Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

A MULTI-FIDELITY APPROACH TO INTEGER-ORDERED SIMULATION OPTIMIZATION USING GAUSSIAN MARKOV RANDOM FIELDS

Graham Burgess[1], Luke Rhodes-Leader[2], Rob Shone[2], and Dashi Singham[3]

[1]STOR-i Centre for Doctoral Training, Lancaster University, Lancaster, UK
[2]Dept. of Management Science, Lancaster University, Lancaster, UK
[3]Operations Research Dept., Naval Postgraduate School, CA, USA

ABSTRACT

We are interested in discrete simulation optimization problems with a large solution space where decision variables are integer-ordered and models of both low and high fidelity are available to evaluate the objective function. We model the error of a low-fidelity deterministic model with respect to a high-fidelity stochastic simulation model using a Gaussian Markov random field. In a Bayesian optimization framework, using the low-fidelity model and the model of its error, we reduce the reliance on the expensive high-fidelity model.

1 INTRODUCTION

Simulation optimization problems arise when one cannot directly evaluate an objective function but can only estimate it via stochastic simulation. When the decision variables take discrete values, this is known as discrete simulation optimization. Furthermore, discrete decision variables may be integer-ordered, such as the number of beds in a hospital ward. In this case, one can take advantage of spatial relationships between feasible solutions to understand the behavior of the objective function across the solution space while only simulating a fraction of solutions.

Salemi et al. (2019) exploit spatial relationships by modelling the objective function across an integer-ordered solution space using a Gaussian Markov random field (GMRF). A GMRF $\mathbb{Y}$ is a multivariate normal distribution $\mathscr{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ associated with an undirected labeled graph $\mathscr{G}=\{\mathscr{V},\mathscr{E}\}$. Nodes $\mathscr{V}$ correspond to feasible solutions and edges $\mathscr{E}$ define the neighborhood structure. The Markov property asserts that any given node, conditional on its neighborhood, is independent of all other nodes. Conditional correlations between neighbors, given all other nodes, are captured in the precision matrix $\boldsymbol{Q}=\boldsymbol{\Sigma}^{-1}$. Salemi et al. (2019) use a GMRF with a suitable neighborhood structure in a Bayesian optimization framework.

The simulation model used in simulation optimization is typically considered to be of high fidelity. A high-fidelity model is good at emulating the input-output behavior of the real system in question, but can be expensive to run. Low-fidelity models, such as simplified simulation models or analytical models, are typically biased, but cheaper to run. A low-fidelity model can help guide the optimization of a high-fidelity model. In continuous simulation optimization, Huang et al. (2006) model the error (or bias) of low-fidelity models with respect to higher-fidelity models using Gaussian process models. We use a GMRF to model low-fidelity error in integer-ordered simulation optimization.

2 METHODOLOGY

The basic ideas behind our method for integer-ordered simulation optimization are to model low-fidelity error using a GMRF and to incorporate both low-fidelity (LF) and high-fidelity (HF) evaluations into a conditional distribution for the HF surface. We iteratively update this distribution to drive the search.

Formally, we wish to solve $\min_{\boldsymbol{x}\in\mathscr{X}} y(\boldsymbol{x})=\mathbb{E}[Y(\boldsymbol{x})]$ where the solution space $\mathscr{X}$ is a subset of the $d-$dimensional integer lattice $\mathbb{Z}^d$ and $Y(\cdot)$ denotes some (random) quantity of interest. Suppose we can estimate $y(\boldsymbol{x})$ with an unbiased high-fidelity (HF) stochastic simulation model or a biased low-fidelity (LF) deterministic model. We model the HF response surface with a GMRF $\mathbb{Y}$. We model the LF error (or bias) with respect to the HF model with another GMRF $\mathbb{Y}_E$.

We evaluate a subset of solutions with both the LF and HF models. This gives us noisy observations of the HF surface, $\boldsymbol{y}_H$, and noisy observations of the LF error, $\boldsymbol{y}_E$. We use these to estimate parameters of the two GMRFs. We collect LF observations $\boldsymbol{y}_L$ at an additional subset of solutions. We derive a conditional distribution for the HF surface given multi-fidelity observations, $\mathbb{Y}|\boldsymbol{y}_H, \boldsymbol{y}_L$. The conditional distribution of Salemi et al. (2019) given single-fidelity observations is partitioned into a block for unsimulated solutions and a block for simulated solutions. Our conditional distribution also contains a block for solutions evaluated with the LF model. This block incorporates information from the LF model and the model of its error.

We use this conditional distribution to guide the search in a multi-fidelity version of the Gaussian Markov improvement algorithm (GMIA) (Salemi et al. 2019). In each iteration, we update our conditional distribution and use it to compute the complete expected improvement (CEI) of all solutions with respect to the sample best solution according to the HF model. We then use information from the LF model to decide whether or not to evaluate the solution with the best CEI using the HF model.

3 ILLUSTRATIVE EXAMPLE

We illustrate the workings of our multi-fidelity approach using a two-dimensional capacity planning problem for homelessness. Our decision variables are the number of housing and shelter units to have in the first year of a five-year modelling horizon. We use a fluid flow queueing model of unsheltered people being 'served' in housing (Burgess et al. 2025). Part of the queue for housing is sheltered, the remainder being unsheltered. The objective function is quadratic in the number of unsheltered people and number of shelters used over a five-year horizon. Our HF stochastic model is this fluid model plus mean-zero Gaussian noise. Our LF deterministic model is this fluid model without noise but with an altered service rate for each house. On the left in Figure 1, we show the true HF surface, without noise. In the middle, we show the conditional mean of the HF surface given multi-fidelity observations. On the right, we show the resulting CEI values and the proposed sampling step.

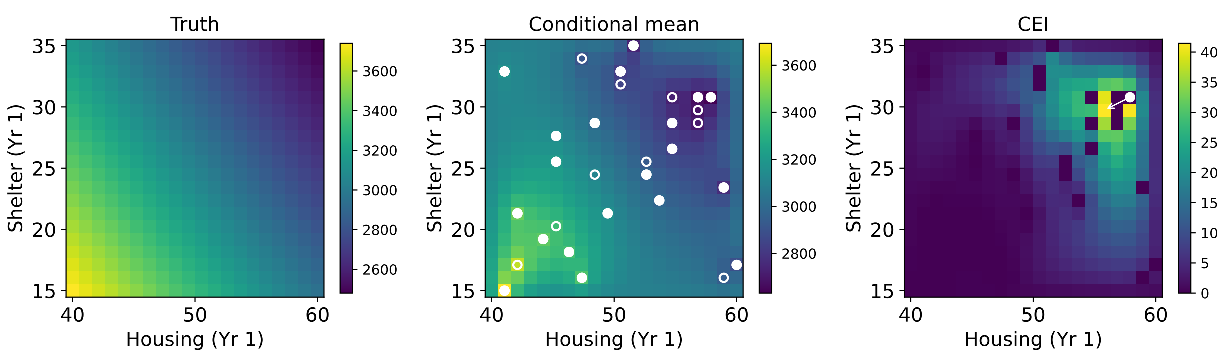

Figure 1: **Left**: True HF surface without noise. **Middle**: Conditional mean of $\mathbb{Y}$ given HF evaluations $\boldsymbol{y}_H$ (filled circles) and LF evaluations $\boldsymbol{y}_L$ (unfilled circles). **Right**: $\text{CEI}(\boldsymbol{x})\ \forall\, \boldsymbol{x} \in \mathscr{X}$. Filled circle is sample best solution. Arrow points towards solution with best CEI.

REFERENCES

Burgess, G., D. I. Singham, and L. Rhodes-Leader. 2025. "Time-Varying Capacity Planning for Designing Large-Scale Homeless Care Systems". *Under Review*.

Huang, D., T. T. Allen, W. I. Notz, and R. A. Miller. 2006. "Sequential Kriging Optimization using Multiple-Fidelity Evaluations". *Structural and Multidisciplinary Optimization* 32:369–382.

Salemi, P. L., E. Song, B. L. Nelson, and J. Staum. 2019. "Gaussian Markov Random Fields for Discrete Optimization via Simulation: Framework and Algorithms". *Operations Research* 67(1):250–266.