SIMULATION METAMODELS FOR MODELING OUTPUT DISTRIBUTION PARAMETERS

Isabel R. Santos

Dept. of Mathematics, Centre for Management Studies
Technical University of Lisbon (IST)
Av. Rovisco Pais, 1049-001 Lisboa, PORTUGAL

Pedro R. Santos

Department of Informatics, INESC-ID
Technical University of Lisbon (IST)
Av. Rovisco Pais, 1049-001 Lisboa, PORTUGAL

ABSTRACT

Metamodels are functions with calibrated parameters, used as abstractions and simplifications of the simulation model. A metamodel exposes the system's input-output relationship and can be used as an analysis tool for solving optimization problems or as a surrogate for building blocks in larger scale simulations. Our approach is to analyze statistically the response by modeling the normal distribution mean and variance functions, in order to better depict the problem and improve the knowledge about the system. The metamodel is checked using the confidence intervals of the estimated distribution parameters, and new design points are employed for predictive validation. An example is used to illustrate the development of analysis and surrogate metamodels.

1 INTRODUCTION

Complex processes are often described by large models, considered as good reality approximations. However, decision support, exploratory analysis and rapid adaptive calculations frequently require easy to understand and explain representations. A simulation metamodel (Barton 1992) is a model of the simulation model, and exposes the fundamental nature of the system's input-output relationship through a simple mathematical function. Graphical representations of the metamodel are also useful for providing a simple and easy form to analyze and communicate.

A simulation metamodel is defined as an abstraction and simplification of the simulation model. A metamodel is built to meet a particular set of objectives, and a metamodel valid for one objective might not be for another (Davis and Bigelow 2003). Metamodels are not necessarily rigorous representations of the original model but they should strive to offer a small degree of uncertainty. Even though the most cost-effective metamodel is not necessarily the most valid, the information provided must be reliable up to a certain degree.

The construction of a metamodel starts with the simulation model that is run several times for different input values. Preferably, these input values should be determined by an appropriate experimental design. A robust design experiment works well across a broad range of scenarios and provides solutions that are less likely to produce unanticipated results (Kleijnen et al. 2005). Then the resulting simulation responses, or output data, and a statistical model are inferred from those responses, using some form of regression. To simplify the metamodel, one may combine some of the inputs and eliminate others that proved to be redundant. Finally, an appropriate mathematical function with a few undetermined parameters must be calibrated, by estimating those parameters, so that the metamodel fits the response.

Different techniques have been proposed for the construction of simulation metamodels. Linear polynomial approximations are easy to implement and are frequently used by simulation researchers (Kleijnen and Sargent 2000, Cheng and Kleijnen 1999). Nonlinear approaches provide more flexible and intuitive approximations but are more complex. These methods include nonlinear regression (Santos and Nova 1999, Santos and Nova 2006), Kriging (Kleijnen and van Beers 2005, Allen, Bernshteyn, and Kabiri 2003) and neural networks (Badiru and Sieger 1998). Other mechanisms include rational metamodels (Hendrickx and Dhaene 2005), radial basis functions (Jin, Chen, and Simpson 2000) and Bayesian approaches (Cheng 1999, Chick 1997). Metamodels have been used for gaining an overview of the system's behavior within an experimental region and to iteratively select new inputs at which to evaluate the system (Kleijnen and van Beers 2004, Santos and Santos 2007). Metamodels can be used as building blocks in larger scale simulations in order to increase the efficiency of computer experiments (Barton 1997). Metamodels are also useful for solving optimizations problems using less expensive computer resources (Jin, Du, and Chen 2001, Cheng and Currie 2004).

In spite of their shortcomings, metamodels can provide, in an easy to grasp form, the behavior of some useful measure of performance, or response, of the simulation model. However, rather than merely estimating the future average

performance of a system, given a particular design and control strategy, additional information should be provided as to how common is such behavior to be expected. Improving the metamodel, relative to a baseline of simple mathematical functions, requires the use and understanding of the statistical nature of the output. Our approach is to analyze statistically the response and to model besides the mean value, of some quantity, its variance. This complements the usual metamodel representation of the mean value with an indication of how reliable the mean value is and how common are similar to be expected. The resulting metamodel uses mean and variance functions to better depict the problem and improves the knowledge about the system. As a consequence a well informed decision maker can have greater confidence and assurance when judging the behavior of the system.

The next section discusses the use of metamodels as analysis tools and as simulation surrogates or building blocks. Section 3 describes the process for modeling output distribution parameters and the construction of the resulting metamodel. In section 4 an example is used to illustrate the building of such a metamodel and depict the resulting expressive capacity. Section 5 is reserved for conclusions.

2 DISTRIBUTION BASED METAMODELS

An important task in simulation is to discover the relationship, if any, between the simulation input X and the response of interest Y. If the input-output is plotted and a visually check suggests some relationship between Y and the X, then we would expect to express this relationship through some function f, that is,

$$Y \approx f(X). \tag{1}$$

The utilization of the function f allows the prediction of Y for a given X. For example, Y could be the steady-state average waiting time in system and X the mean time between arrivals. for a queuing system (Law and Kelton 2000). Mathematically, the relation (1) can be written as the metamodel

$$Y = f(X; \theta) + \varepsilon$$

where ε represent the noise and f belongs to a known parametric family of functions, $\mathscr{S}$. Finding the metamodel reduces to estimating the parameter vector θ. The family $\mathscr{S}$ should be large and flexible enough to approximate a comprehensive variety of functional relationships. If two or more metamodels fit the data equally well, it is desirable to choose the simplest one. Even when a linear approximation works well, a nonlinear metamodel may still be used to retain a clear interpretation of the parameters (Santos and Nova 2001).

The proposed distribution based metamodel emulates not only a mean value but also the estimated standard deviation, i. e., a dispersion measure at each design point. In practice, besides fitting the estimated averages, the estimated variances are also fitted. This approach is useful when the model has unequal variances along the region of interest, as is frequently the case when the system exhibits a complex behavior. These regions require a more careful and detailed study in order to better understand the evolution and performance of the system.

2.1 Analysis Metamodels

The goal of a metamodel is to aid in the purpose for which the simulation model was built. So, the metamodel is developed as a mechanism designed to improve the knowledge and to better analyze the simulation model. In this context, it may be important to use metamodels that describe the mean and the variance of the response.

Suppose a simulation experiment performed according to some experimental design $\{X_p : p = 1, \ldots, n_p\}$. For each design point, n_r independent replications of the simulation model are carried out and a set of output values

$$\{Y_{prk} : p = 1, \ldots, n_p; r = 1, \ldots, n_r; k = 1, \ldots, n_k\}$$

is obtained. Y_{prk} is the k-th observation correspondent to the r-th replication of the experimental point p.

If our goal is to study the mean response, then we consider the output

$$\{\bar{Y}_{pr.} = \sum_{k=1}^{n_k} Y_{ijk}/n_k, p = 1, \ldots, n_p, r = 1, \ldots, n_r\}.$$

If the output values, $\bar{Y}_{pr.}$, are obtained from a sample of dimension sufficiently large, then we may apply the central limit theorem and consider the hypothetical approximated distribution

$$\bar{Y}_{pr.} \sim \mathrm{N}(\mu_p, \sigma_p^2) \tag{2}$$

where

$$\left[\begin{array}{c} \mu_p \\ \sigma_p \end{array} \right] \approx \left[\begin{array}{c} g_1(X_p; \theta_\mu) \\ g_2(X_p; \theta_\sigma) \end{array} \right].$$

In this case, the goal is to estimate the parameter vectors θ_μ and θ_σ. The resulting metamodel needs two functions to model the normal distribution estimated parameters, providing a better understanding of the simulation response. Although, the metamodel values given by the evaluation of the functional relationship in any of the points within the design region are parameters of the normal distribution, the metamodel is not considered a statistical model.

2.2 Surrogate Metamodels

A metamodel can also be thought as a simulation component that may be used and reused instead of the original simulation model. The metamodel can be reused stand-alone as a surrogate for a complete simulation model or may be reused as a building block in a larger simulation model. To be reused, a metamodel is encapsulated in a module with a defined interface, providing limited functionality and able to be used within a defined architecture. A component may be regarded as an input-output transfer function, since its reuse should offer all the apparent benefits of function reuse.

A successful reuse strategy should include features to support technical aspects like abstraction, selection, specialization and integration (Pidd 2002). A metamodel is itself an abstraction providing succinct, high level descriptions of reusable artifacts that are essential in assisting the developers in understanding their purpose, nature and behavior. Selection enables the identification of previously developed metamodels as Components Off The Shelf (COTS) to aid the developer in performing reuse. Parameterization of metamodels, such as θ_μ and θ_σ parameters, support specialization of components since its is unlikely to reuse without modification. Integration provides a mechanism for the combination, connection and communication between reuse artifacts enabling the reuse of a metamodel in a simulation context.

When a metamodel is reused, the mean and variance parameters obtained when evaluating the metamodel are used to drive a normal distribution generator (see equation (2)). The values resulting from the normal distribution exhibit a stochastic behavior similar as to what is to be expect from the original simulation model. In this context, the analysis of the system may take into account deviations from the average behavior and the occurrence of rare events.

The reuse of a metamodel, as building block in an event driven simulation model, requires the construction of the metamodel as a delay line. This implies that the input variable is the time between arrivals and that the response, or output variable, is the time spent inside the metamodel. Furthermore, the observed responses must be observations of individual entities rather than means or other quantities. The output of the metamodel, when used as a black-box building block, corresponds to a delay assigned to each arriving entity. The delayed entity is then scheduled to the metamodel output node, using the standard simulation scheduling procedure.

For this approach, we consider only one simulation run, and the sequence to be analyzed at each design point p should be $Y_{p11}, \ldots, Y_{p1n_k}$. The metamodel also results from the estimation of the vector parameters θ_μ and θ_σ and by hypothesis the output values are

$$\bar{Y}_{p1k} \sim N(\mu_p, \sigma_p^2). \qquad (3)$$

3 METAMODEL CONSTRUCTION

The process of building a successful distribution based metamodel requires a sequence of steps. As in other metamodeling processes, a region of interest must be defined and an appropriate experimental design should be developed, resulting in a set of considered relevant design points or inputs. Then the necessary data is collected from a predefined number of replications executed at each design points. In order to proceed with the construction of the metamodel we must be assured that the data is sufficiently independent and identically distributed (*iid*), and that the resulting mean and variance estimators are reasonably valid at each design point. This step is performed using classical input modeling techniques to assess if the estimators can be used to approximate the collected assuming a normal distribution (Leemis 2004).

Once the data is collected and validated, the construction of the metamodel requires the selection of a curve to represent each of the normal distribution parameters. Each of the parameters is modeled independently and different metamodeling techniques may be used, depending weather the behavior of the parameter is linear or exhibits non-linearities. After each distribution parameter is fitted, the resulting curve must be validated in order to ensure that the expected output does not deviate significantly from the simulation data collected.

3.1 Gather Simulation Data

The collection of data must address a number of issues such as the initial and final conditions, whether to perform terminating or steady-state simulations and decide the required number of replications and run length. When executing terminating simulations, that are run a predetermined amount of time or until a certain event occurs, it may be necessary to censor the results if rare events are simulated. The results obtained from steady-state simulation, that have no natural point of termination, are biased by the warm-up period. In this case, an initial data deletion must be performed to control the bias and reduce the estimation error in the parameters of the metamodel. Techniques, such as the Welch's procedure, are used to eliminate the initial bias (Alexopoulos 2006).

When assessing the average behavior of the system with respect to a quantity of interest, for instance the time in system, the sample mean values can be obtained in different ways. Independent replications, when each run provides a single data value, provide *iid* data that enables the direct use of classical statistical techniques. Overlapping and non-overlapping batches, when the simulation is divided into sections each of these providing a single data values, can be used in steady-state simulations but the data collected

might not be considered *iid* from a simulation point of view (Alexopoulos 2006).

As a result of data collection, a set of values W_{pi} is associated with each design point p ($p = 1, \ldots, n_p$ and $i = 1, \ldots, n_i$), denoted by W_i or just W if no specific point is addressed. If an analysis model is being constructed, then the values correspond to the mean responses from (2), i. e., $W_{pr} = \bar{Y}_{pr}$. If the metamodel is to used as building block, the observation values of a single run correspond to $W_{pk} = Y_{p1k}$ from (3). In either case, the behavior of these sets of points, one set for each design point, represents the quantity to be modeled.

3.2 Output Analysis

The output analysis is used to estimate the mean and variance parameters at each design point. These values will then be used to construct the transfer functions that describe the metamodel. There are several ready-to-use readily available distribution fitting software packages on the market or included with commercial simulation languages, like BetaFit, UniFit or ExpertFit (Barton et al. 2002), that are used to find the distribution that best fits the data. These packages take the sampled data set and select, from a set of common scalar probability distributions, those that give the better values for several goodness-of-fit (*gof*) statistics. In this paper, however, we are mainly concerned with the mean and variance parameters of the normal distribution, although mean and variance values may be computed from other distributions. Also, we must assume the same distribution for all design points, with presumably different distribution parameters for each point.

In order to apply classical statistical techniques we must assess weather the observations, in the collected order, are *iid*. One technique applies linear regression to the output W_i versus the collect value number i, $i = 1, \ldots, n_i$. The regressed first degree polynomial

$$W_i = \beta_0 + \beta_1 i + \varepsilon$$

assumes that data is *iid* if the hypothesis

$$H_0 : \beta_1 = 0 \tag{4}$$

holds. If β_1 is significantly different from zero the assumption is probably not appropriate, corresponding to learning and fatigue curves for negative and positive values of β_1, respectively (Leemis 2004). However, if output data has cycles and trends the above technique does not detect such dependencies. The correlation plot is another technique that determines the correlation between adjacent observations, for different levels of separation or lags. The graphical plot represents $\hat{\rho}_j = \widehat{\text{Cor}}(W_i, W_{i+j})$ as a function of

$j = 1, \ldots, n_i - 1$. and all represented values must be close to zero if data is to be assumed *iid* (Law and Kelton 2000).

Assuming that the collected data is approximately *iid*, we may used Maximum Likelihood Estimators (MLE*s*) to determine the parameter values. For each experimental point, we may obtain the MLE*s* for the parameters of the hypothetical normal distribution of W_{pi}:

$$\hat{\mu}_p = \bar{W}_{p.} = \frac{1}{n_i} \sum_{i=1}^{n_i} W_{pi}, \qquad \hat{\sigma}_p = \left(\frac{1}{n_i} \sum_{i=1}^{n_i} (W_{pi} - \bar{W}_{p.})^2 \right)^{1/2}.$$

In the final stage of the output analysis we must ensure that the normal distribution with the estimated parameters is, in fact, a good approximation of the original output data. Statistical *gof* tests, like chi-square and Kolmogorov-Smirnov, are used to investigate whether to reject, or not, the choice of the hypothetical distribution and its parameters. The Kolmogorov-Smirnov test statistic evaluates the maximum vertical difference between the empirical and the fitted cumulative distribution functions. However, even if the *gof* test rejects the distribution and its parameters, which is likely if the data set is large, some authors suggest that the result should only be advisory (Biller and Nelson 2002). Alternatively, a graphical comparison based on the graphs of the cumulative distribution function is also sensitive to lack of fit.

3.3 Curve Selection and Parameter Estimation

The construction of a distribution metamodel requires the fitting of one curve per selected distribution parameter, mean and variance for a normal distribution. Each parameter is fitted independently and requires, in the context of regression metamodeling, the selection of a curve that approximates the input-output behavior inferred from the values obtained from each design point. Linear regression using polynomials is limited by the inability of low degree polynomials to sufficiently approximate data values and the tendency of high degree polynomials to fluctuate for values between the provided data, specially in the boundaries. If polynomials provide a poor fit, frequent when the output exhibits flat regions, other metamodel types must be used. The regression of nonlinear functions, on the other hand, requires an extensive and flexible catalog of curve candidates as well as a careful choice of good approximate initial values.

The distribution metamodel estimation uses the least squares method from numerical analysis. The least-squares criterion of estimation in the linear model remains the most commonly used method. The least squares method uses a direct method if a polynomial fit is being used, while an iterative method is required for a nonlinear function fitting. This method consists of minimizing the error sum of squares. In our case, we minimize $\text{SSE}(\mu, \theta_\mu)$ and

SSE(σ, θ_σ), where

$$\text{SSE}(\eta, \theta_\eta) = \sum_{p=1}^{n_p} [\hat{\eta}_p - g_1(X_p; \theta_\eta)]^2.$$

That is, we obtain the vectors $\hat{\theta}_\mu$ and $\hat{\theta}_\sigma$ such that $\text{SSE}(\eta, \hat{\theta}_\eta) < \text{SSE}(\eta, \theta_\eta)$ for all $\theta_\eta \in \mathbb{R}^{m_\eta}$, with both $\eta = \mu$ and $\eta = \sigma$.

3.4 Metamodel Validation

It makes sense to have some form of quality assurance so as to ensure that the metamodel is fit for its intended purpose. Metamodel validation is responsible for establishing that its output data closely resembles the output data that should be observed from the simulation model. The accuracy required from the metamodel depends on its intended use and the utility function of the decision makers. The comparison of the metamodel and simulation output data could be done using numerical statistics up to a significance level. Alternatively, the assessment could be made by using graphical plots.

When MLE*s* are used to determine the values associated with each design point, it makes sense to use the respective confidence intervals as validation boundaries. The intervals $I_p(\mu) = \hat{\mu}_p \pm \delta_p(\mu)$ and $I_p(\sigma) = \hat{\sigma}_p \pm \delta_p(\sigma)$ where

$$\delta_p(\eta) = \left(-\text{E}\left[\frac{\partial^2 \ln L_p(\hat{\mu}_p, \hat{\sigma}_p)}{\partial \eta^2} \right] \right)^{-1/2} \qquad (5)$$

and $L_p(\hat{\mu}_p, \hat{\sigma}_p)$ is the likelihood function evaluated at $(\hat{\mu}_p, \hat{\sigma}_p)$, based on the p-th design point data. For n_i large, these confidence intervals are approximate $100(1 - \alpha)\%$ (Law and Kelton 2000). If the metamodel's output, computed with the estimated parameters, is within the confidence interval, we do not reject the metamodel when H_0 holds for a given significance level, for instance 95% of the design points $p = 1, \ldots, n_p$, where

$$H_0: \quad g_1(X_p; \hat{\theta}_\mu) \in I_p(\mu) \quad \wedge \quad g_2(X_p; \hat{\theta}_\sigma) \in I_p(\sigma). \qquad (6)$$

Similarly, a predictive validation can be inferred by using the same hypothesis for a set of inputs other than the design points used for the metamodel estimation.

3.5 Data Presentation

Metamodels fill in the gaps by estimating the expected behavior. The usual metamodels are mathematical constructs, often with little, if any, intuitive value to decision makers. Mean and standard deviation from a normal distribution are tangible and easily understood quantities that complement each other to form a more complete description of the

system. Analytical representations are guided towards optimization procedures and detailed analysis, while graphical representation provide an easy to grasp overview.

A graphical representation of the metamodel's output provide an intuitive and easy to communicate depiction of the expected behavior of the target system. Although, the metamodel models two distinct quantities, mean and standard deviation, they are closely related and may graphically represented in a single plot. The output can be displayed by plotting a central mean value within an uncertainty band of variable width, whose breadth is controlled by the standard deviation value; a 2σ band width corresponds to a probability of approximately 68%. Such approach improves significantly the available information without increasing excessively the graphical presentation.

4 APPLICATION EXAMPLE

The construction of distribution metamodels is illustrated using a simple parts painting processing unit. The mean time between arrivals of the parts, in minutes, is the decision variable X modeled with an exponential distribution. The painting time is triangularly distributed between 2 and 10 minutes, with a mode of 5 minutes. However, 20 percent of the painting operations have to be repeated due to imperfections. The purpose of the simulation experiment is to express the time in the system, Y, as a function of the mean time between arrivals of parts. Two examples were developed to exercise either, analysis and surrogate, metamodels. For the development of the analysis metamodel, where the response is the mean time in system, 20 independent replications of 2000 observations were collected for each design point. The construction of the surrogate metamodel collects the time in the system of the last 50 observations of the first replication used to build the analysis metamodel, for each design point. The region of interest, or experimental region, assumes X values between 7 and 11 minutes, where the lowest interarrival time considered is near the system's saturation point.

The adopted experimental design considers a set of 9 evenly spaced designs points {7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5, 11}. Another experimental design is used to assess the predictive capability of the developed metamodels and uses the middle points of the previous design {7.25, 7.75, 8.25, 8.75, 9.25, 9.75, 10.25, 10.75}. The selection of the middle points assumes a worst case behavior of polynomial fitting, specially for high degree polynomials. In order to reduce the bias, an initial-data deletion was performed using the Welch's moving average, where 300 observations were deleted from the first design point $X = 7$ and 100 from the last point $X = 11$, for example. The model of the painting processing unit was simulated using AweSim 3.0 and the metamodels built in MATLAB 6.5 using some custom made routines.

Table 1: Design points β_1 values for *iid* hypothesis and MLE*s* estimators with the respective $\delta_p(\eta)$, given by (5), for the analysis metamodel.

p	X_p	β_{1p}	$\hat{\mu}_p$	$\delta_p(\mu)$	$\hat{\sigma}_p$	$\delta_p(\sigma)$
1	7.0	1.25	128.52	30.65	65.50	22.93
2	7.5	2.92	69.79	20.43	43.65	15.28
3	8.0	1.73	44.37	9.33	19.94	6.98
4	8.5	−0.56	29.14	3.14	6.72	2.35
5	9.0	−0.05	26.38	3.23	6.91	2.42
6	9.5	−0.03	20.95	1.89	4.03	1.41
7	10.0	0.17	17.72	1.22	2.60	0.91
8	10.5	−0.06	16.23	0.80	1.71	0.60
9	11.0	0.14	15.50	0.89	1.90	0.67

The data collected to build the analysis metamodel W_{pi} consists of 20 means for 9 design points. Even though each mean was collect from an independent replication, the values of β_1 obtained to assess the *iid* hypothesis (4) can be as high as 2.92, as depicted in the third column of Table 1. The Kolmogorov-Smirnov test does not reject the choice of a normal distribution ($\alpha = 0.05$), with the obtained MLE parameters, for every design point.
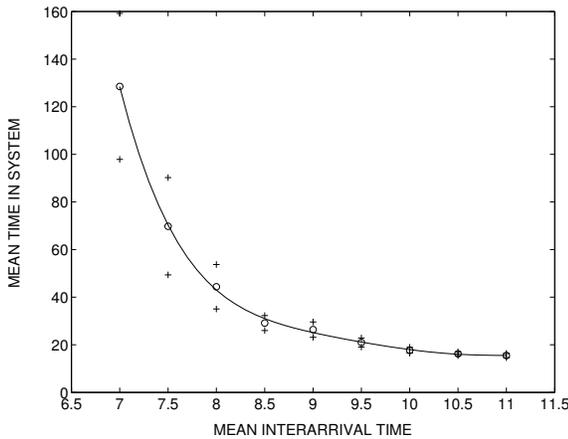

Figure 1: Estimated mean values, with the respective confidence intervals, and the resulting fitted fifth degree polynomial.

The polynomials to be selected correspond to the lowest degree that meets the hypothesis (6), as shown in Table 2.

Table 2: Number of design point rejections of the hypothesis (6) for different polynomial degrees.

degree	1	2	3	4	5	6	7	8
mean	7	7	4	1	0	0	0	0
stdev	7	6	4	5	3	0	0	0

The selected polynomials, of degrees 5 and 6 for the mean and standard deviation parameters, are depicted in

Table 3: Estimated parameters for the selected polynomials used to describe the analysis metamodel.

coefficient	mean	stdev
X^0	$4.260e4$	$−4.509e5$
X^1	$−2.150e4$	$3.017e5$
X^2	$4.343e3$	$−8.352e4$
X^3	$−438.20$	$1.225e4$
X^4	22.067	$−1.004e3$
X^5	$−0.4435$	43.681
X^6		$−0.7874$

Figures 1 and 2, their coefficients presented in Table 3, and the resulting metamodel in Figure 5.

In order to gain additional insight about the metamodel's validity, the experimental design with middle points was used to verify if the hypothesis (6) holds. The mean polynomial holds the hypothesis (6) for every middle design point, giving a high validity assurance; see Figure 3. However, in the standard deviation case, depicted in Figure 4, there are two design points that do not hold the hypothesis (6). Specifically, for $X = 8.75$ the polynomial undervalues the estimated confidence interval, while for $X = 10.75$ the polynomial overvalues the respective interval.
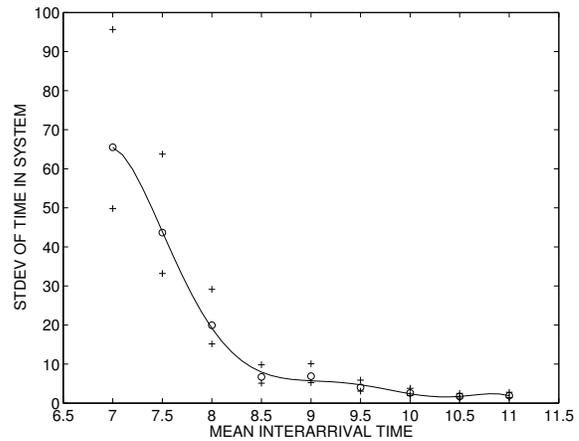

Figure 2: Estimated standard deviation values, with the respective confidence intervals, and the resulting fitted polynomial.

The data used to build the surrogate metamodel W_{pi} consists of 50 observations for 9 design points. Since every β_1 value is smaller than the highest value observed in the analysis metamodel, we do not reject the hypothesis (4) of *iid*; see Table 4. However, the Kolmogorov-Smirnov test, with a critical value of 0.1884, rejects the hypothesis of a normal distribution ($\alpha = 0.05$), for five of the nine design points. In fact, three of the rejected design points have values close to the critical, and only the first two points ($p = 1, 2$) differ significantly from the critical value. These two values correspond to input values near the system's saturation point
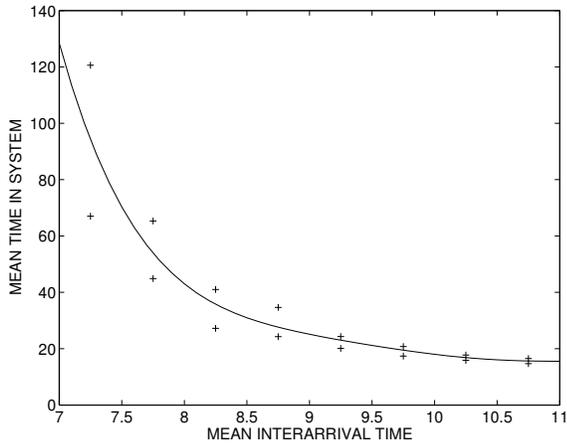
Figure 3: Comparison of the confidence intervals for the mean values, from the predictive experimental design, against the fitted polynomial.
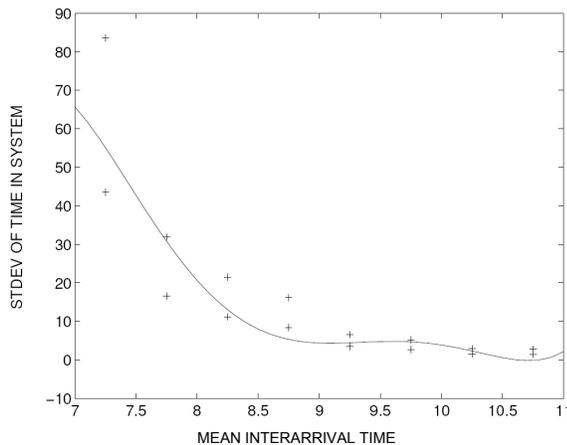

Figure 4: Comparison of the confidence intervals for the standard deviation values, from the predictive experimental design, against the fitted polynomial.

and exhibit discrepancies in their tails. Nevertheless, we are going to accept the normal distribution for two reasons. First because the proposed method requires a single distribution for every design point and some flexibility is needed. Second, some authors suggest that these tests should only be advisory, especially for large data sets, in particular if additional evidence supports the selection (Biller and Nelson 2002).

The estimated parameters of the selected polynomials, of degrees 7 and 5 for mean and standard deviation parameters (lowest degrees for hypothesis (6)), are shown in Table 5 and the resulting metamodel depicted in Figure 6.

5   CONCLUSIONS

Metamodels are abstractions of the simulation model that expose the system's input-output relationship through simple mathematical functions. The proposed distribution based metamodels analyze statistically the response using classical
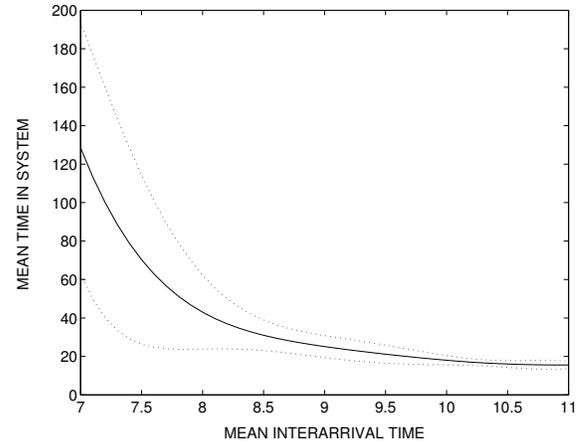

Figure 5: Final representation of the analysis metamodel, with the expected response and the associated uncertainty band.

Table 4: Design points β_1 values for *iid* hypothesis, MLE*s* estimators with the respective value of Kolmogorov-Smirnov test, for the surrogate metamodel.

p	X_p	β_{1p}	$\hat{\mu}_p$	$\hat{\sigma}_p$	K-S test
1	7.0	1.11	579.93	317.92	0.4226
2	7.5	−0.65	221.57	113.19	0.3643
3	8.0	0.68	62.98	36.92	0.2154
4	8.5	0.69	45.02	23.04	0.1830
5	9.0	0.70	24.75	24.21	0.2017
6	9.5	−0.35	15.29	17.18	0.2362
7	10.0	−0.07	13.56	9.69	0.1488
8	10.5	0.10	13.44	7.72	0.1139
9	11.0	−0.28	15.08	12.75	0.1802

statistical techniques. The response is modeled as a normal distribution, at each design point, and the estimated mean and standard deviation MLE values used for the metamodel construction. Two polynomial functions fit the distribution parameters evaluated at each experimental design point. This approach, estimates variances besides averages, giving better insight of the system's behavior in regions where the model exhibits unequal variances, as is frequently the case in systems with complex behavior.

The accuracy of the polynomial functions are checked using the confidence intervals obtained when estimating the MLE values, for every design point. The pair of polynomials to be selected, are obtained by linear regression, and correspond to the lowest polynomial degree that ensures a sufficiently accurate fit. A predictive validation test is performed, employing the same technique developed for the polynomial accuracy, using an experimental design built with the middle points of the initial design.

A distribution metamodel may be developed as an analysis tool designed to improve the knowledge and judgment of the system's behavior, for example in solving optimiza-

Table 5: Estimated parameters for the selected polynomials used to describe the surrogate metamodel.

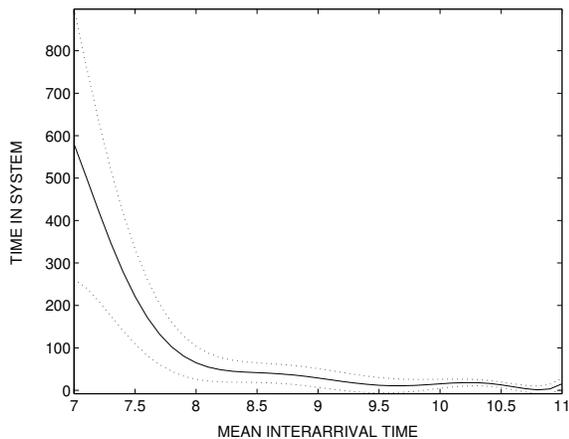| coefficient	mean	stdev
X^0	$-2.347e7$	$2.134e5$
X^1	$1.869e7$	$-1.092e5$
X^2	$-6.343e6$	$2.224e4$
X^3	$1.189e6$	$-2.254e3$
X^4	$-1.331e5$	113.62
X^5	$8.897e3$	-2.2788
X^6	-328.82	
X^7	5.1849	

Figure 6: Final representation of the surrogate metamodel, with the expected response and the associated uncertainty band.

tion problems. Likewise, the metamodel can be reused as a simulation model surrogate stand-alone or as a larger model building block. When used as a building block the metamodel behaves as a delay line, relating the interarrival time with the time in system.

A parts painting processing unit example is used to illustrated the distribution based metamodel development. The resulting distribution metamodels are depicted using a single diagram where the mean behavior is complemented by an uncertainty band representing the expected dispersion measure of the mean within the experimental region.

REFERENCES

Alexopoulos, C. 2006. A comprehensive review of methods for simulation output analysis. In *Proceedings of the Winter Simulation Conference*, ed. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 168–178.

Allen, T. T., M. A. Bernshteyn, and K. K. Y. Kabiri. 2003. A comparison of alternative methods for constructing meta-models for computer experiments. *Journal of Quality Technology* 35 (2): 1–17.

Badiru, A. B., and D. B. Sieger. 1998. Neural network as a simulation metamodel in economic analysis of risky projects. *European Journal of Operational Research* 105:130–142.

Barton, R., R. Cheng, S. Chick, S. Henderson, A. Law, L. Leemis, B. Schmeiser, L. Schruben, and J. Wilson. 2002. Panel on current issues in simulation input modeling. In *Proceedings of the Winter Simulation Conference*, ed. F. B. A. E. Yücesan, C.-H. Chen and J. M. Charnes, 353–369. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Barton, R. R. 1992. Metamodels for simulation input-output relations. In *Proceedings of the Winter Simulation Conference*, ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson, 289–299. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Barton, R. R. 1997. Design of experiments for fitting subsystem metamodels. In *Proceedings of the Winter Simulation Conference*, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, 303–310. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Biller, B., and B. Nelson. 2002. Answers to the top ten input modeling questions. In *Proceedings of the Winter Simulation Conference*, ed. F. B. A. E. Yücesan, C.-H. Chen and J. M. Charnes, 35–40. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Cheng, R. C. H. 1999. Regression metamodelling in simulation using bayesian methods. In *Proceedings of the Winter Simulation Conference*, ed. P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 330–335. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Cheng, R. C. H., and C. S. M. Currie. 2004. Optimization by simulation metamodelling methods. In *Proceedings of the Winter Simulation Conference*, ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 485–490. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Cheng, R. C. H., and J. P. C. Kleijnen. 1999. Improved design of queueing simulation experiments with highly heteroscedastic responses. *Operations Research* 47 (5): 762–777.

Chick, S. E. 1997. Bayesian methods for discrete event simulation. In *Proceedings of the Winter Simulation Conference*, ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 89–100. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Davis, P. K., and J. H. Bigelow. 2003. *Motivated metamodels*. Rand.

Hendrickx, W., and T. Dhaene. 2005. Sequential design and rational metamodeling. In *Proceedings of the Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 290–298. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Jin, R., W. Chen, and T. W. Simpson. 2000. Comparative studies of metamodeling techniques under multiple modeling criteria. Technical Report AIAA-2000-4801, AIAA.

Jin, R., X. Du, and W. Chen. 2001. The use of metamodeling techniques for optimization under uncertainty. In *Proceedings of Design Engineering Technical Conferences and Computer Information in Engineering Conferences (DETC)*. Sacramento, CA: ASME.

Kleijnen, J. P. C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa. 2005. A user's guide to the brave new world of designing simulation experiments. *INFORMS Jornal on Computing* 17 (3): 263–289.

Kleijnen, J. P. C., and R. G. Sargent. 2000. A methodology for fitting and validating metamodels in simulation. *European Journal of Operational Research* 120:14–29.

Kleijnen, J. P. C., and W. C. M. van Beers. 2004. Application-driven sequential designs for simulation experiments: Kriging metamodeling. *Journal of the Operational Research Society* 55 (8): 876–883.

Kleijnen, J. P. C., and W. C. M. van Beers. 2005. Robustness of kriging when interpolating in random simulation with heterogeneos variances: Some experiments. *European Journal of Operational Research* 165 (3): 826–834.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*. third ed. New York, NY, USA: McGraw-Hill.

Leemis, L. 2004. Building credible input models. In *Proceedings of the Winter Simulation Conference*, ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 29–40. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Pidd, M. 2002. Simulation software and model reuse: a polemic. In *Proceedings of the Winter Simulation Conference*, ed. F. B. A. E. Yücesan, C.-H. Chen and J. M. Charnes, 772–775. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Santos, M. I., and A. M. P. Nova. 1999. The main issues in nonlinear simulation metamodel estimation. In *Proceedings of the Winter Simulation Conference*, ed. P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 502–509. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Santos, M. I., and A. M. P. Nova. 2001. Validation of nonlinear simulation metamodels. In *Proceedings of Applied Simulation and Modeling*, ed. M. H. Hamza, 421–425. IASTED.

Santos, M. I., and A. M. P. Nova. 2006. Statistical fitting and validation of nonlinear simulation metamodels: a case study. *European Journal of Operational Research* 171 (1): 53–63.

Santos, M. I., and P. M. Santos. 2007, February. Sequential designs for simulation experiments: Nonlinear regression metamodeling. In *Proceedings of the 26th IASTED International Conference Modelling, Identification and Control*, 88–93. IASTED.

AUTHOR BIOGRAPHIES

M. ISABEL REIS DOS SANTOS is an assistant professor in the Department of Mathematics at the Technical University of Lisbon. She received a B.S. and M.S. degrees in Applied Mathematics and Computation from Technical University of Lisbon, and Ph.D. degree in Industrial Engineering from Technical University of Lisbon. She is a member of the Center for Management Studies of IST. Her research interests include probabilistic and statistical aspects of simulation, applications of simulation, stochastic models and simulation metamodels. Her web page can be found via <web.tagus.ist.utl.pt/~maria.santos>.

PEDRO M. REIS DOS SANTOS is an assistant professor in the Department of Informatics at the Technical University of Lisbon. He received a B.S. and M.S. degrees in Electrical Engineering from the Technical University of Lisbon. He received a Ph.D. in Computer Science from the Technical University of Lisbon. He is an invited member of the INESC-ID. His research interests include simulation, especially its programming, statistical aspects, applications, and metamodels. His web page can be found via <web.tagus.ist.utl.pt/~pedro.reis>.