Optimal Parameter Upscaling for Partial Differential Equation Models in Mathematical Biology

Chinedu Eleh
Dr. Hans-Werner van Wyk

Department of Mathematics and Statistics, Auburn University

Joint Mathematical Meetings

April 2022

Outline of Presentation

Motivation I

Let

$$u = \mathcal{G}(x; \theta)$$

and consider the problem of finding θ, an input to a mathematical model, given u an observation of solution to the model at point x.

Motivation II

- Consider the most basic coupled epidemic model

$$\begin{aligned}
\frac{ds(t)}{dt} &= -\alpha\, s(t)\, i(t), & s(0) &= s_0 \\
\frac{di(t)}{dt} &= -\alpha\, s(t)\, i(t) - \beta i(t), & i(0) &= i_0 \\
\frac{dr(t)}{dt} &= -\beta\, i(t), & r(0) &= r_0
\end{aligned}$$

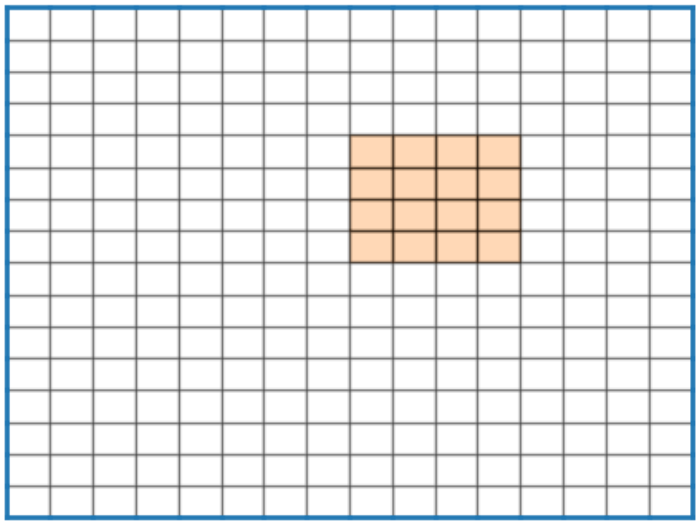
where
$s =$ susceptible population density,
$i =$ infected population density,
$r =$ recovered population density,
$\theta = (\alpha, \beta)$ is a random field that may depend on t, i, r and s.

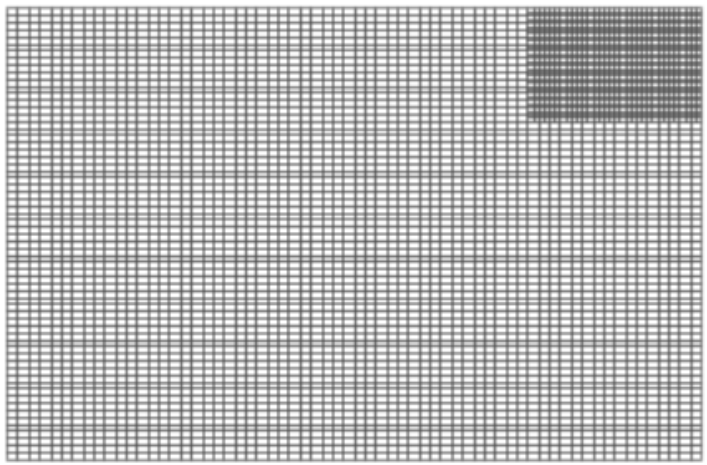
Remarks

- We are interested in estimating the model parameters given some realizations $\{u_i\}_{i=1}^{m}$.
- There are uncertainties in the governing equation due to incomplete knowledge of the underlying physics and/or inevitable errors in measurements.
- These uncertainties are encapsulated in the model parameters $\alpha,\ \beta$ and initial and boundary conditions, for some problems

Brief History

Finite element requires high resolution to capture stochastic information

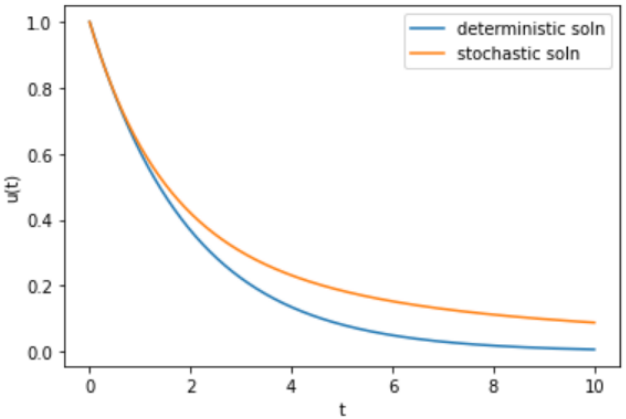




Naive Approach

- First, we explore the forward model.
- A naive approach is to convert the problem into a deterministic one by replacing θ with a statistic and solving analytically or numerically using finite difference, finite element or finite volume method. Simple example:

$$\frac{du(t)}{dt} = -\theta u(t); \qquad u(0) = u_0$$



Naive solution:
$u(t) = u_0 e^{-\bar{\theta} t}$

Stochastic solution:
$u(t) = u_0 \int_\Omega e^{-\theta t} d\mu(\theta)$

- Perturbation and homogenization methods have also been considered.

Monte Carlo Sampling (MCS) Methods

- The discovery of MCS algorithms was a huge breakthrough in better explaining the effects of uncertainties in models.
- Metropolis-Hastings update step for the inverse case:

$$\theta^{(k+1)} = \begin{cases} \theta^* & \text{with probability} \quad \min\left\{1, \ \frac{\mathcal{L}(u|\theta^*)\pi(\theta^*)}{\mathcal{L}(u|\theta^k)\pi(\theta^k)}\right\} \\ \theta^k & \text{otherwise} \end{cases}$$

  assuming a symmetric proposal.
- MCMC sampling has a very slow convergence rate, it is mostly not suitable for practical purposes.
- MCMC samples are often correlated

Karhunen-Loeve Expansions I

- A very successful approach also considered stem from the Kosambi–Karhunen–Loève theorem
- For a random field $\theta(t)$, let $\mu_\theta(t)$ be the mean of the process and let $C(t,s) = \text{cov}(\theta_t, \theta_s)$ be its covariance function. The Karhunen–Loève expansion of θ_t is

$$\theta_t(\omega) = \mu_\theta(t) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\phi_i(t)\theta_i(\omega)$$

where ϕ_i's are the orthogonal eigenfunctions and λ_i's are the corresponding eigenvalues of the eigenvalue problem

$$\int_T C(t,s)\phi_i(s)ds = \lambda_i\phi_i(t), \qquad t \in T$$

Karhunen-Loeve Expansions II

and

$$\theta_i(\omega) = \frac{1}{\sqrt{\lambda}} \int_T (\theta_t(\omega) - \mu_\theta(t))\phi_i(t)dt$$

are mutually uncorrelated random variables with zero mean and unit variance.

- KKL modes are the principal components of the covariance kernel and are expensive to compute
- KKL requires prior knowledge of the covariance kernel $C(t, s)$ and the underlying distribution.

Pause and Ponder

- Naive approach does not carry forward any stochastic information
- Monte Carlo Simulation is very slow
- Karhunen-Loeve expansions require prior knowledge of the underlying distribution, which is not known for many problems.

What then is a way forward?

Polynomial Chaos

- Generalized polynomial chaos, in many cases, is arguably the only feasible method for stochastic simulations of complex systems [Xiu, 2010] [1]

Theorem (Cameron & Martin, 1947)

Let $L^2(\Omega, \mathcal{A}, P)$ *be a Hilbert space of real-valued random variables and* $\mathfrak{D} \subset L^2(\Omega, \mathcal{A}, P)$*, a complete subspace. Suppose for any* $n \in \mathbb{N}$*,* $\mathcal{B}_n(\mathfrak{D}) = \{f(\xi_1, \cdots, \xi_m) : f$ *an* m*-variate polynomial of degree* $\leq m$ *,* $\xi_i \in \mathfrak{D},\ i \in [m],\ m \in \mathbb{N}\}$ *and* $\{\overline{\mathcal{B}}_n\}_{n\in\mathbb{N}} \subset L^2(\Omega, \mathcal{A}, P)$ *is a strictly increasing complete subspace and Gaussian for* $n = 1$ *. Then, there exists* $\{\mathfrak{D}_n\}_{n\in\mathbb{N}}$ *such that*

$$\bigoplus_{n=0}^{\infty} \mathfrak{D}_n = L^2(\Omega, \sigma(\mathfrak{D}), P)$$

Specifically, for $\sigma(\mathfrak{D}) = \mathcal{A}$*,* $\bigoplus_{n=0}^{\infty} \mathfrak{D}_n = L^2(\Omega, \mathcal{A}, P)$

Generalized Polynomial Chaos

- The work of [Xiu & Karniadakis, 2002] is pivotal in understanding, applying and generalizing polynomial chaos method. They coined the term *generalized polynomial chaos expansion*
- Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence of (not necessarily identically distributed) basic random variables satisfying conditions
  1. $\mathbb{E}[|\xi_n|^m] < \infty$ for all $n, m \in \mathbb{N}$
  2. $F_{\xi_n}(x) := P(\xi_n \leq x)$ is continuous for each $n \in \mathbb{N}$

Definition

A distribution is said to be *determinate*, in the Hamburger sense (aka *solvability of the moment problem* if the distribution function is *uniquely* defined by the sequence of its moments

$$\mathbb{E}[\xi^m] = \int_{\mathbb{R}} x^m dF_\xi(x)$$

Generalized Polynomial Chaos

With conditions (1) and (2) and the preceding definition, it follows that:

Theorem

The sequence of orthogonal polynomials associated with a real random variable ξ satisfying the two conditions above is dense in $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), dF_\xi)$ if and only if the moment problem is uniquely solvable for its distribution.

Following this development and other equivalent formulations, [Ernst etal, 2012] proved conditions under which a random vector admits generalized polynomial chaos expansion

Generalized Polynomial Chaos

Theorem

If the distribution function F_ξ of a random vector $\xi = (\xi_1, \cdots, \xi_n)$ with continuous distribution and finite moments of all orders satisfies one of the following conditions, then the multivariate polynomials in $\xi_1, \cdots, \xi_n$ are dense in $L^2(\Omega, \sigma(\xi), P)$. In this case any random variable $\zeta \in L^2(\Omega, \sigma(\xi), P)$ is the limit of its generalized polynomial chaos expansion, which converges in quadratic mean.

(a) *The distributiion function F_ξ has compact support, i.e., there exists a compact set $S \subset \mathbb{R}^n$ such that $P(\xi \in S) = 1$*

(b) *The random vector is exponentially integrable, i.e., there exists $c > 0$ such that*

$$\mathbb{E}(e^{c\|\xi\|}) < \infty$$

Wiener-Askey Scheme

Courtesy: [Ernst etal, 2010]

Distribution	polynomials	density
Gaussian	Hermite	$\rho(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}$
Gamma(α, λ)	Laguerre	$\rho(\xi) = \frac{\lambda}{\xi(\alpha)} (\lambda\xi)^{\alpha-1} e^{-\lambda\xi}$
Beta(α, β)	Jacobi	$\rho(\xi) = \frac{(1-\xi)^{\alpha}(1+\xi)^{\beta}}{2^{\alpha+\beta+1} B(\alpha+1,\beta+1)}$
Uniform(α, β)	Legendre	$\rho(\xi) = \frac{1}{\beta-\alpha}$
Arcsin	Chebyshev	$\rho(\xi) = \frac{1}{\sqrt{1-\xi^2}}$

Transport Maps

- Polynomial chaos expansion method is not without faults
- Its computational complexity increases fast with increase in the number of parameters θ [Moselhy & Marzouk, 2012]
- Transport maps were introduced to mitigate the computational burden involved modeling with polynomial chaos
- A transport map, $T : \mathbb{R}^n \to \mathbb{R}^n$ is a deterministic transformation that pushes forward μ to ν, yielding [Parno & Marzouk, 2012]

$$\nu(B) = \mu(T^{-1}(B))$$

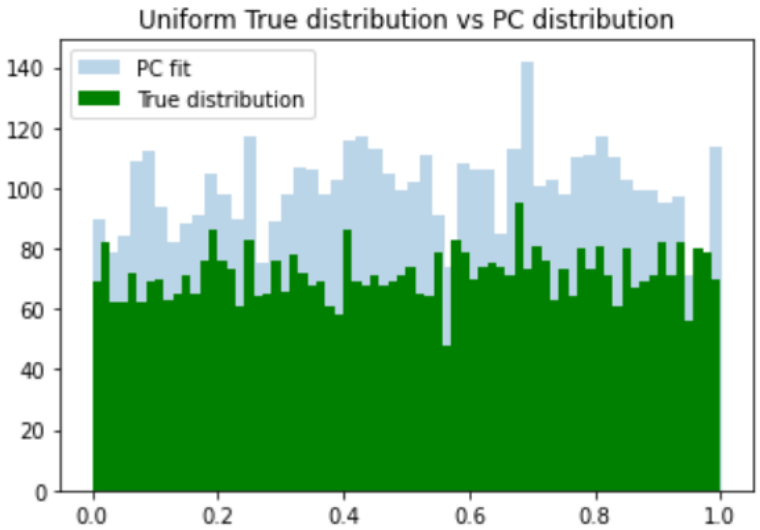
- Uniqueness of T is ensured by assuming the *triangular formulation*

$$T(x_1, x_2, \ldots, x_d) = \begin{bmatrix} T_1(x_1) \\ T_2(x_1, x_2) \\ \vdots \\ T_d(x_1, x_2, \ldots, x_d) \end{bmatrix}$$

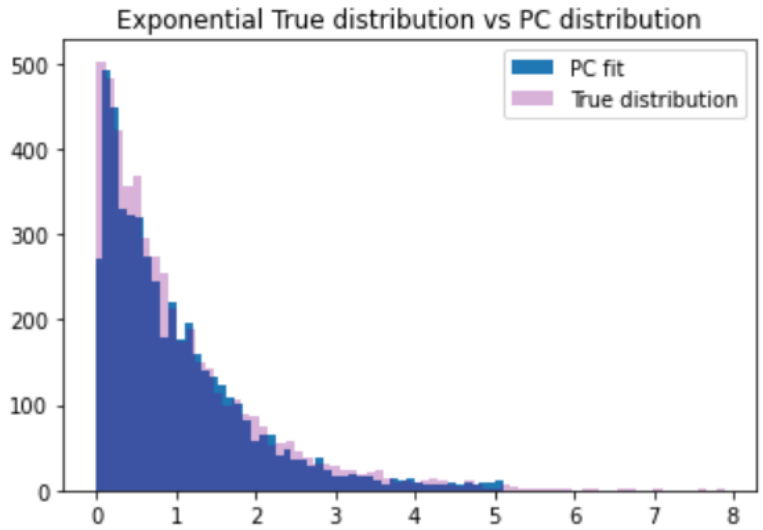
Numerical Experiments

First, we consider the case where θ is a scalar random variable

$\theta \sim Uniform(0, 1)$



$\theta \sim exp(0.5)$

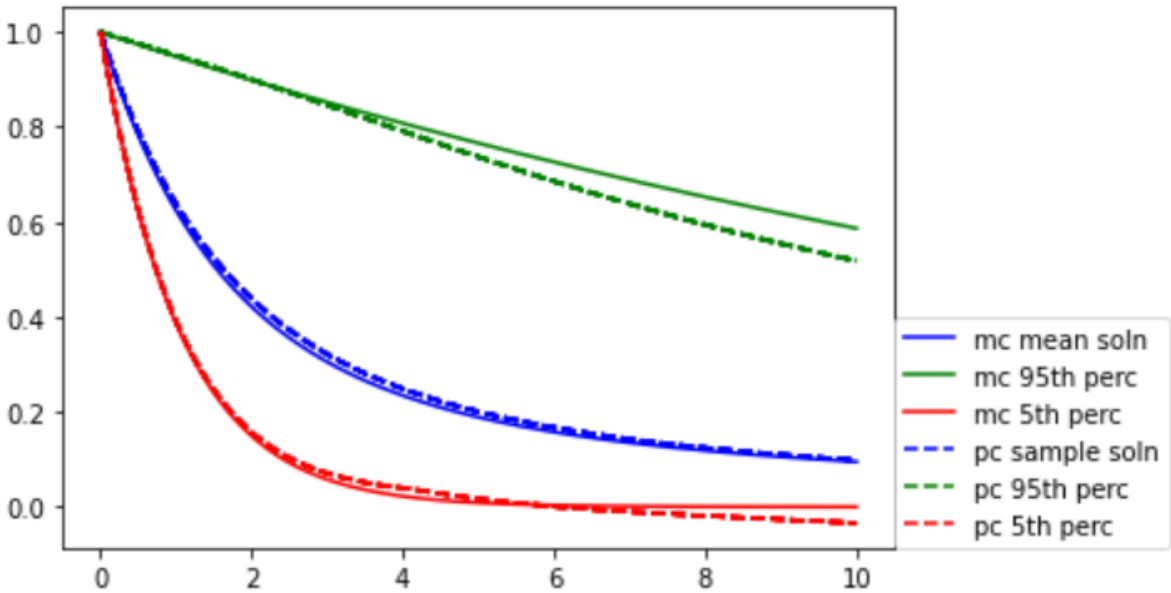


Numerical Experiments

Consider the exponential decay model given by the differential equation

$$\frac{du(t;\theta)}{dt} = -\theta u(t,\theta); \qquad u_0 = c$$

where $\theta > 0$ is a random variable.



Conclusion

- Most, if not all biology inspired mathematical models depend on certain random parameter(s)
- Successes in making inference from or validating these models depend on how well the stochastic information from these parameters are propagated into the state variables
- We demonstrated that transport maps are powerful and handy in this regard
- In progress, we are looking to leverage the expressive power of Deep Neural Networks in constructing transport maps

[1] Dongbin Xiu. *Numerical methods for stochastic computations*. Princeton university press, 2010.