^c|\boldsymbol{\beta}_S; D^c] = \boldsymbol{m}_c(D^c) = \boldsymbol{H}_S(D^c)\boldsymbol{\beta}_S\,.$$

GP emulation

We derive

$$\mathbf{y}^e|\mathbf{y}^c \sim \mathcal{N}(\boldsymbol{\mu}_{exp|c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c),\Sigma_{exp|c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c))$$

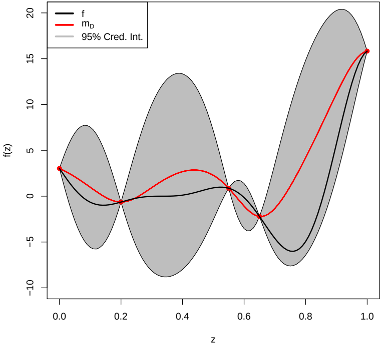
with

$$\begin{aligned}&\boldsymbol{\mu}_{exp|c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c)\\&=\boldsymbol{H}_S(\mathbf{X}^e,\boldsymbol{\theta})\boldsymbol{\beta}_S+\boldsymbol{H}_\delta(\mathbf{X}^e)\boldsymbol{\beta}_\delta+\Sigma_{exp,c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c)\Sigma_{c,c}(D^c)^{-1}(\mathbf{y}^c-\boldsymbol{m}_c(D^c)),\end{aligned}$$

$$\Sigma_{exp|c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c)=\boldsymbol{V}_{exp,exp}(\mathbf{X}^e,\boldsymbol{\theta})-\Sigma_{exp,c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c)\Sigma_{c,c}(D^c)^{-1}\Sigma_{exp,c}((\mathbf{X}^e,\boldsymbol{\theta}),D^c)^T,$$

with
$\boldsymbol{V}_{exp,exp}(\mathbf{X}^e,\boldsymbol{\theta})=\Sigma_{exp,exp}(\mathbf{X}^e,\boldsymbol{\theta})+\Sigma_\delta(\mathbf{X}^e)+\sigma^2_{err}\boldsymbol{I}_{n_e}.$

GP emulator illustrated



Conditional Likelihood

GP emulator plugged into

$$\mathcal{L}^C(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta, \sigma^2_{err}; \hat{\boldsymbol{\beta}}_S, \hat{\Phi}_S, \mathbf{y}^e | \mathbf{y}^c, \mathbf{X}^e, D^c)$$
$$\propto |\Sigma_{exp|c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)|^{-1/2}$$
$$\exp\Big\{ -\frac{1}{2}(\mathbf{y}^e - \mu_{exp|c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c))^T \Sigma_{exp|c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)^{-1} (\mathbf{y}^e - \mu_{exp|c}((\mathbf{X}^e, \boldsymbol{\theta}), D^c)) \Big\}.$$

Frequentist estimation: least squares estimation

[Cox et al., 2001, Wong et al., 2017]

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \mathcal{Q}}{argmin}\, M_n(\boldsymbol{\theta}) \quad \text{with} \quad M_n(\boldsymbol{\theta}) = \frac{1}{n_e} \sum_{i=1}^{n_e} \{y_i^e - f(\mathbf{x}_i^e, \boldsymbol{\theta})\}^2, \tag{2}$$

and
estimation of δ_0 applying any nonparametric regression method to the “data”

$$\{\mathbf{x}_i, y_i^e - f(\mathbf{x}_i^e, \hat{\boldsymbol{\theta}})\}_{i=1,\ldots,n_e}.$$

Bayesian estimation for $\mathcal{M}_4$

Prior information:

$$\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \Phi, \sigma_{err}^2) = \pi(\boldsymbol{\theta}) \times 1 \times \pi(\Phi) \times \pi(\sigma_{err}^2).$$

Estimation

- Full Bayesian: For a full Bayesian analysis, integrating other parameters out is needed to finally get $\pi(\boldsymbol{\theta}|\boldsymbol{y})$,
- Modular:
  1. maximizing the likelihood $\mathcal{L}^M(\boldsymbol{\beta}_S, \Phi_S|\mathbf{y}^c; D^c)$ to get the maximum likelihood estimates (MLE) $\hat{\boldsymbol{\beta}}_S$ and $\hat{\Phi}_S$ of $\boldsymbol{\beta}_S$ and Φ_S
  2. plugged into the conditional likelihood $\mathcal{L}^C(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta, \sigma_{err}^2; \hat{\boldsymbol{\beta}}_S, \hat{\Phi_S}, \mathbf{y}^e|\mathbf{y}^c, \mathbf{X}^e, D^c)$
  3. sampled with MCMC methods:
     $\pi(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta, \sigma_{err}^2|\mathbf{y}^e, \mathbf{y}^c, \mathbf{X}^e, D^c) \propto$
     $\mathcal{L}^C(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta, \sigma_{err}^2; \hat{\boldsymbol{\beta}}_S, \hat{\Phi_S}, \mathbf{y}^e|\mathbf{y}^c, \mathbf{X}^e, D^c) \cdot \pi(\boldsymbol{\theta}, \boldsymbol{\beta}_\delta, \Phi_\delta, \sigma_{err}^2)$
- Generate explicitely realizations of $(\delta(\mathbf{x}_i^e))_{1 \leq i \leq n_e}$ conditionally on the current parameters values in a Gibbs sampling algorithm. [Bayarri et al., 2007].

for other models

For $\mathcal{M}_1$ and $\mathcal{M}_3$ no modularization use the actual code in the likelihood.

Outline

A word on history matching

- common alternative to KOH calibration [Craig et al., 1997, Vernon et al., 2010, Boukouvalas et al., 2014, Andrianakis et al., 2017].
- searches for inputs where the simulator outputs closely match observed data, while recognizing the presence of the various uncertainties, including model discrepancy
- rules out “implausible” inputs,

$\boldsymbol{\theta}$ is deemed implausible if:

$$\frac{||y^e - f(\mathbf{x}^e, \boldsymbol{\theta})||}{\sqrt{\sigma_S^2(\mathbf{x}^e, \boldsymbol{\theta}) + \sigma_\delta^2(\mathbf{x}^e) + \sigma_{err}^2}} \geq 3, \tag{3}$$

where $\sigma_S^2, \sigma_\delta^2$, and σ_{err}^2 are the variances of the surrogate, the model discrepancy, and the observational error.

- number 3 comes from [Pukelsheim, 1994] who shows that at least 95% of any unimodal distribution is contained within three standard deviations,
- HM can be repeated in so-called “waves”, using non-implausible $\boldsymbol{\theta}$ found at one wave to generate simulation runs for the next wave
- simulator not valid if plausible space is void.

Andrianakis, I., McCreesh, N., Vernon, I., McKinley, T. J., Oakley, J. E., Nsubuga, R. N., Goldstein, M., and White, R. G. (2017).
Efficient history matching of a high dimensional individual-based HIV transmission model.
SIAM/ASA Journal on Uncertainty Quantification, 5(1):694–719.

Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C.-H., and Tu, J. (2007).
A framework for validation of computer models.
Technometrics, 49(2):138–154.

Boukouvalas, A., Sykes, P., Cornford, D., and Maruri-Aguilar, H. (2014).
Bayesian precalibration of a large stochastic microsimulation model.
IEEE Transactions on Intelligent Transportation Systems, 15(3):1337–1347.

Carmassi, M., Barbillon, P., Chiodetti, M., Keller, M., and Parent, E. (2019).
Bayesian calibration of a numerical code for prediction.
Journal de la société française de statistique, 160(1):1–30.

Cox, D. D., Park, J.-S., and Singer, C. E. (2001).
A statistical method for tuning a computer code to a data base.
Computational statistics & data analysis, 37(1):77–92.

Craig, P. S., Goldstein, M., Seheult, A. H., and Smith, J. A. (1997).

Pressure matching for hydrocarbon reservoirs: a case study in the use of bayes linear strategies for large computer experiments.
In Case Studies in Bayesian Statistics, pages 37–93. Springer.

Gramacy, R. B. (2020).
Surrogates: Gaussian process modeling, design, and optimization for the applied sciences.
CRC press.

Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., and Ryne, R. D. (2004).
Combining field data and computer simulations for calibration and prediction.
SIAM Journal on Scientific Computing, 26(2):448–466.

Kennedy, M. C. and O'Hagan, A. (2001).
Bayesian calibration of computer models.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3):425–464.

Liu, F., Bayarri, M., Berger, J., et al. (2009).
Modularization in bayesian analysis, with emphasis on analysis of computer models.
Bayesian Analysis, 4(1):119–150.

Pukelsheim, F. (1994).
The three sigma rule.

The American Statistician, 48(2):88–91.

Robert, C. P., Casella, G., and Casella, G. (1999).
Monte Carlo statistical methods, volume 2.
Springer.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989).
Design and analysis of computer experiments.
Statistical science, pages 409–423.

Vernon, I., Goldstein, M., Bower, R. G., et al. (2010).
Galaxy formation: a Bayesian uncertainty analysis.
Bayesian Analysis, 5(4):619–669.

Wong, R. K., Storlie, C. B., and Lee, T. (2017).
A frequentist approach to computer model calibration.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79(2):635–648.