1. Bayesian Calibration
2. Dimension-Reduction approach to Calibration
3. Bayesian Calibration of Stochastic Simulators
4. Future work, etc.

M. T. Pratola
Email: mpratola@stat.osu.edu
and
(3) O. Chkrebtii
Email: chkrebtii@stat.osu.edu
and (4) Y. Sun
Email: sunwards@kaust.edu

Monday February 26^{th}, 2018

Bayesian Calibration of Mathematical-Physical Models

A scientist/engineer/etc. postulates a theoretical model, $x^*, \frac{\partial x^*}{\partial z}, \frac{\partial^2 x^*}{\partial z \partial s}$, to explain a real-world process, y.

For instance, the model may explain the time evolution of a spatial field $x^*(s, z)$ using constraints $\frac{\partial x^*}{\partial z}|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = g(x^*, f(x^*, \theta), \theta)$ parameterized by $\boldsymbol{\theta}$. This is often the case in ODE/PDE models.

In calibration experiments, one wishes to combine x^* with observations y^f to estimate $\boldsymbol{\theta}$, predict the field at out-of-sample locations and/or times, etc...

Bayesian Calibration of Mathematical-Physical Models

There are, of course, a few issues:

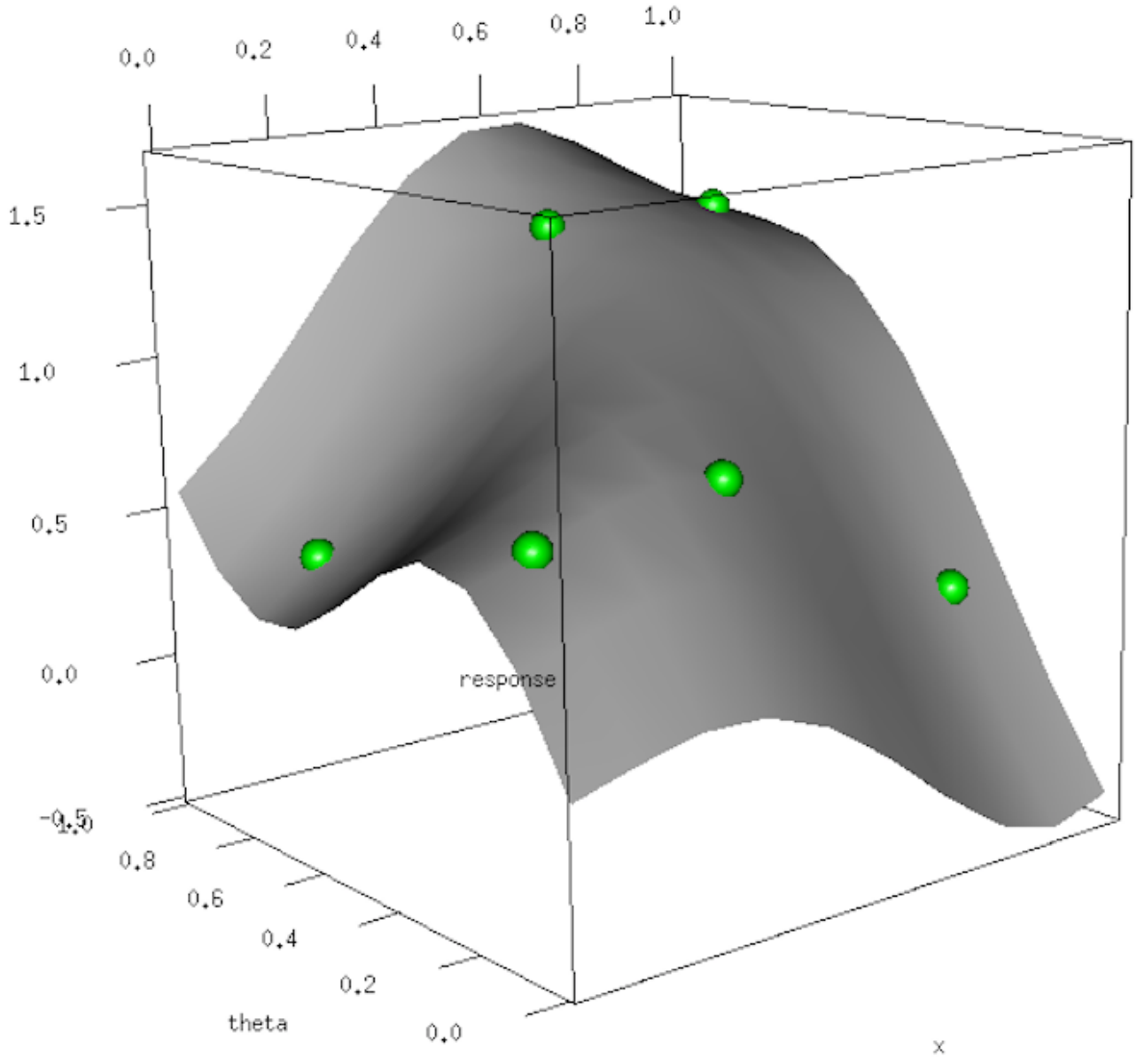
- The theoretical model need be approximated on computer
  $\longrightarrow$ the simulation model $x(\mathbf{s};\boldsymbol{\theta}) \approx x^*$
- The theoretical model is wrong
  $\longrightarrow$ $E[y^f] \neq x^*$, the case of model discrepancy
- x can only be sparsely sampled in $\boldsymbol{\theta}$-space
  $\longrightarrow$ the case of emulation
- The observational process is observed with error $\boldsymbol{\epsilon}$
- The datasets and simulator outputs can be huge
  $\longrightarrow$ Big Data
- The model output may consist of multiple states

Bayesian Calibration of Mathematical-Physical Models

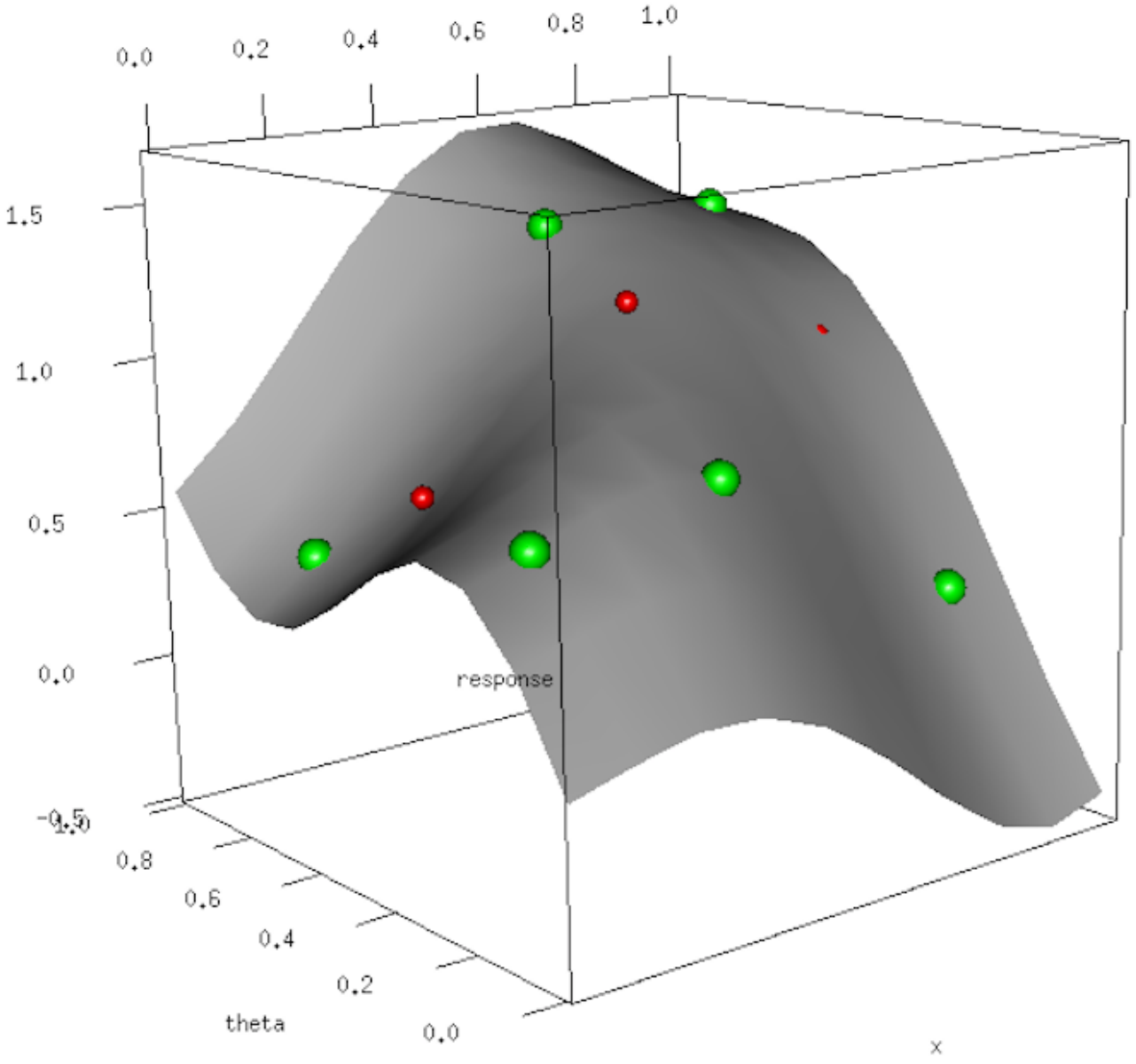
In addition...

- Exploring the tradeoff between discrepancy and parameter estimates is hard
- Existing statistical calibration techniques do not incorporate all of the above uncertainties

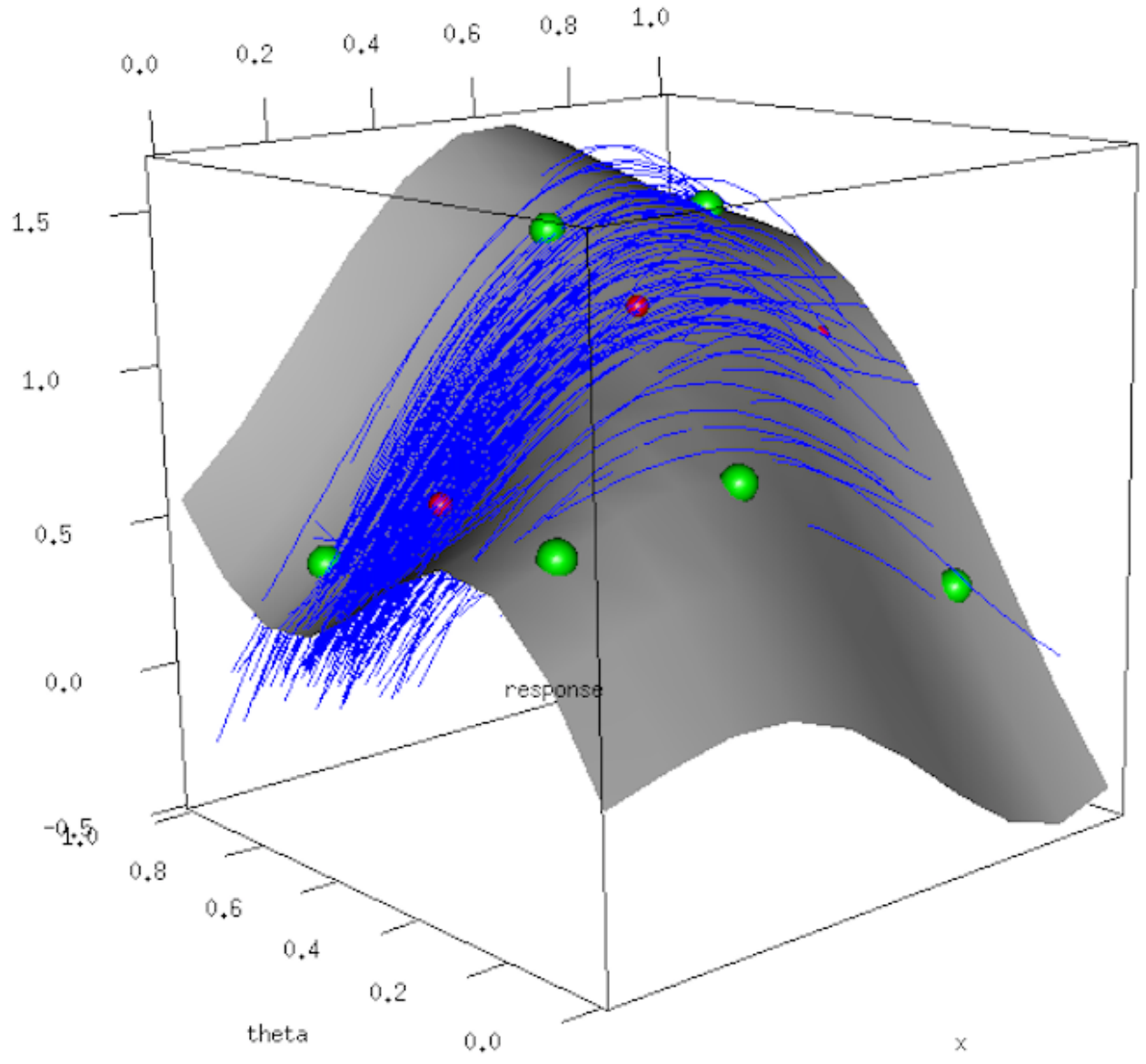
Calibration in Pictures



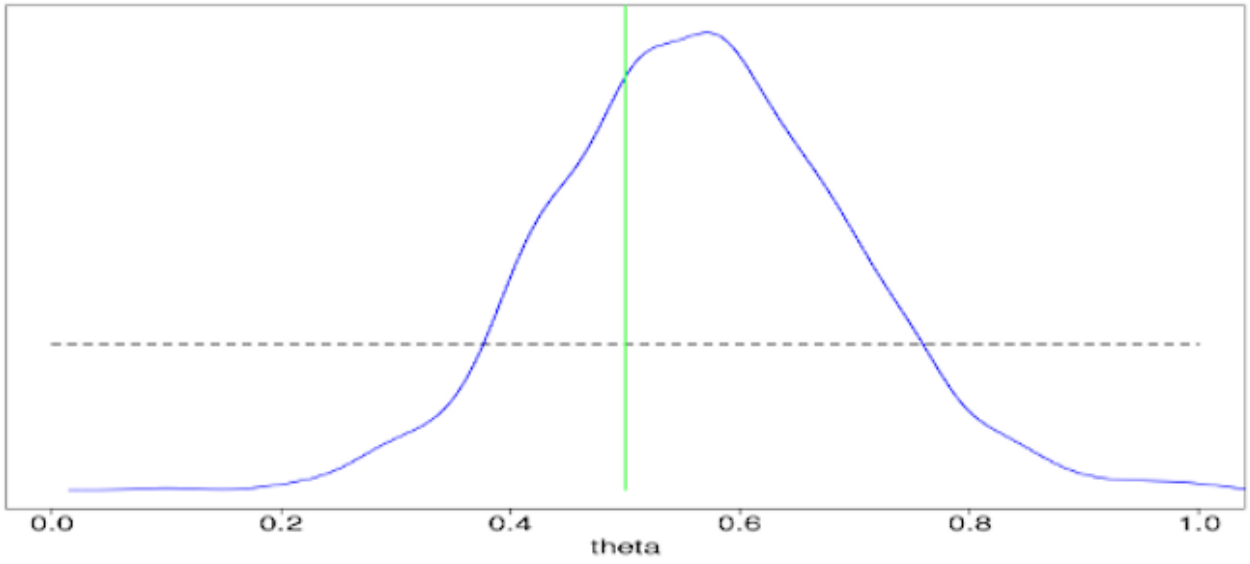
Calibration in Pictures

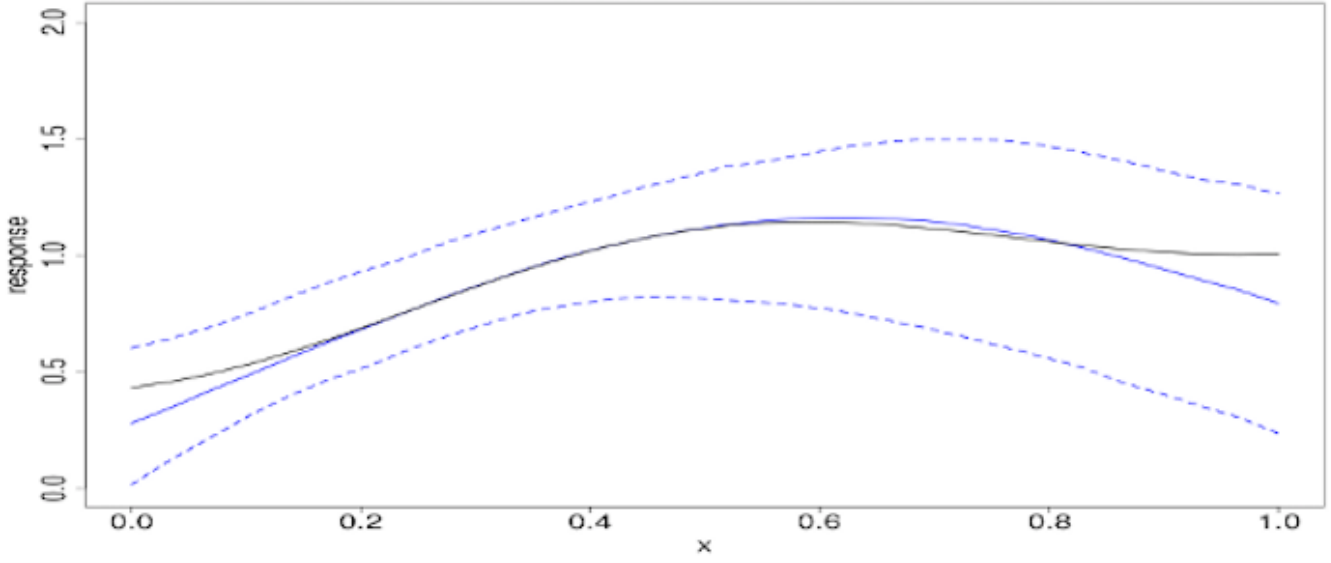


Calibration in Pictures



Calibration in Pictures





1. The Kennedy& O'Hagan Framework

- Model for the field observations is

$$y^f(\mathbf{s}_i) = \eta(\mathbf{s}_i, \boldsymbol{\theta}) + \delta(\mathbf{s}_i) + \epsilon(\mathbf{s}_i), \quad i = 1, \ldots, n$$

where $\epsilon(\mathbf{s}_i) \sim N(0, \lambda_f^{-1})$, $\delta(\mathbf{s}_i)$ accounts for the discrepancy between the simulator and reality and $\boldsymbol{\theta}$ denotes the "true" (or best in some sense) setting of the calibration parameter $\mathbf{t}$.

† M.A. Kennedy and T. O'Hagan: *Bayesian Calibration of Computer Models (with discussion)*, Journal of the Royal Statistical Society, Series B, vol.68, pp.425–464 (2001).

CMCE Model, no discrepancy ($\delta(\mathbf{s}) = 0$).

- Besides our model for the observations, we also need a model for the simulator outputs.
- Since the simulator is slow, we will have to emulate it.
- We have field data,

$$\mathbf{y}^f = (y^f(\mathbf{s}_1), \ldots, y^f(\mathbf{s}_n))^T$$

- And simulator output,

$$\mathbf{y}^c = (y^c(\mathbf{s}_1, \mathbf{t}_1), \ldots, y^c(\mathbf{s}_m, \mathbf{t}_m))^T$$

- With no discrepancy, our model for the field is

$$y^f(\mathbf{s}_i) = \eta(\mathbf{s}_i, \boldsymbol{\theta}) + \epsilon_i$$

and our model for the simulator is

$$y^c(\mathbf{s}_i, \mathbf{t}_i) = \eta(\mathbf{s}_i, \mathbf{t}_i)$$

CMCE Model, no discrepancy ($\delta(\mathbf{s}) = 0$).

- Use our usual emulator model for the simulator, a GP:

$$\eta(\mathbf{s}, \mathbf{t}) \sim GP(\mu(\mathbf{s}, \mathbf{t}), \lambda^{-1}\mathbf{R}(\mathbf{s}, \mathbf{t}; \boldsymbol{\rho}))$$

  where $\mathbf{R}(\mathbf{s}, \mathbf{t}; \boldsymbol{\rho})$ is formed as

$$\text{cor}(\eta(\mathbf{s}, \mathbf{t}), \eta(\mathbf{s}', \mathbf{t}')) = \prod_{i=1}^{d} c(\mathbf{s} - \mathbf{s}') \prod_{j=1}^{k} c(\mathbf{t} - \mathbf{t}')$$

  for $\mathbf{s} \in \mathbb{R}^d$ and $\mathbf{t} \in \mathbb{R}^k$.

- A typical choice for the correlation function $c()$ will be the Gaussian:

$$c(h_i) = \rho_i^{||h_i||^2}$$

  where $h_i = s_i - s_i'$ for correlation parameter $\rho_i \in (0, 1)$.

CMCE Model, no discrepancy ($\delta(\mathbf{s}) = 0$).

- This gives us our model (and correspondingly the likelihood) for the field and simulator data,

$$\begin{pmatrix} \mathbf{y}^f \\ \mathbf{y}^c \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(\mathbf{s}, \boldsymbol{\theta}) \\ \mu(\mathbf{s}, \mathbf{t}) \end{pmatrix}, \lambda^{-1} \begin{bmatrix} \mathbf{R}^{ff} & \mathbf{R}^{fc} \\ \mathbf{R}^{cf} & \mathbf{R}^{cc} \end{bmatrix} + \begin{bmatrix} \lambda_f^{-1}\mathbf{I} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

  Here, $\mathbf{R}^{ff}$ denotes the correlation elements between field observations, $\mathbf{R}^{cc}$ the correlation between simulator outputs and $\mathbf{R}^{fc}$ the cross-correlation between field observations and simulator outputs.

CMCE Model, no discrepancy ($\delta(\mathbf{s}) = 0$).

- For simplicity let's take $\mu(\mathbf{s}, \mathbf{t}) = 0$.
- Specifying priors on the parameters $\boldsymbol{\rho}, \lambda, \lambda_f$ and the calibration parameters $\boldsymbol{\theta}$ we have

$$\pi(\boldsymbol{\theta}, \lambda, \lambda_f, \boldsymbol{\rho} | \mathbf{y}^f, \mathbf{y}^c) \propto L(\cdot | \mathbf{y}^f, \mathbf{y}^c) \pi(\lambda) \pi(\lambda^f) \prod_{i=1}^{k} \pi(\theta_i) \prod_{j=1}^{d+k} \pi(\rho_j)$$

- Common prior specification is

$$\pi(\lambda) = \text{Gamma}(a, b)$$

$$\pi(\lambda^f) = \text{Gamma}(a_f, b_f)$$

$$\pi(\rho_j) = \text{Beta}(\alpha_j, \beta_j)$$

- And we also need a prior on the calibration parameters,

$$\pi(\theta_i) = \text{Unif}(0, 1)$$

(assuming the inputs are scaled to the unit hypercube).

CMCE Model, no discrepancy ($\delta(\mathbf{s}) = 0$).

- What does this model do? Consider predicting the field process at a new location $\mathbf{s}^*$ (for a given $\boldsymbol{\theta}$).
- Let $\mathbf{c}^T = (\text{cov}(y^f(\mathbf{s}^*), y^f(\mathbf{s}_1)), \ldots, \text{cov}(y^f(\mathbf{s}^*), y^f(\mathbf{s}_n)), \text{cov}(y^f(\mathbf{s}^*), y^c(\mathbf{s}_1)), \ldots$
- Or in short-hand, $\mathbf{c}^T = (\mathbf{c}^f, \mathbf{c}^c)^T$.
- Then the mean of the conditional predictive distribution is

$$\begin{aligned} E[y^f(\mathbf{s}^*)|\mathbf{y}^f, \mathbf{y}^c, \cdot] &= \mathbf{c}^T \Sigma^{-1} (\mathbf{y}^f, \mathbf{y}^c)^T \\ &= \vdots \\ &= \sum_{i=1}^{n} w_i^f(\boldsymbol{\theta}) y^f(\mathbf{s}_i) + \sum_{j=1}^{m} w_j^c(\boldsymbol{\theta}) y^c(\mathbf{s}_j, \mathbf{t}_j) \end{aligned}$$

CMCE Model, no discrepancy ($\delta(\mathbf{s}) = 0$).

- This shows that the field process is predicted as a weighted combination of the field observations and simulator output.
- The role of the estimated calibration parameter, $\boldsymbol{\theta}$, comes through the cross-covariance terms, $\mathbf{c}^c$ and Σ^{cf} which both depend on $\boldsymbol{\theta}$.
- If the estimated $\boldsymbol{\theta}$ indicates the field data is "far" from the simulator output, i.e. $|\theta_j - t_j|$ is large $\forall j$, then these correlation components will be small and the field prediction is mainly based on the field observations.
- In extreme case of $\mathbf{c}^c = 0$ and $\Sigma^{cf} = 0$ we get $E[y^f(\mathbf{s}^*)] = \mathbf{c}^{f\,T}\Sigma^{f\,-1}\mathbf{y}^f$, the usual GP predictor.
- If the estimate of $\boldsymbol{\theta}$ is poor, the prediction of the field process may be inappropriately influenced by the simulator outputs if they receive too much weight – i.e. model things the outputs and field are "closer" than the actually are.

CMCE Model, with discrepancy

- Popular form of discrepancy is to assume an additive discrepancy,

$$y^f(\mathbf{s}_i) = \eta(\mathbf{s}_i, \boldsymbol{\theta}) + \delta(\mathbf{s}_i) + \epsilon_i$$

- Naturally, we will model the discrepancy, $\boldsymbol{\delta} = (\delta(\mathbf{s}_i), \ldots, \delta(\mathbf{s}_n))$ also as a GP,

$$\boldsymbol{\delta} \sim N\left(\mu_\delta(\mathbf{s}), \lambda_\delta^{-1}\mathbf{R}_\delta(\mathbf{s}; \boldsymbol{\phi})\right)$$

- Assuming η, δ and ϵ are independent, the likelihood becomes

$$\begin{pmatrix} \mathbf{y}^f \\ \mathbf{y}^c \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(\mathbf{s}, \boldsymbol{\theta}) + \mu_\delta(\mathbf{s}) \\ \mu(\mathbf{s}, \mathbf{t}) \end{pmatrix}, \Sigma\right)$$

where

$$\Sigma = \lambda^{-1}\begin{bmatrix} \mathbf{R}^{ff} & \mathbf{R}^{fc} \\ \mathbf{R}^{cf} & \mathbf{R}^{cc} \end{bmatrix} + \begin{bmatrix} \lambda_\delta^{-1}\mathbf{R}_\delta & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \lambda_f^{-1}\mathbf{I} & 0 \\ 0 & 0 \end{bmatrix}$$

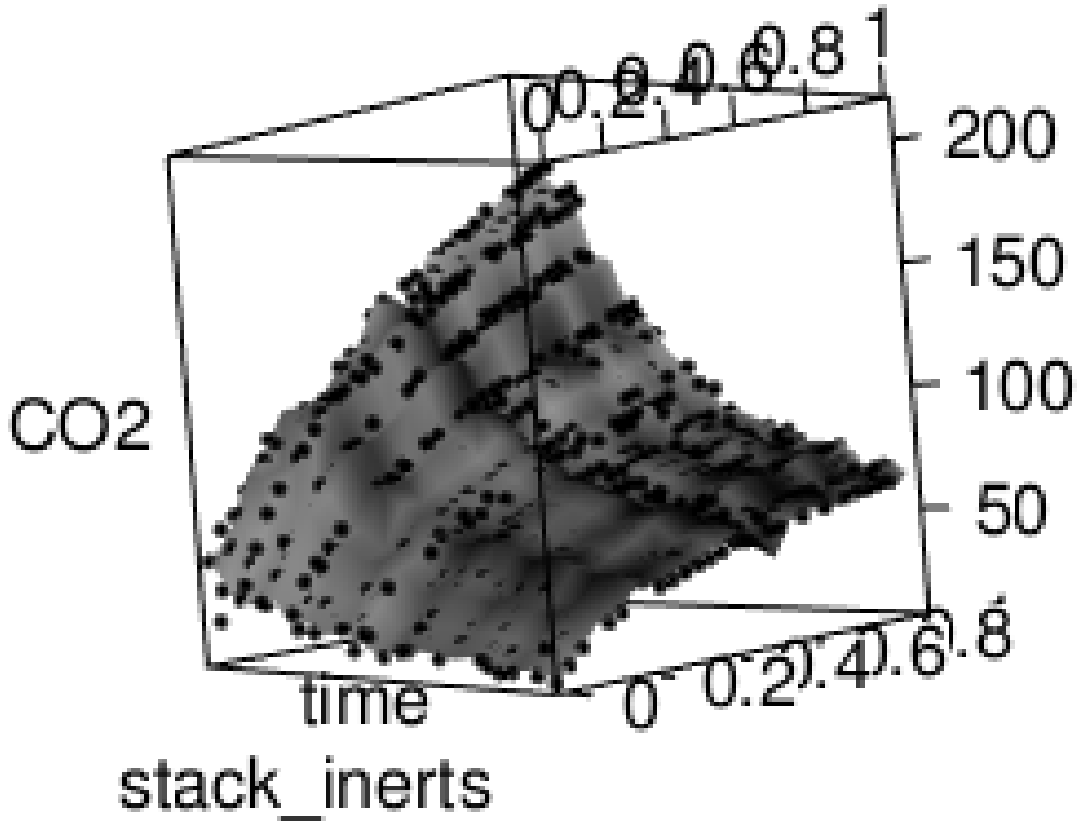
Prediction and Inference

- We are typically interested in:
- the emulated calibrated simulator, $E[\eta(\mathbf{s}, \boldsymbol{\theta})|\mathbf{y}^f, \mathbf{y}^c]$
- the predicted discrepancy, $E[\delta(\mathbf{s})|\mathbf{y}^f, \mathbf{y}^c]$
- the predicted field process, $E[\eta(\mathbf{s}, \boldsymbol{\theta}) + \delta(\mathbf{x})|\mathbf{y}^f, \mathbf{y}^c]$
- the estimated calibration parameter, $E[\boldsymbol{\theta}|\mathbf{y}^f, \mathbf{y}^c]$
- And of course uncertainties in the above.
- There are other forms of discrepancy that have been considered, such as multiplicative and more complex forms, but these are generally less common.

2. The Higdon et al. Model

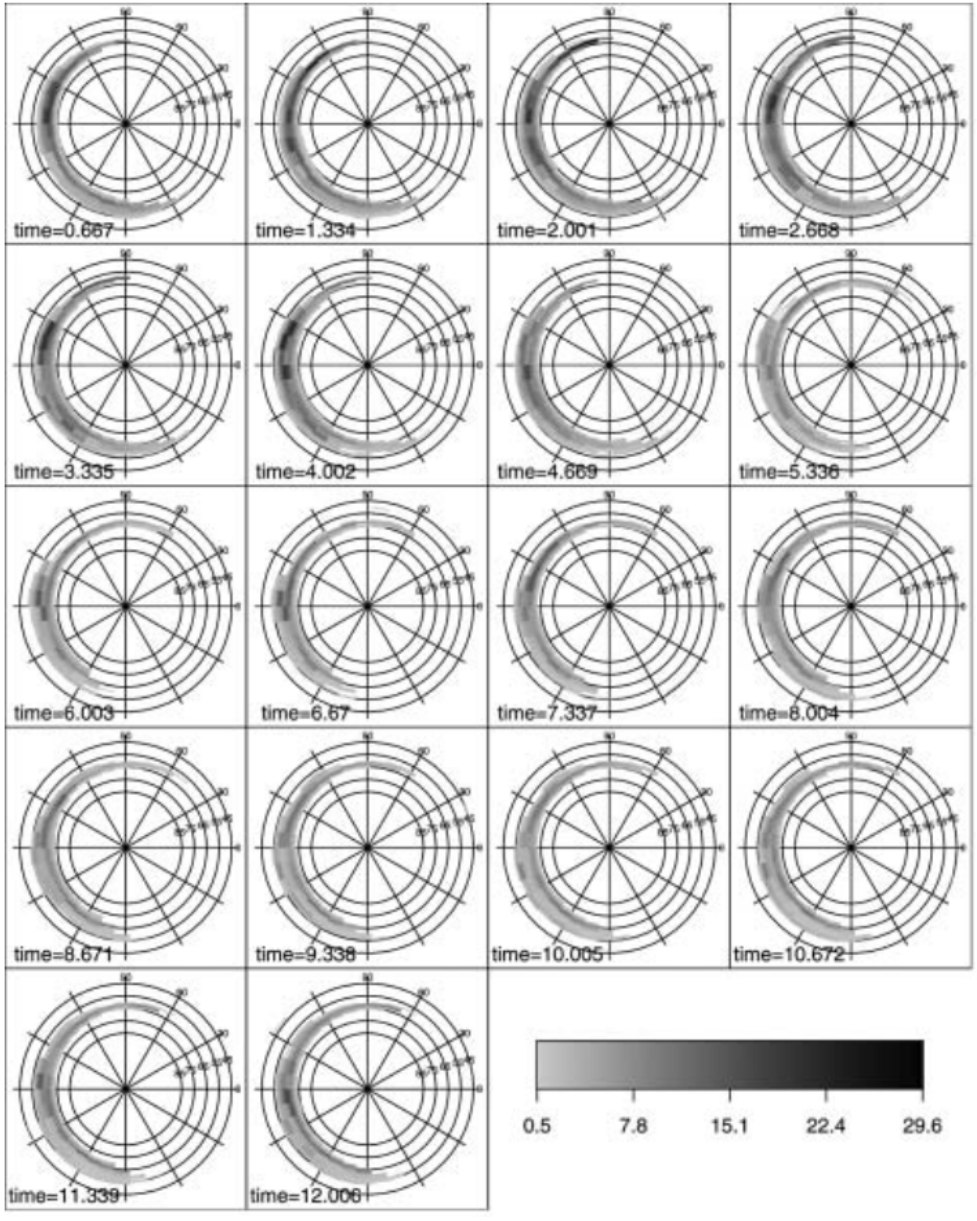
- Often our computer models are, say, spatial-temporal processes observed and computed over a dense grid in $\mathbf{s}$.
- At the same time, we are limited in how many runs of the computer model we can make at different $\boldsymbol{\theta}$'s.
- For instance, for a given set of calibration parameters ($\boldsymbol{\theta}$'s), a climate simulator may generate a dense grid of preciptation and temperature fields over the lattitude, longitude of the Earth and at many timepoints.
- Modeling such objects directly using the KOH approach is going to be infeasible.

The Higdon et al. Model



The Higdon et al. Model

The Higdon et al. Model



The Higdon et al. Model

- Higdon et al. (2008) propose to use a dimension-reduction technique to alleviate this challenge for densely observed model outputs.
- The idea is to remove directly modeling the dense outputs within the GP model, whch creates huge correlation matrices which are difficult and computationally expensive to store, manipulate and invert.

The Higdon et al. Model

- For a given output computing at setting $\boldsymbol{\theta}_j$ over the dense spatial-temporal grid $\mathbf{s}$, take the output and put it in a vector, say $\mathbf{x}_j$.
- Do this for all $j = 1, \ldots, m$ runs of the simulator in your budget for $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m$.
- Stack these vectors in the matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_m]$.
- Decompose the matrix as $\mathbf{X} = \sum_{l=1}^{n_c} V_{lj} U_l$
- Insight: think of V_{lj} as $V_l(\boldsymbol{\theta}_j)$.

The Higdon et al. Model

- The model becomes:

$$y^f(\mathbf{s}_i) = \sum_{l=1}^{n_c} V_l(\boldsymbol{\theta}) U_l + \delta(\mathbf{s}_i) + \epsilon_i$$

  and the discrepancy is modeled as before,

$$\boldsymbol{\delta} \sim \text{GP}(\mu_\delta(\mathbf{s}), \lambda_\delta^{-1} \mathbf{R}_\delta(\mathbf{s}; \phi)$$

- The V_{lj}'s are modeled using a GP. Writing $\mathbf{V}_l = (V_l(\boldsymbol{\theta}_1), \ldots, V_l(\boldsymbol{\theta}_m), V_l(\boldsymbol{\theta}))$, the model for these weights is

$$\mathbf{V}_l \sim \text{GP}(\boldsymbol{\mu}_{V_l}, \lambda_{V_l}^{-1} \mathbf{R}_{V_l})$$

  where $\mathbf{R}_{V_l}$ will again make use of a correlation function.
- Key point is that the $\mathbf{V}_l$'s are of much lower dimension than the actual model outputs!
- Select n_c in some reasonable manner.

Bayesian Calibration of Mathematical-Physical Models

What about simulator uncertainty?

- In some cases, $x(\mathbf{s};\boldsymbol{\theta})$ may be an approximation of the desired mathematical model, say arising from solutions to ODE/PDE's using numerical methods. In our case, these solutions are solved probabilistically using the "PODES" method of Chkrebtii et al. (2014), and we have available posterior solution realizations.

Bayesian Calibration of Mathematical-Physical Models

What about simulator uncertainty?

- Or, the model $x(\mathbf{s};\boldsymbol{\theta})$ may itself be stochastic. For instance, agent based models are stochastic models popular in finance and biology. In this case, we have available solution realizations of the stochastic model.

Stochastic Water Temperature Model

- Models for river water temperature are used for ecological and conservation studies.
- Particularly relevant in research trying to understand the effects of climate change on wildlife, etc.
- State of the art models are made up of a deterministic component and stochastic component:

$$T_w(s) = T_a(s) + R_w(s)$$

where

$$T_a(s) = a_1 + a_2 sin\left(\frac{2\pi}{365}(s - s_0)\right)$$

$$R_w(s) = KR_a(s) + \epsilon$$

Of particular interest is the air-water interface thermal diffusivity parameter, K.

JAK-STAT Intracellular Signaling Pathway Model

- The complex mechanism of gene transcription consists of a series of biochemical reactions. The process is reversible, so that after transcription occurs, the chemicals return to their original state so the process may repeat.
- Currently, this process is modeled using a delay differential equation model consisting of four states:

$$
\begin{array}{lcll}
\frac{d}{ds}x_1(s;\boldsymbol{\theta}) & = & -\theta_1\, x_1(s;\boldsymbol{\theta})\, \mathrm{Epo}R_A(s;\boldsymbol{\theta}) + \ 2\,\theta_4\, x_4(s-\theta_5), & s \in [0,60], \\
\frac{d}{ds}x_2(s;\boldsymbol{\theta}) & = & \theta_1\, x_1(s;\boldsymbol{\theta})\, \mathrm{Epo}R_A(s;\boldsymbol{\theta}) - \theta_2\, {x_2}^2(s;\boldsymbol{\theta}), & s \in [0,60], \\
\frac{d}{ds}x_3(s;\boldsymbol{\theta}) & = & -\theta_3\, x_3(s;\boldsymbol{\theta}) + \ \frac{1}{2}\,\theta_2\, {x_2}^2(s;\boldsymbol{\theta}), & s \in [0,60], \\
\frac{d}{ds}x_4(s;\boldsymbol{\theta}) & = & \theta_3\, x_3(s;\boldsymbol{\theta}) - \theta_4\, x_4(s-\theta_5;\boldsymbol{\theta}), & s \in [0,60], \\
x_1(s;\boldsymbol{\theta}) & = & \theta_6, & s \in [-\theta_5,0], \\
x_i(s;\boldsymbol{\theta}) & = & 0, \ i=2,3,4, & s \in [-\theta_5,0]
\end{array}
$$

JAK-STAT Intracellular Signaling Pathway Model

Symbol	Description	Prior
$\theta_1, \ldots, \theta_4$	Reaction rates of states 1-4	χ^2_1
θ_5	Time delay	χ^2_6
θ_6	Initial conc. of the state 1	$\mathrm{N}\left(y^{(3)}(0), 40^2\right)$
θ_7	Prior precision of probabilistic solver	$100 + \text{Log-N}\,(10, 1)$
θ_8	Length-scale of probabilistic solver	$0.12 + \mathrm{Exp}\,(0.1)$

- Note here that we have parameters of the DE system $(\theta_1, \ldots, \theta_6)$ as well as parameters of the *solver* of the DE system (θ_7, θ_8).

JAK-STAT Intracellular Signaling Pathway Model

- Measurements are available via a process called immunoblotting, with measurement models

$$
\begin{aligned}
y_1(s_{1,i}) &= \kappa_1 \left(x_2(s_{1,i}) + 2x_3(s_{1,i})\right) + \epsilon_1(s_{1,i}), && 1 \le i \le S_1, \\
y_2(s_{2,i}) &= \kappa_2 \left(x_1(s_{2,i}) + x_2(s_{2,i}) + 2x_3(s_{2,i})\right) + \epsilon_2(s_{2,i}), && 1 \le i \le S_2, \\
y_3(s_{3,i}) &= x_1(s_{3,i}) + \epsilon_3(s_{3,i}), && 1 \le i \le S_3, \\
y_4(s_{4,i}) &= x_3(s_{4,i}) \left(x_2(s_{4,i}) + x_3(s_{4,i})\right)^{-1} + \epsilon_4(s_{4,i}), && 1 \le i \le S_4,
\end{aligned}
$$

- It is of interest to recover the unknown model parameters $\boldsymbol{\theta}$ and the (multiplicative) discrepancies $\boldsymbol{\kappa}$ based on the measured data $\boldsymbol{Y}$.

Modeling *Cheap*† Stochastic Simulators

- If realizations of the simulator can be cheaply sampled, say

$$\boldsymbol{\chi}(\boldsymbol{\theta}) \sim \pi(\boldsymbol{\chi}|\boldsymbol{\theta})$$

and the likelihood for the observations is

$$\mathbf{Y}|\boldsymbol{\chi}(\boldsymbol{\theta}), \boldsymbol{\theta}, \lambda_f \sim N(\boldsymbol{\chi}(\boldsymbol{\theta}), \lambda_f^{-1}\mathbf{I})$$

where λ_f is the precision of the observations.

- Then it is feasible to directly sample the posterior using a Metropolis-within-Gibbs algorithm when the prior on λ_f is conjugate.

† ala Higdon et al. (2004)

Modeling *Cheap* Stochastic Simulators

1. Sample $\theta_j^*|\cdot$ (MH step) for $j = 1, 2, \ldots$
2. Sample $\lambda_j^*|\cdot$ (Gibbs step)

However, each draw θ_j^* requires evaluating $\boldsymbol{\chi}(\boldsymbol{\theta}_{-j}, \theta_j^*)$. When an evaluation of $\boldsymbol{\chi}$ is expensive, this is not feasible.

3. Modeling Stochastic Simulators

- Assume N independent realizations of the simulator at m settings of parameters θ are available as n-vector outputs.
- The kth state of the simulator output at parameter setting θ_j and spatial-temporal setting $\mathbf{s}_i$ is denoted $\chi_{uk}(\mathbf{s_i}, \boldsymbol{\theta}_j) \equiv \chi_{ukij}$ where $k = 1, \ldots, n_s; j = 1, \ldots, m; i = 1, \ldots, n; u = 1, \ldots, N$.
- These outputs are modelled according to an n_c-component orthogonal basis expansion,

$$\chi_{ukij} = \sum_{l=1}^{n_c} V_l(\boldsymbol{\theta}_j) U_{ukil}$$

- Here, the V_l's are assumed realizations of a Gaussian Process,

$$\mathbf{V}_l|\cdot \sim GP(0, \lambda_{v_l}^{-1}\mathbf{R}_{v_l})$$

additionally specified by correlation scale parameters.

Modeling Stochastic Simulators

- The state observations are modeled as

$$\mathbf{Y}_k|\mathbf{U}_{u,k}, \mathbf{V}(\boldsymbol{\theta}), \boldsymbol{\delta}_k \sim N(\kappa_k \mathbf{U}_{uk}\mathbf{V}(\boldsymbol{\theta}) + \boldsymbol{\delta}_k, \lambda_{f,k}^{-1}\mathbf{I}_n)$$

  where $\lambda_{f,k}$ is the observation precision of the kth state, $\boldsymbol{\delta}_k$ is an additive discrepancy and κ_k is a multiplicative discrepancy.
- We utilize conjugate priors on $\lambda_{f,k}$'s and Gaussian priors on κ_k's. GP priors are specified for $\boldsymbol{\delta}_k$,

$$\boldsymbol{\delta}_k \sim GP(\mu_{\delta_k}, \lambda_{\delta_k}^{-1}\mathbf{R}_{\delta_k}).$$

Modeling Stochastic Simulators

- The posterior distribution,

$$\pi\left(\{\boldsymbol{\theta}_j\}_{j=1}^{m}, \boldsymbol{\delta}, \boldsymbol{\kappa}, \{\lambda_{v_l}\}_{l=1}^{n_c}, \{\boldsymbol{\rho}\}_{l=1}^{n_c}, \lambda_f, \{\lambda_{\delta_k}\}_{k=1}^{n_s}\{\boldsymbol{\psi}_k\}_{k=1}^{n_s} | \mathbf{Y}, \boldsymbol{\Phi}\right)$$

$$\propto \pi\left(\mathbf{Y}|\mathbf{U}_u, \mathbf{V}(\boldsymbol{\theta}), \boldsymbol{\delta}, \boldsymbol{\kappa}\right) \prod_{l=1}^{n_c} \left(\pi\left(V_l(\boldsymbol{\theta})|\mathbf{V}_l, \lambda_{v_l}, \boldsymbol{\rho}_l, \boldsymbol{\theta}\right) \pi\left(\mathbf{V}_l|\lambda_{v_l}, \boldsymbol{\rho}_l, \boldsymbol{\theta}\right)\right)$$

$$\times \prod_{k=1}^{n_s} \pi\left(\boldsymbol{\delta}_k|\mu_{\delta_k}, \lambda_{\delta_k}, \boldsymbol{\psi}_k\right) \prod_{l=1}^{n_c} \left(\pi\left(\lambda_{v_l}\right) \pi\left(\boldsymbol{\rho}_l\right)\right) \pi\left(\lambda_f\right)$$

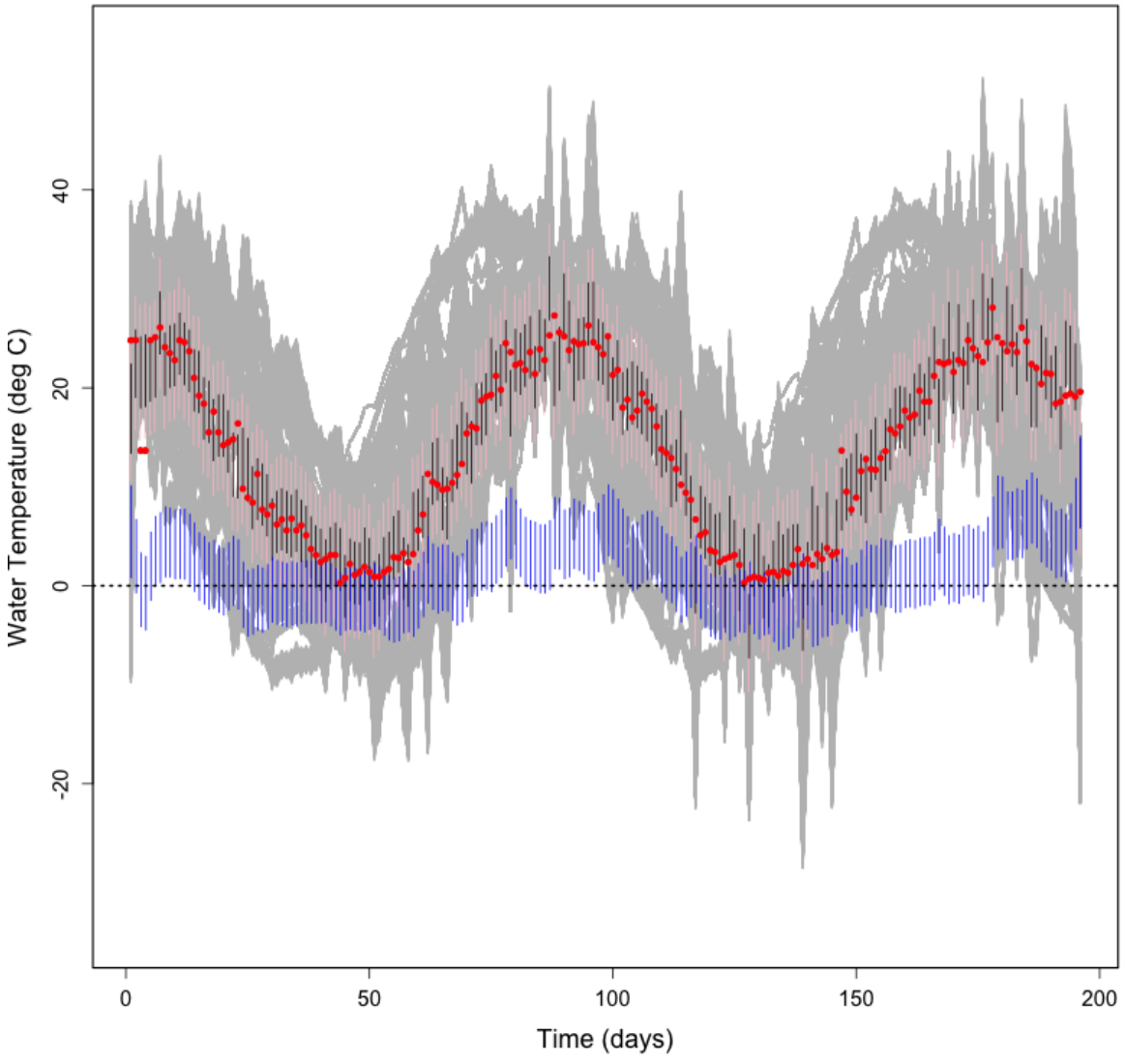
$$\times \prod_{k=1}^{n_s} \left(\pi\left(\lambda_{\delta_k}\right) \pi\left(\boldsymbol{\psi}_k\right) \pi\left(\kappa_k|\mu_{\kappa_k}, \lambda_{\kappa_k}^{-1}\right)\right) \prod_{t=1}^{q} \pi\left(\theta_t\right)$$

is sampled using an MCMC algorithm. Specifying appropriate priors on additive and multiplicative discrepancies is particularly important.

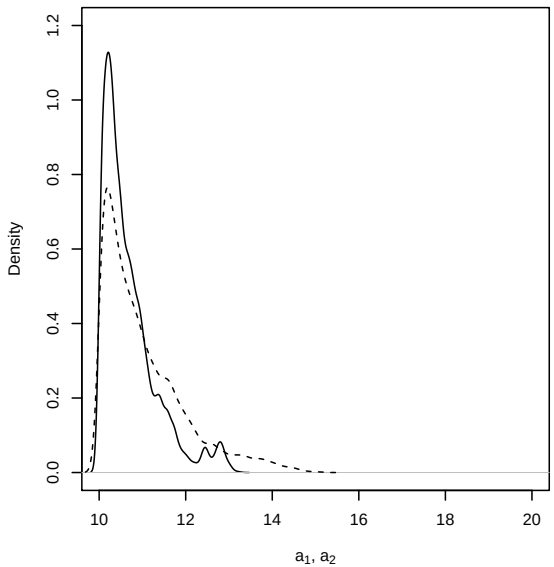
Water Temperature Example

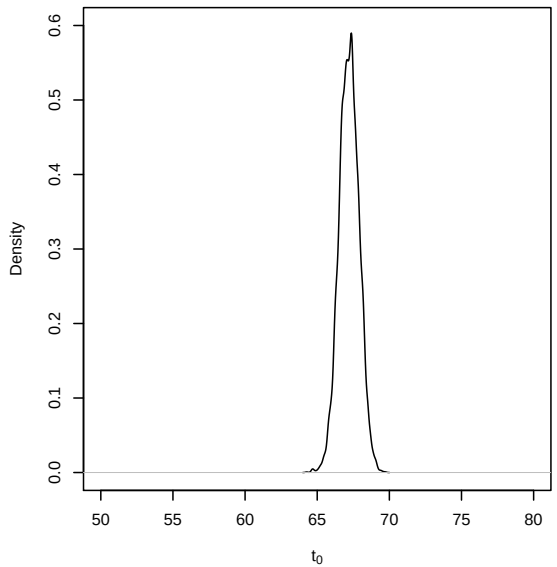
- Cassie et al (1998) consider the problem of constructing a stochastic model of small river water temperature given air temperature data.
- The stochastic model consists of 2 parameters (a_1, a_2) that shift and scale a sinusoidal time component, and a "diffusivity" parameter (K) which scales the air temperature data.
- Previous investigations show that the interpolation ability is very good. Expect weak evidence of a discrepancy.

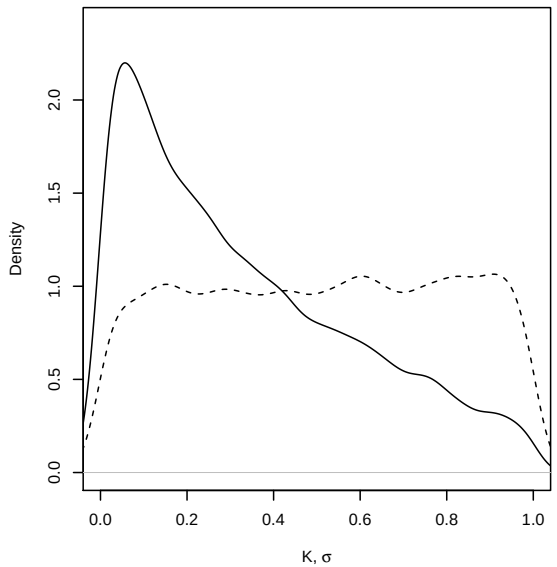
Stochastic Water Temperature Model



Stochastic Water Temperature Model



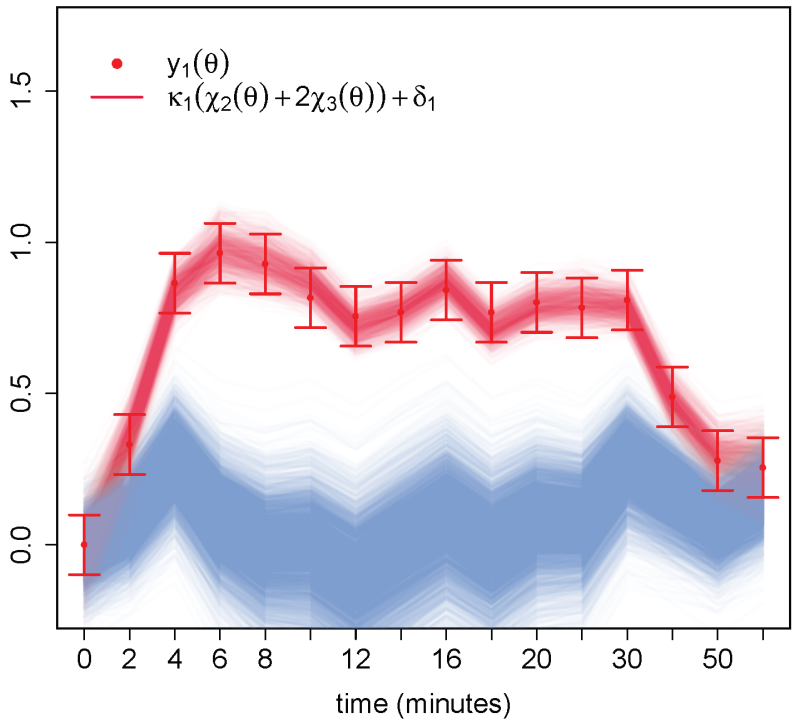


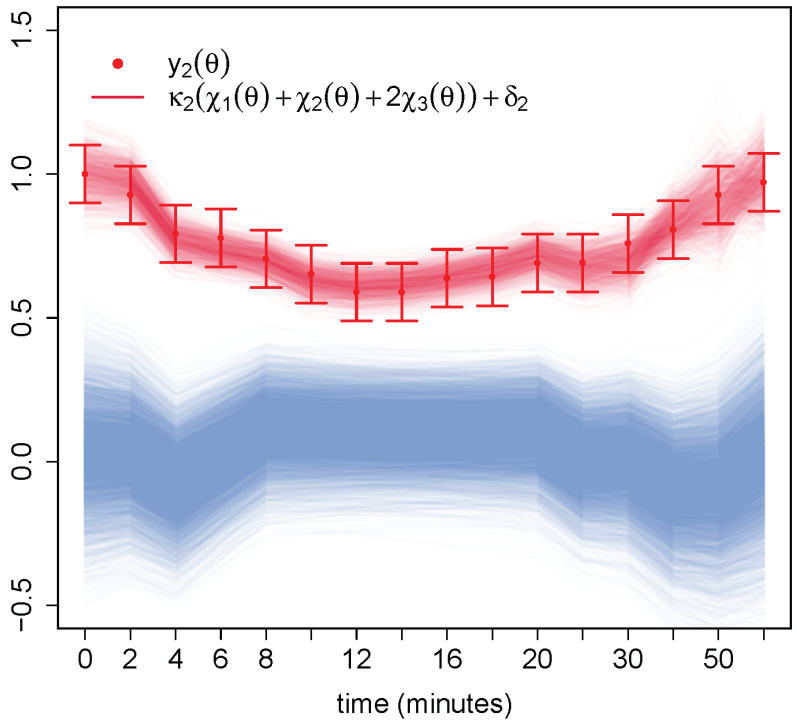


JAK-STAT Intracellular Signaling Pathway Model

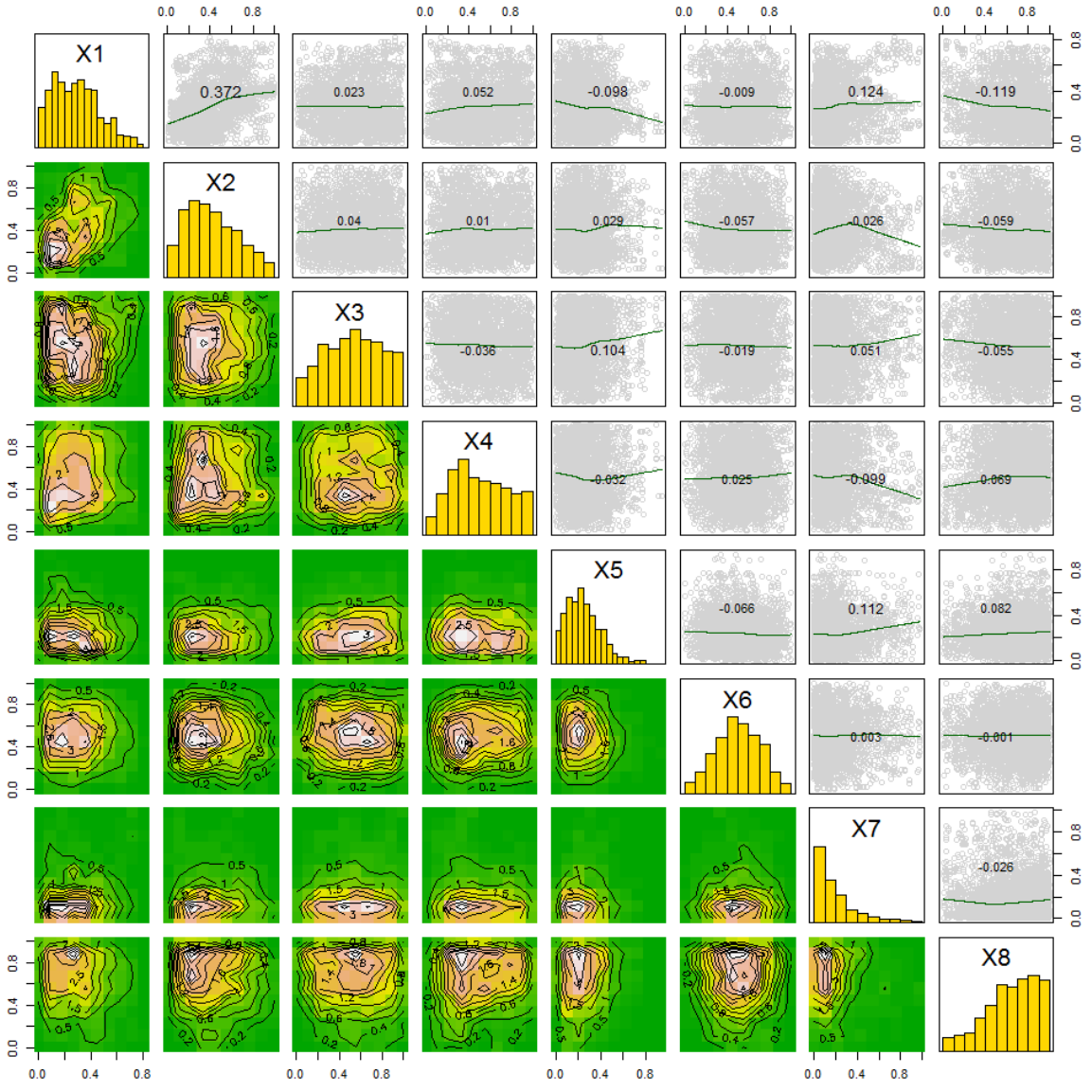
- In the JAK-STAT exampmle, we construct a design of $m = 100$ settings of $\boldsymbol{\theta}$ and at each setting, we have an ensemble of 10 solution realizations from the "PODES" solver of Chkrebtii et al (2014).
- It is known that the multiplicative scaling (κ's) are an important discrepancy in this model, while it is assumed no additive discrepancy (δ's) is present.
- Solving this system is very expensive! The PODES method took ~ 1 day to compute $20,000$ posterior realizations.

JAK-STAT Intracellular Signaling Pathway Model





JAK-STAT Intracellular Signaling Pathway Model



4. Future work, etc.

Come back in the future.

Conclusion

- In calibration experiments, want predictions of real world process by leveraging a calibrated mathematical-physical model of reality.
- Need to quantify uncertainties in order to perform inference.
- In many cases, the simulator itself is stochastic and has uncertainties. We have outlined a Bayesian approach that allows us to calibrate such models while capturing these additional uncertainties.
- In complex stochastic models, the proposed statistical approach enables more efficient use of computational resources.