Discrete Event Simulation and Metamodels

Ibrahim B. Topal , Graham Burgess (Supervisor)

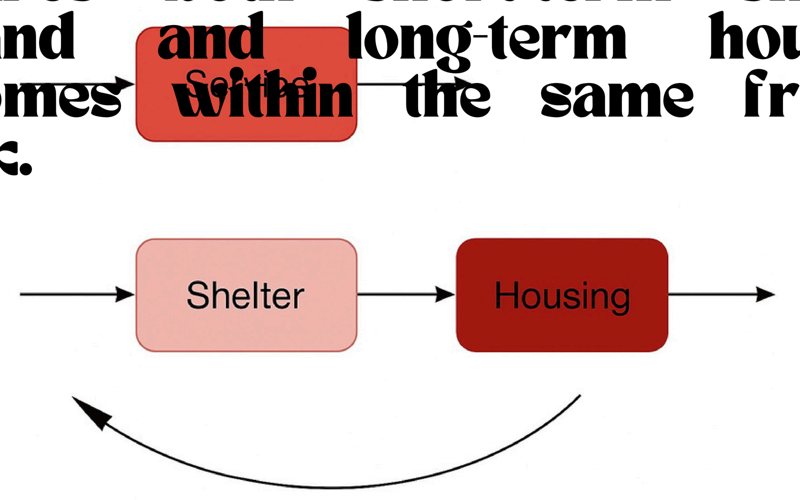
[1]STOR-i CDT, Lancaster University

Introduction

Discrete-Event Simulation (DES) is a powerful tool for modelling complex systems where changes occur at discrete points in time, such as in healthcare, manufacturing, and housing. While DES provides detailed insights, it can be computationally expensive when many replications are required. To address this, researchers use metamodels—simplified statistical or mathematical approximations of the simulation—that provide faster predictions while preserving key system behaviour. To build accurate metamodels, it is essential to cover the model input space efficiently with a limited number of simulation runs. As the number of input dimensions grows, this becomes increasingly difficult; Latin Hypercubes addresses this by ensuring good coverage of the input space with fewer runs.

Queueing Models

Two queueing structures were studied to demonstrate the use of DES and metamodels. A single-server queue was used for hospital waiting times, where individuals remain in a single line until service is received, and the mean waiting time is the key measure. A tandem queue was used for homelessness: the first state represents waiting for shelter, and the second represents waiting for permanent housing after leaving shelter. This layered design captures both short-term shelter demand and long-term housing outcomes within the same framework.



Constructing One-Dimensional Metamodels

To generate the metamodel plots we used different approaches in R.

For the Gaussian Process (GP) model, we used the mlegp package. A GP assumes outputs follow:

$$z \sim N(\mu, \Sigma(\beta) + N)$$

where $\Sigma(\beta)$ is the GP covariance matrix (parameterised by the length-scale β) with marginal variance, and N is a diagonal nugget matrix with ith diagonal element $\sigma^2_\varepsilon(i)$, capturing stochastic simulation noise. For deterministic outputs $N = 0$; for constant noise $N = \sigma^2_\varepsilon I$. mlegp allows N to be specified exactly or up to a multiplicative constant.

For the Linear and Quadratic regression metamodels, we used standard polynomial regression in R:

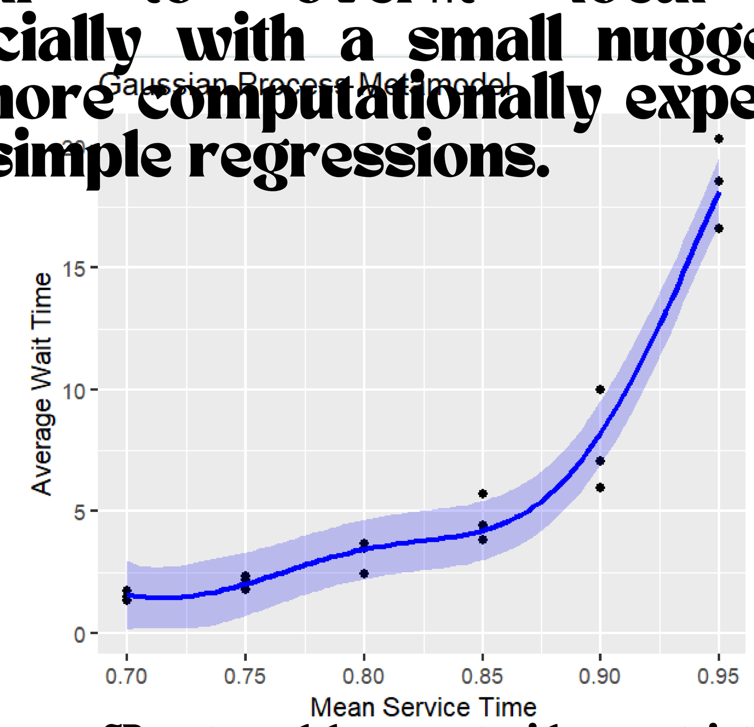
$$y = \beta_0 + \beta_1 x, \qquad y = \beta_0 + \beta_1 x + \beta_2 x^2.$$

These were fitted using lm() with polynomial terms in service time, providing simpler benchmarks against the GP.

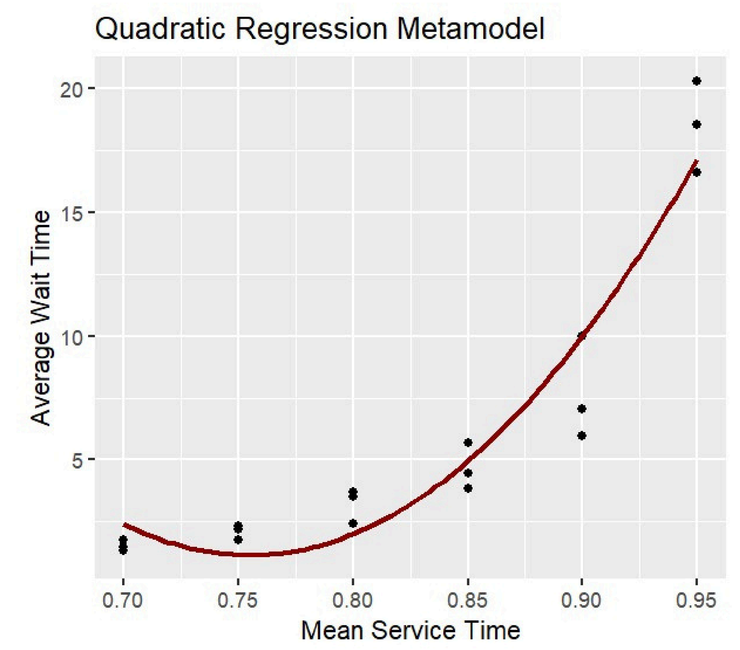
Together, these approaches allow comparison of simple polynomial regressions with the more flexible GP model.

Metamodel Results

We evaluate metamodels for predicting mean hospital waiting time in a single-server queue (M/M/1-type) DES. Linear and Quadratic regressions capture broad trends but struggle near high utilisation. Gaussian Process (GP) metamodels achieved the lowest MSE by adapting to curvature; however, GPs can appear to overfit local noise (especially with a small nugget) and are more computationally expensive than simple regressions.



GP metamodel: mean fit with uncertainty band.



Quadratic metamodel: smooth trend but underfits extremes.

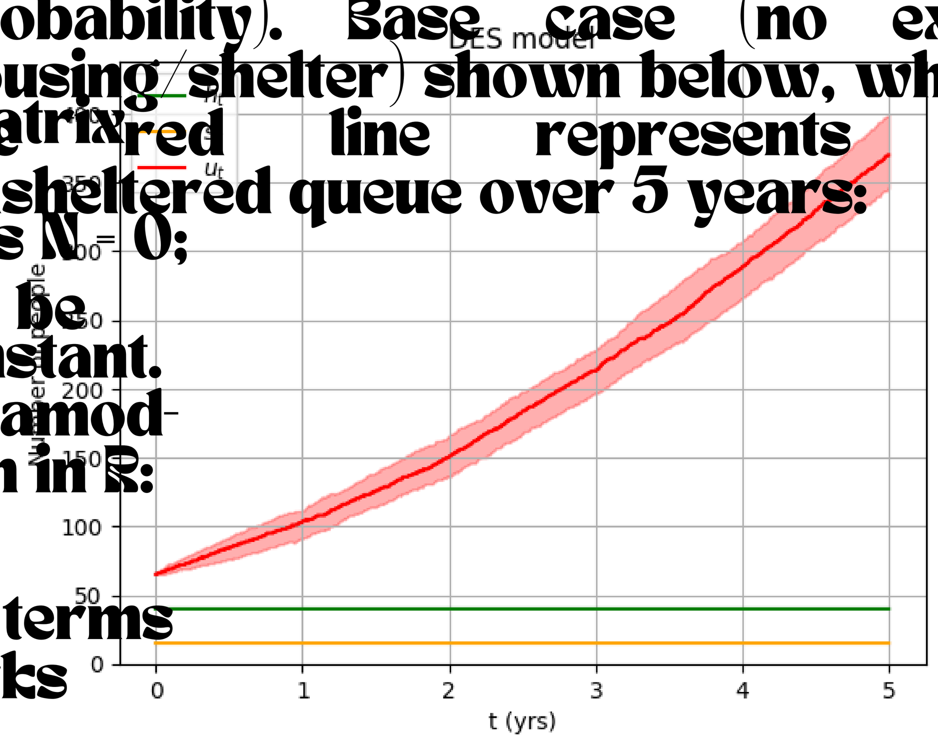
MSE Comparison (lower is better)

Metamodel	MSE
Gaussian Process	1.1328
Quadratic Regression	2.7552
Linear Regression	9.6135

MSE computed on DES outputs: $\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$.

Homelessness DES

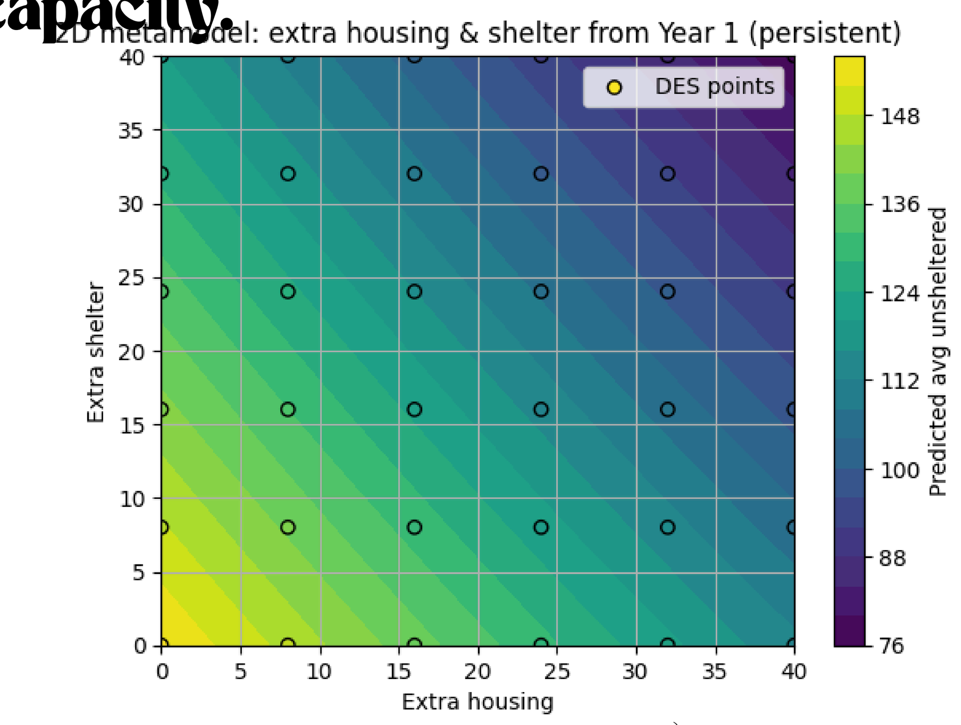
We simulate a tandem queue with a potential loop from Housing back to Shelter. Key inputs are: H (extra housing), S (extra shelter), Mh (mean housing service time), Ms (mean shelter service time), A (arrival rate), and R (re-entry probability). Base case (no extra housing/shelter) shown below, where the red line represents the unsheltered queue over 5 years:



By varying these inputs we assess impacts on unsheltered queueing times; we then build metamodels to approximate these DES outputs efficiently.

Higher Dimensional Metamodels

2D(H, S) : For two dimensions, we tested the impact of adding extra housing (H) versus shelter (S) capacity.



2D quadratic surface for (H, S)

The fitted quadratic response surface is

$$\hat{y} = 155.8 - 1.17\,H - 0.93\,S + 0.0025\,H^2 + 0.001\,S^2 + 0.000200\,HS.$$

with MSE $= 1.485$ and $R^2 = 0.996$. Adding one unit of housing reduces the amount of unsheltered homelessness more than adding one unit of shelter (since $|-1.17| > |-0.93|$). Also, the quadratic is stronger in H (coefficient on H2 is 0.0025) than in S (coefficient on S2 is 0.001), i.e., as housing increases further, the quadratic effect is greater.

4D(H, S, A, R) : We then expanded to four inputs by adding the arrival rate (A) and re-entry probability (R). A full-factorial design was used and a quadratic (squares + pairwise interactions) fitted to capture curvature and key interactions.

6D(H, S, Mh, Ms, A, R) : Full factorial designs become computationally expensive in six dimensions. We therefore used Latin Hypercube to achieve good space-filling coverage with fewer runs, then fitted a quadratic with selected interactions. Finally, we applied stepAIC, another package in R, to automatically remove negligible terms (Akaike Information Criterion-based stepwise selection), yielding a concise metamodel with predictive accuracy.

Conclusions

- DES + Metamodels = fast, accurate what-ifs. GP gave the lowest MSE vs. linear/quadratic but is more computationally expensive.
- 2D(H, S) : $|\beta_H| > |\beta_S|$ ⇒ adding housing reduces the unsheltered homelessness slightly more compared to adding shelter.
- Scaling up: Latin Hypercube + stepAIC enabled a concise 6D model with far fewer runs than full factorial.

References

Singham, D. I., Lucky, J., and Reinauer, S. Discrete-event simulation modeling for housing of homeless populations.
Dancik, G. mlegp: An R package for Gaussian process modeling and sensitivity analysis.
Barton, R. R. Tutorial: Basics of Metamodeling. Proceedings of the Winter Simulation Conference, 2023.
Wolf, ... martens, ... A Tutorial on Designing and Conducting Simulation Experiments. Proceedings of the Winter Simulation Conference.
Nelson, B. L., and Pei, L. Foundations and Methods of Stochastic Simulation. Springer.