Generic simulation metamodeling: Making metamodeling accessible for simulation users

M.B. Kolenbrander, L.J. Kortmann, S.W. Cunningham, A. Verbraeck
Faculty of Technology, Policy and Management,
Delft University of Technology
Jaffalaan 5, 2628 BX Delft, the Netherlands

Abstract: Metamodeling can reduce the cost of experimenting with discrete event simulation models. While potentially valuable, metamodeling is not used in commercial settings because of the lack of expertise and effort surrounding the setup. In order to make the power of metamodeling more accessible to simulation users, our goal was to find a metamodeling technique that can be used with any discrete event simulation model. After thoroughly testing both regression metamodeling and Kriging metamodeling techniques, the performance of Kriging, especially with erratic model behaviour, allowed it to be chosen as more suitable to be used in a generic experimental design environment for simulation users. The resulting environment has the potential to reduce the cost of experimenting, increase the knowledge gained from a simulation model and could, with further research, support less costly optimization studies.

Keywords: Metamodeling, Kriging, Regression, Experimental design, Simulation

1. Introduction
Simulation models are used in many fields to experiment with real-world systems to gain insight into their behaviour. Based on this insight one can make strategic, operational, policy, construction or design decisions without the need to experiment with real-world system itself, which could be too costly, inconvenient or simply impossible. Experimenting with simulation models can be time consuming and costly since one experiment can take minutes to even hours to complete (Wang & Shan, 2007). When fully investigating the behaviour of a simulation model takes thousands of experiments, it presents an unworkable situation. In order to save costs, experimenters reduce their scope of research or otherwise conduct less thorough investigations into the behaviour of a simulation model, increasing the risk of overlooking valuable information. A possible solution to this particular problem is metamodeling. Metamodeling is mentioned in literature as a way to reduce the number of required experiments. Only a reduced sample of experiments is run, after which the other results are estimated by certain interpolation techniques.

1.1 Research problem
The problem dealt with in this paper is that, while metamodeling has shown to be a promising technique, it has only been used in single, isolated studies. Metamodels are currently purpose-built and calibrated specifically to handle one model or set of experiments, which is one of the reasons that metamodeling is not used in commercial simulation studies. To set up a metamodel for each project
would be time consuming beyond the point of profitability. Also it requires specific knowledge that not all commercial simulation teams or simulation end-users have in-house. An automated way of setting up and calibrating a metamodel would greatly benefit its general use.

Many metamodeling techniques exist, each with their own advantages and disadvantages (Barton, 1994; Barton, 1998). The choice of the metamodeling technique and its subsequent calibration are normally done based on the goal of the experiments, expected model behaviour and user requirements. In order for laymen to be able to use metamodeling in simulation projects, a metamodeling technique needs to be found that is best suited for general use, i.e. does not require expertise (in metamodeling or even simulation) for its use. It is currently unknown which metamodeling technique is best suited for general use.

1.2 Research question and approach

The research described in this paper is guided by the following question.

*Which metamodeling technique is the most appropriate for generic use in a commercial simulation project?*

The question can be answered by testing multiple metamodeling techniques on criteria that focus on the fact that the metamodeling technique needs to deal with many kinds of simulation models and many kinds of responses, while at the same time it needs to be used by users without knowledge about metamodeling.

The approach that was used to find the most suitable modelling technique is as follows. First a long list of metamodeling techniques is selected based on references in literature. After a qualitative analysis these techniques are scored and reduced to a short list based on a multi-criteria analysis. The short list of metamodeling techniques is analysed quantitatively by comparing their accuracy when applied to artificial results and actual simulation model results. Based on the results of the quantitative analysis the most suitable metamodeling technique can be chosen.

This paper is structured in the following way. First metamodels and metamodeling will be explained more thoroughly in section 2. Afterwards in section 3, a long list of metamodeling techniques will be presented, scored and reduced based on a multi-criteria analysis. In section 4 the remaining metamodeling techniques are compared quantitatively, after which the conclusions are presented in section 5.

2 Metamodels and metamodeling

Metamodeling is essentially the process of using metamodels to reduce the number of experiments, create a metamodel based on experiment results and use the metamodel to estimate results.

2.1 Metamodels

Metamodels are fundamentally a model of a model. In this case, a metamodel is a representation of the input/output function of a simulation model. The simulation model is regarded as a black box that represents a certain relationship between its input and output variables.
There are many different metamodeling techniques that can be used in the field of simulation. Generally, metamodels can be denoted as follows.

\[ g(x) \approx f(x) = y \]

Equation 1 - A general description of a metamodel

Where \( g \) is the simulation model, which is approximated by the metamodel \( f \), producing results \( y \) based on the set of input variables \( x \).

2.2 metamodeling

The process of metamodeling consists of three parts; sampling, fitting and reproducing.

Sampling is the process of selecting a sample (i.e. reduced) set of experiments from the original set of experiments representing all possible combinations of input factors. Sampling is where the cost-saving aspect of metamodeling originates. A sampling method can have a large impact on the accuracy of the resulting metamodel.

Fitting is the process of creating the metamodel. The metamodel is fitted to the results of the sample set of experiments. The method of fitting depends directly on the chosen type of metamodel. This can for instance consist of adjusting the function’s coefficients according to a least squares method.

Once a mathematical representation exists between the input variables and the model’s results, all desired results can be estimated by using the input variables as arguments in the metamodel’s function. This reproduced (i.e. predicted) data can be subsequently presented to users in reports, graphs and charts.

3. Long list of metamodeling techniques

Initially four metamodeling techniques were considered based on references in literature; polynomial regression metamodels (otherwise known as response surface metamodels), spline metamodels, artificial neural networks and lastly Kriging metamodeling. Reducing the list of suitable metamodeling technique to be used in a generic, automatic experimental design environment, is done by use of a multi-criteria analysis.

Four criteria were chosen as follows. Since users do not necessarily have first-hand knowledge about the underlying simulation model it cannot be assumed that the user knows in advance what the model behaviour will be like. This results in the first criterion. The second criterion stems from the potential erratic behaviour by discrete event simulation models, caused by failures and other disturbances that exist in the model. The third criterion reflects how often a technique is used or mentioned in literature. The last criterion reflects the purpose of metamodeling, i.e. cost saving. This results in the following criteria.

- No prior knowledge required
- Ability to model capricious behaviour
- Usage
- Cost saving
3.1 Metamodeling techniques
The four metamodeling techniques, as well as their scoring on the criteria are briefly described below. For a more comprehensive description refer to Kolenbrander (2012).

3.1.1 Polynomial regression metamodeling
Polynomial regression metamodeling (also known as response surface methodology) is a technique that has been used for over 30 years (Box & Draper, 1987). Polynomial regression metamodeling essentially consists of fitting a polynomial function to a simulation model’s response. Each variable \( x_n \) of a polynomial, similar to Equation 2, represents an input factor. Determining the result of an unknown part of the response surface can be accomplished by filling in the known data points into the fitted polynomial.

\[
Y = m_1 x_1^2 + m_2 x_2^2 + m_3 x_1 + m_4 x_2 + m_5 x_1 x_2 + b
\]

Equation 2 - Standard second order polynomial model with two factors and an interaction effect

3.1.2 Spline metamodeling
Spline is an interpolation technique that connects two nearest known points with a separate polynomial function while maximizing the smoothness of the overall curve at the intersections at each point. In other words, the derivatives of polynomials intersecting at a known data point must be equal to one another at that point.

3.1.3 Kriging
Kriging is an interpolation technique, which estimates unknown points based on the value and proximity to known points on the response surface. Originally used in geostatistics (Krige, 1951) it has been suggested to be used in simulation metamodeling (Stein, 1987; van Beers & Kleijnen, 2004). Kriging assumes that the closer the input data are, the more positively correlated the outputs are (van Beers & Kleijnen, 2004).

3.1.4 Artificial Neural Networks
Artificial Neural Networks are networks of simulated neurons, interconnected in a similar way as the human brain’s neurons (Fonseca, Navaresse, & Moynihan, 2003). It relies on creating causal relationships between the input and output factors of a simulation model by identifying the unknown (hidden) internal variables within the model. The Artificial Neural Network (ANN) method has been used successfully as a metamodeling technique for discrete event simulation with high accuracy (Kilmer, Smith, & Shuman, 1994). These neural networks are incrementally constructed by gradual learning, which can require a large number experiments (Al-Hindi, 2004).

3.1 Multi criteria analysis
In table 1 below is shown how each of the listed metamodeling techniques scores on each of the criteria. Refer to Kolenbrander (2012) for a structured and more detailed description of the scoring process.
Not one metamodelling technique scores perfectly on all the criteria. Still, the best performing techniques can be chosen. Polynomial Regression and Kriging can both be seen as having the highest potential for generic usage in a metamodelling environment and were therefore chosen to be compared on their accuracy, described in the next section.

4 Quantitative comparison of metamodelling techniques

4.1. Research setup

The performance of the two metamodelling techniques, polynomial regression metamodelling and kriging metamodelling are compared on their accuracy using eight different responses from both simulation models and custom artificial data.

4.1.1 Criteria

Four criteria are used to determine the suitability of each of the chosen metamodelling techniques for general use in a commercial simulation project, which are; (1) Prediction accuracy, (2) Cost saving, (3) Handling of erratic responses and (4) No prior knowledge required.

Prediction accuracy and handling of erratic responses are both captured in the quantitative analysis performed in this section. The accuracy is determined using a range of coefficients that measure the goodness of fit between estimated data and actual data, which are described thoroughly by Kolenbrander (2012), of which $R^2$ is the most important in this paper. The measurement of performance handling erratic responses is captured with a number of artificial responses such as a step function and constrained functions. The cost saving criterion is handled by using a similar amount of experiments for each metamodelling technique. The criterion regarding the necessity of prior knowledge is discussed in section 3.

4.1.2 Research steps

In order to determine the accuracy of the metamodelling techniques on each response five steps used; (1) a sample set of experiments is created using one of the sampling techniques described earlier, (2)
these experiments are run, meaning that the accompanying output is retrieved, (3) the metamodel is fitted to the available results, (4) the experiments outside the sample set are estimated using the metamodel, (5) the results can be compared with the actual (i.e. known) data from either the simulation model or pre-set artificial responses. For Kriging this process is repeated 100 times for each simulation model, since the sampling step is based on a random process.

4.1.3 Testing data
Four artificial responses are used to compare performance differences when estimating fundamental response structures such as a step function, a constrained and unconstrained polynomial function, and a logistics function. Four simulation model responses are used to compare performance differences when estimating actual discrete event simulation models. Relatively basic models such as an M/M/1 system and an inventory system also used by Biles et al. (2007) and Zakerifar et al. (2009) have been used as well as a model simulating gas station queues and processes, provided by simulation consultancy firm Systems Navigator, which ships the model with their scenario management software. For a more thorough description of the responses used in the research refer to Kolenbrander (2012).

4.2 Metamodeling technique-specific choices

4.2.1 Regression metamodeling
Polynomial regression metamodels require knowledge about the behaviour of the model to select the order of the used polynomial function. In research done by Kolenbrander (2012) it was determined that it is not possible to select the order of the polynomial based on the goodness of fit of the regression function on the observed data. Since higher order polynomials could result in inherent instabilities (Jin, Chen, & Simpson, 2001) both second and third order regression metamodels are used in the research done in this section.

4.2.2 Kriging metamodeling
There exist different types of Kriging (Kleijnen, 2009; Kleijnen & Beers, 2004): Simple Kriging, Ordinary Kriging and Universal Kriging. The differences have to do with expected value of the underlying model i.e. the ‘default’ value for a Kriging model in case a known data point is too far away to exert any influence. In the research of Kriging accuracy ordinary Kriging is used. Also, according to Kleijnen (2009), it suffices in practice.

4.3 Sampling
As said in section 2.2, a sampling technique can have a large impact on the cost saving aspect as well as the accuracy of the metamodel. This is why the sampling methods most used for both metamodeling techniques are described below.

4.3.1 Regression metamodeling
One of the most used sampling techniques for regression metamodeling, and the sampling technique best suitable for generic use is the central composite design (see figure 1) (Biles, et al., 2007).
This design essentially forms a circle around the centre point and has 13 design points for a two-dimensional design space (two factors).

### 4.3.2 Kriging metamodeling

Latin Hypercube Sampling (LHS) is a randomly generated design and ensures that every level of every factor is used exactly once (see figure 2). It is widely promoted as the best design to be used in conjunction with Kriging (Biles, et al., 2007; Kleijnen, 2009; Kleijnen & Beers, 2004; van Beers & Kleijnen, 2004; Zakerifar, et al., 2009). However, since it is generated randomly, its quality is not consistent. Gunzburger and Burkardt (2004) describe Latin hypercube design quality, or ‘coverage’ to have the following three attributes; (1) the points are equally spaced, (2) the points cover the region and (3) the points are isotropically distributed, i.e., there is no directional bias. Since various methods exist to construct a Latin Hypercube, it is important to determine which method produces high quality designs. Figure 4 shows the results of a comparison of the coverage of six different LHS design methods. Note that since the coverage is not constant, a box plot is used to display both the median as the spread of the coverage one thousand randomly generated designs.

The design method OptimumLHS, which uses an algorithm “to maximize the harmonic mean distance from each design point to all the other points in the design” (Stocki, 2005), performs best in terms of average accuracy of the resulting metamodel and has been used in the research of Kriging performance.

### 4.4 Research results

In Figure 4 the results are shown of the comparison between the accuracy of regression metamodeling and Kriging metamodeling.
Figure 4 - Comparison of metamodeling techniques in terms of artificial function estimation accuracy

There is a **statistically significant difference** between all pairs of results (Kolenbrander, 2012). One result stands out, which is the large difference in accuracy between the two techniques regarding the Step function response. This difference is shown in Figure 5. Regression metamodeling clearly has trouble handling erratic data, while Kriging handles it rather well.

Figure 5 - Comparison of step function response estimation by regression and Kriging metamodeling techniques

On overall, Kriging has a better average performance than regression metamodeling. Looking at the responses of the simulation models the difference between the techniques is largest at the Inventory and Gas Station 2 models. The responses of these models were not easily described by a second or third order polynomial function. An example of this can be seen in Figure 6, which compares the original response of the Gas Station 2 model with both metamodeling techniques.

Figure 6 - Comparison of step Gas Station 2 model estimation by regression and Kriging metamodeling techniques
5. Conclusions

Based on the qualitative analysis in section 3 and the quantitative analysis in section 4 it can be concluded that Kriging metamodeling is a suitable technique for generic use in a commercial simulation project. The strongest advantages of Kriging over other metamodeling techniques are the fact that it can be used without any prior knowledge of the behaviour of a simulation model, its overall accuracy and its ability to handle erratic response behaviour. These characteristics make it possible to use Kriging as a metamodeling technique to handle many kinds of simulation models, regardless of the expected behaviour of the responses, without the need for specific metamodeling expertise.

A direct application of the outcome of this research is the implementation of the Kriging algorithm in a scenario management tool that assists users with creating experiments / scenarios for simulation. This is where a generic metamodeling technique can be used to create a sample set prior to simulation and reproduce the non-simulated data prior to the display of the results.

6. Discussion points and recommendations

There are a number of discussion points that focus on limitations or assumptions done in this paper that could restrict applicability of the outcomes to specific situations. At the same time recommendations are done for further research.

No prior knowledge required

An influential criterion in this paper was focused on whether or not setting up the metamodel requires prior knowledge about the model’s behaviour. This criterion was chosen based on the assumption that the users of the metamodeling environment do not require to be knowledgeable in doing simulation projects. Simulation experts could have the possibility to have obtained already more knowledge about the behaviour of system that is observed. Given this prior knowledge, other metamodeling techniques might be available to give more accurate results than a technique that assumes ignorance. Further research on the effects of having prior knowledge on the chosen metamodeling technique can be done.

Using metamodels for optimisation

While not discussed in this paper, metamodels have the potential to be used in (constrained) optimization (Biles, et al., 2007; Zakerifar, et al., 2009). Traditionally (constrained) optimization is performed based on various algorithms that incrementally ‘search’ for the optimum. The cost of such optimisation can be reduced by using metamodeling. The algorithms used to find the optimum is applied on the metamodel instead of the simulation model. Further research can determine what kind of metamodeling techniques are best suited to be used together with optimization.
References


