Active Learning for Multi-Fidelity Computer Experiments with Different Mesh Densities

Romain Boutelet and Chih-Li Sung*

Department of Statistics and Probability
Michigan State University

2024 INFORMS Annual Meeting

*This author gratefully acknowledges the support provided by NSF DMS 2113407

Table of Contents

Outline

Introduction

- Often, real world phenomena are described with a set of Partial Differential Equations (PDEs) which are impossible to solve exactly in practice.
- Finite Element Methods (FEM) offer an approximation of the solution through the use of a **tuning parameter** (e.g. mesh size) to control the numerical accuracy as well as the computational cost.
- We propose a class of **adaptive** non-stationary Gaussian process models to link the outputs of simulation runs with different mesh densities, as well as an **Active Learning** method to improve the performance of this model.

Setting

- Two types of inputs:
  - **Input variables**, which can be control variables, environmental variables, or calibration variables.
  - **Tuning variables**, which determine the performance of the numerical computations.
- Computer simulations are run at different **input** locations and with different **tuning parameter** values.
- The **output values** must be a real number.

Outline

Multi-Fidelity Emulators

Objective:

Use information from **different fidelity levels** to make accurate predictions on the underlying model.

Existing methods:

- Auto-regressive model (O'Hara and Kennedy, 2000)
- Non-stationnary model (Tuo et al., 2014)

Non-Stationary Model (Tuo et al., 2014)

Non-Stationary Model

The **response variable** y, at input location $\mathbf{x} \in \mathcal{D}$ and with mesh size $t \in \mathcal{T}$, is assumed to be the sum of the exact solution φ, and the error δ, such that,

$$y(\mathbf{x}, t) = \varphi(\mathbf{x}) + \delta(\mathbf{x}, t),$$

where $\varphi(\mathbf{x})$ and $\delta(\mathbf{x}, t)$ are realizations of two mutually independent GPs V and Z, respectively.

<u>Remark:</u> Since y must equal the exact solution φ as $t \to 0$, we need δ to satisfy $\delta(\mathbf{x}, t) \xrightarrow[t \to 0]{} 0$, for all $\mathbf{x}$.

More in detail

- The **mean function** is assumed to have a separable form, such that

$$\mathbb{E}[V(\mathbf{x})] = f_1^T(\mathbf{x})\beta_1, \quad \mathbb{E}[Z(\mathbf{x},t)] = f_2^T(\mathbf{x},t)\beta_2$$

- The **covariance function** of our response variable is

$$k(\mathbf{x},\mathbf{x}',t_1,t_2) = k_1(\mathbf{x},\mathbf{x}') + k_2(\mathbf{x},\mathbf{x}',t_1,t_2),$$

where

  - V has a stationary covariance function of the form
    $k_1(\mathbf{x},\mathbf{x}') = \sigma_1^2 \prod_{i=1}^d e^{-\phi_1^2(x_i - x_i')^2}$
  - Z has a non-stationary covariance function of the form
    $k_2(\mathbf{x},\mathbf{x}',t_1,t_2) = \sigma_2^2 F_H(t_1,t_2) \prod_{i=1}^d e^{-\phi_2^2(x_i - x_i')^2}$

Extending the Non-Stationary model

We introduce a new covariance function on the mesh size t, adapted from the Fractional Brownian Motion.

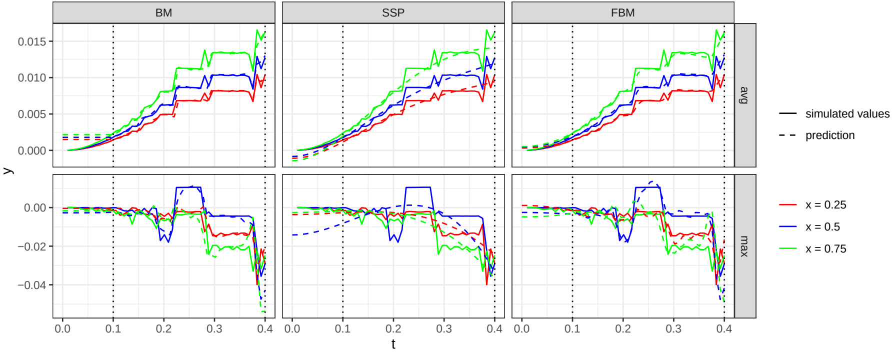
$$F_H(t_1, t_2) = \left\{ \frac{1}{2}(t_1^{2H} + t_2^{2H} + |t_1 - t_2|^{2H}) \right\}^{\frac{l}{2H}}, \quad 0.5 \leq H \leq 1$$

The parameter H can be estimated with MLE.

Remark: This covariance function replaces the two covariance functions proposed in Tuo, Wu and Yu (2014):

- The Brownian Motion, $BM(t_1, t_2) = \min(t_1, t_2)^l$.
- The Scaled Stationary Process, $SSP(t_1, t_2) = \left(\sqrt{t_1, t_2} e^{-\phi(t_1 - t_2)^2}\right)^l$

To illustrate the behavior of FEM output we show simulated values from a simple PDE approximated with FEM for two different responses (*average* or *maximum*), and three different input values.



Outline

Active Learning and Multi Fidelity Emulators

- **Active Learning (AL) designs**
  - optimize the choice of the next design point by a given **criterion**,
  - must take the **computational cost** into account for multi-fidelity simulations.
- **Current methods**
  - He et al. (2017) used the EQIE criterion for optimization,
  - Stroh et al. (2022) gave a general framework to perform sequential design using the non-stationary model.

Integrated Mean Square Prediction Error (IMSPE)

For GPs, the IMSPE has a **closed-form expression** (Binois et al. 2019).

This result can be extended to the non-stationary model.

Lemma 1

The IMSPE from the n initial design elements, I_n, can be written as

$$I_n := \int_{\mathbf{x} \in \mathcal{D}} \sigma_n^2(\mathbf{x}, 0) d\mathbf{x}$$

Calculating the IMSPE sequentially

Our goal is to compute the IMSPE where the new candidate $(\tilde{\mathbf{x}}, \tilde{t})$ is added to the design, $I_{n+1}(\tilde{\mathbf{x}}, \tilde{t})$.

Proposition 1 (Sequential IMSPE)

Let $(\tilde{\mathbf{x}}, \tilde{t})$ be a candidate design point.
Then the IMSPE can be written as

$$I_{n+1}(\tilde{\mathbf{x}}, \tilde{t}) = I_n - R_n(\tilde{\mathbf{x}}, \tilde{t}),$$

where $R_n(\tilde{\mathbf{x}}, \tilde{t})$ only requires $\mathcal{O}(n^2)$ computation.

IMSPER to Cost Ratio

To take the **computational cost** into account for our criterion, we choose the next point $(\mathbf{x}_{n+1}, t_{n+1})$ by maximizing the ratio between the IMSPER $R_n(\tilde{\mathbf{x}}, \tilde{t})$ and the cost $C(\tilde{t})$:

$$(\mathbf{x}_{n+1}, t_{n+1}) = \underset{(\mathbf{x},t)\in\mathcal{D}\times\mathcal{T}}{\operatorname{argmax}} \frac{R_n(\mathbf{x}, t)}{C(t)}$$

Outline

Synthetic example

We use a synthetic example, simulated using our FBM model, with $H = 1$.

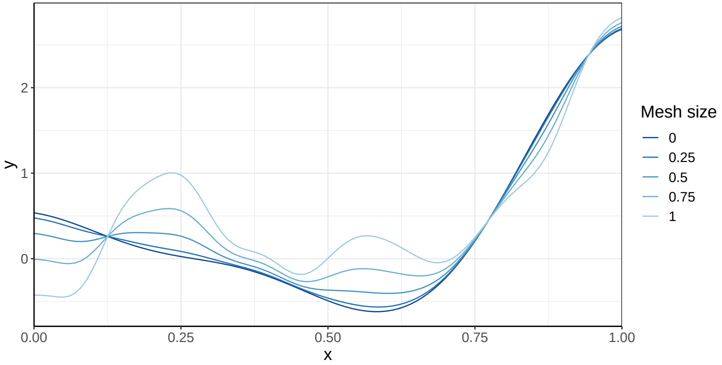
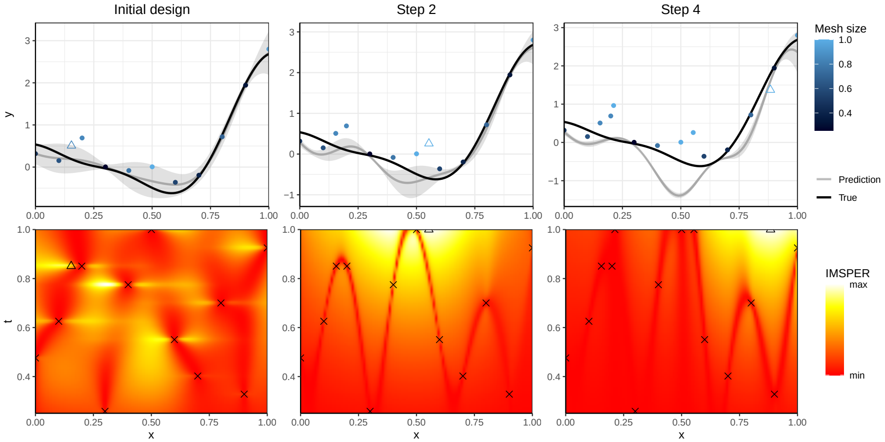


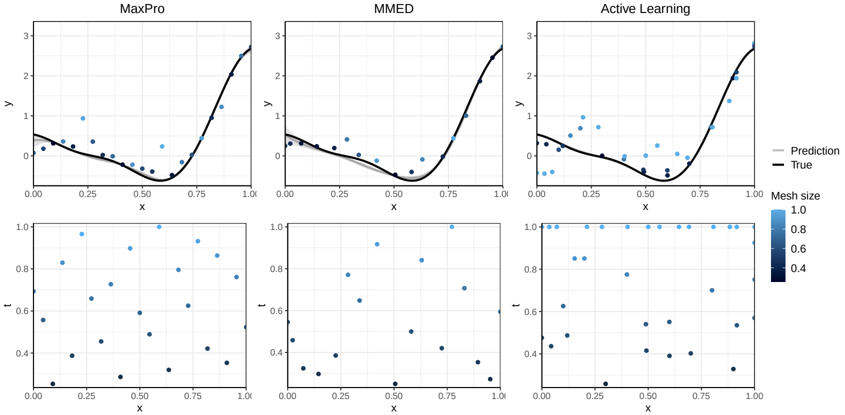
Illustration of the Active Learning



Issue: This method leverages the reduction of the prediction error, but relies on the initial design for robust parameter estimation.

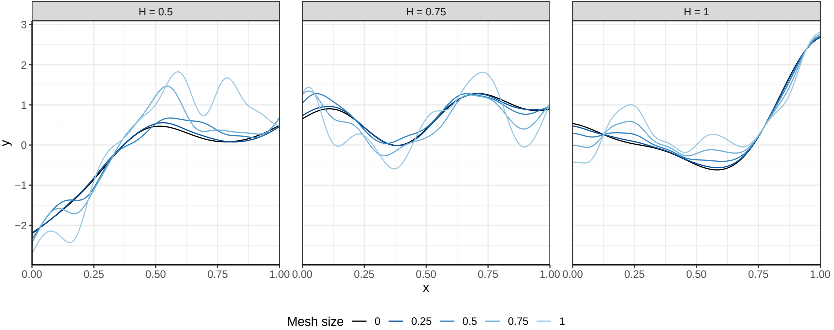
Comparing One-shot Design vs Active Learning

- The one-shot designs are the Multi-Mesh Experimental Design (Yuchi et al. 2023), and MaxPro (Joseph, Gul, and Ba, 2015).
- The Active Learning design opts for more cheap design points, for a better coverage on the input space.



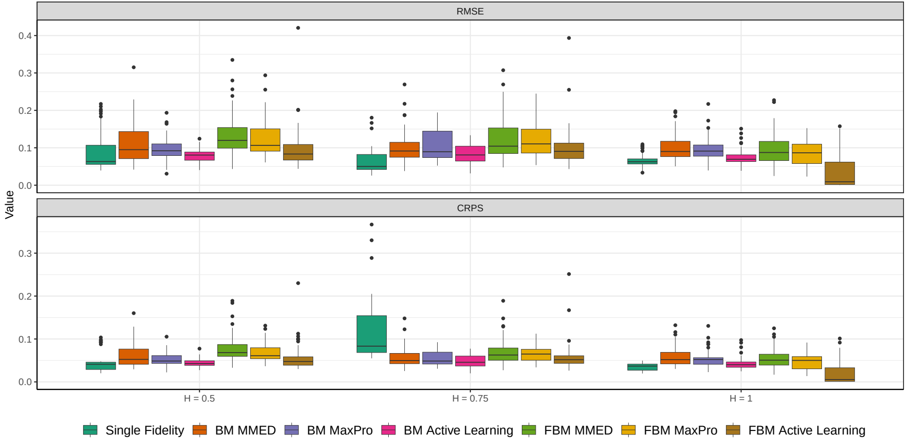
Numerical simulations using our model

Using our FBM surrogate model, we produce a set of 5 sample functions for 3 different values of H.



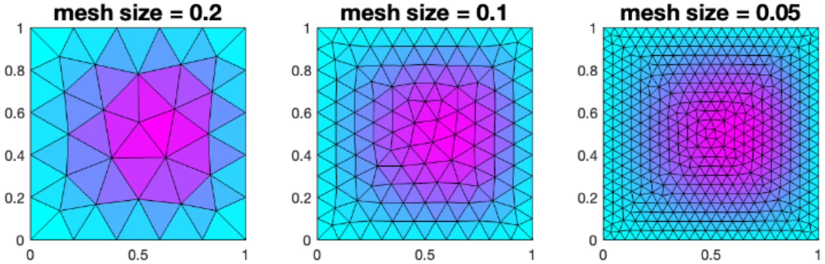
Results from numerical simulations

We compare the prediction results from our model with the BM model, using the 3 different designs in each case.



Poisson's equation on a square membrane:

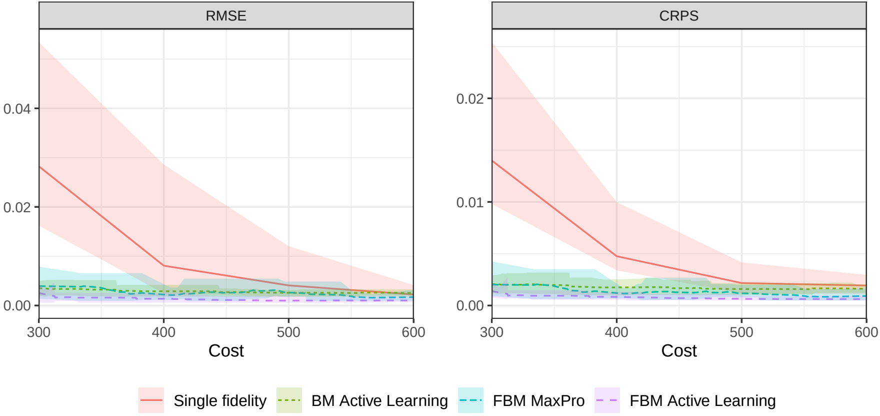
$$\Delta u = (x^2 + 2\pi^2)e^{xz_1}\sin(\pi z_1)\sin(\pi z_2) + 2x\pi e^{xz_1}\cos(\pi z_1)\sin(\pi z_2)$$
$$u = 0 \text{ on } \partial\mathcal{D}$$



The response of interest is $\int_{(z_1,z_2)\in\mathcal{D}} u(z_1, z_2; x)dz_1dz_2$, for $x \in [-1, 1]$.

Results of Elliptic PDE

We now compare the results from our method to the Single Fidelity, MaxPro and BM model with Active Learning.



Conclusion

What we did:

- We introduced a novel and adaptive covariance function for the non-stationary model.
- We proposed an **active learning** strategy adapted to the Multi-Fidelity setting.
- Numerical studies show the **effectiveness** of our method compared other methods.

What we need to explore:

- Developing an **initial design** that gives more robust parameter estimation.

Thank You!