



2008

Comparing designs for computer simulation experiments

Johnson, Rachel T.

IEEE

Proceedings of the 2008 Winter Simulation Conference, ed. Mason, S.J., Hill, R.R., Monch, L.,



Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

COMPARING DESIGNS FOR COMPUTER SIMULATION EXPERIMENTS

Rachel T. Johnson
Douglas C. Montgomery

Bradley Jones

John W. Fowler

Department of Industrial Engineering
Arizona State University
Tempe, AZ 85287, U.S.A

100 SAS Campus Drive
SAS Institute
Cary, NC 27513, U.S.A

Department of Industrial Engineering
Arizona State University
Tempe, AZ 85287, U.S.A

ABSTRACT

The use of simulation as a modeling and analysis tool is wide spread. Simulation is an enabling tool for experimenting virtually on a validated computer environment. Often the underlying function for the results of a computer simulation experiment has too much curvature to be adequately modeled by a low order polynomial. In such cases finding an appropriate experimental design is not easy. This research uses prediction variance over the volume of the design region to evaluate computer simulation experiments assuming the modeler is interested in fitting a second order polynomial or a Gaussian Process model to the response data. Both space-filling and optimal designs are considered.

1 INTRODUCTION

Technology has had a tremendous impact on the way problems are viewed today. Computers have changed the way systems are analyzed. Prior to the use of computers, there was limited ability to study and analyze complex scientific problems which required intensive mathematical analysis or offered no closed form mathematical solution for the problem under investigation. Now, almost every field of science and engineering makes use of computer programs and models that allow the simulation of a system.

Examples of applications of computer simulation models includes circuit simulation, stress analysis testing, hurricane tracking, turbulent flow studies, and manufacturing environments. An integrated circuit simulation is described in Currin et al. (1991). This circuit simulation was also presented and studied in Currin et al. (1988) and Sacks et al. (1989). Allen et al. (2003) describe a Finite Element Analysis (FEA) model used for designing an “interference fit” plastic seal. Modeling weather patterns over the entire globe is another application of computer simulation. Several papers on the modeling and analysis of computer hurricane models include Watson and John-

son (2004) and Iman et al. (2006). Simulation of turbulent mixing in jet engines is presented in Xiao et al. (2006) who use computational fluid dynamics (CFD) models; these models predict turbulence properties in a physical experiment. The modeling of manufacturing environments is well published, especially in the semiconductor manufacturing field, where queuing analysis is inadequate. Johnson et al. (2004a), Johnson et al. (2004b), and Johnson et al. (2005) described semiconductor manufacturing simulations.

Experiments carried out on simulation models are similar to physical experiments. The researcher performs a computer simulation experiment by making a number of systematic changes to the parameters (or inputs) of a computer simulation model. Computer simulation experiments provide several advantages over physical experimentation. First, computer simulation experiments only require the programming of the model and are limited only by the speed of the processor(s). Second, prototypes used for physical experimentation are generally expensive and require substantial time to model and build. Computer simulation experiments are comparatively cheap. They only involve the cost of a computer, the time duration to build the simulation, and the time duration to execute the runs. One disadvantage of the computer simulation experiments is their questionable ability to accurately predict the real world. While physical experiments have empirical validity, the question of whether or not a computer model is an adequate surrogate for the real system is an important consideration.

There are many studies in the literature that address the adequacy of computer simulation models. Three important topics dealing with this are the calibration, verification, and validation of computer simulation experiments via sophisticated statistical techniques. These topics are critical to the delivery of adequate computer simulation models that will be used for predicting the behavior of real processes. This research assumes that the computer simulation has been calibrated, verified, and validated,

and that it can be used to make accurate predictions about the behavior of the physical system it models.

Once an adequate computer simulation model is obtained, the next major consideration is to decide how experiments will be designed and carried out. The use of designed experiments is equally important when studying a computer simulation as it is for a physical system. The complexity of the underlying model necessitates a carefully designed experiment. Allen et al. (2003) point out that, “even though FEA is intended to reduce costs compared with physical experimentation, finite element experiments are often time consuming and costly.” Thus, computer simulations, while cheap compared to physical experiments, can often have very long run times, which requires additional planning on behalf of the experimenter. Moreover, computer simulations often have numerous input variables.

Because of the complexity of computer simulation experiments and the time consumed in making runs, it can be very convenient to develop a simpler surrogate for the computer simulation model. A surrogate model (also known as a meta-model) is a closed form mathematical expression that relates the input variables to the output response. Using a surrogate model of the computer simulation, which is also then a surrogate for the physical system, allows for very fast (microseconds) predictions of new responses at design points not yet tested. This is a very cheap alternative to running the computer code, which may take hours or days. The drawback to using a surrogate model is that it requires additional verification to demonstrate that it provides adequate approximations of the computer simulation model and the physical system of interest.

Computer simulation models used to create mathematical relationship between the inputs and response variables require the determination of an appropriate design and analysis techniques. The issues associated with the creation of designs and analysis of computer simulation experiments are sometimes different than those encountered in the physical domain.

Computer simulation models can be divided into many different categories. For instance, there are many different types of computer simulation models. Several were described earlier in the introduction. Discrete Event Simulation (DES) is frequently used to model manufacturing environments and Finite Element Analysis (FEA) is frequently used in engineering research and product development. A broader categorization of computer simulation experiments labels them as either stochastic or deterministic as illustrated in Figure 1. The division between deterministic and stochastic simulations may require completely different design and analysis techniques.

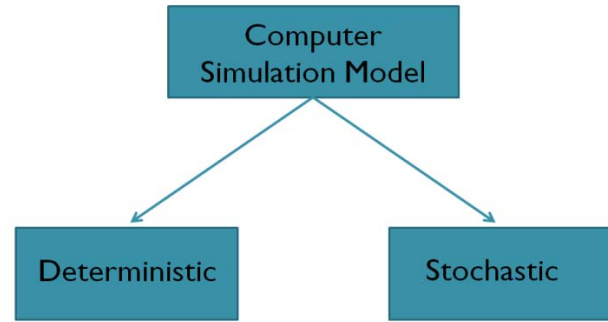


Figure 1: Computer simulation subcategories

If one is interested in the design and analysis for computer simulation models that are deterministic in nature, the designs prescribed in the current literature would fall under the category known as space filling designs. Analysis strategies for deterministic computer output include a host of model fitting techniques ranging from Kriging to Fourier regression. Space filling designs are a logical choice because they provide properties that are desirable for deterministic models. One desirable property is that each design point is unique. Another desirable property is that the uniqueness is held even if input variables are eliminated from the analysis. In other words, if the design reduces in dimension, each design will remain unique. This is an attractive property because a repeated design point would provide no additional information to the analyst. Note the properties required by deterministic simulation render traditional design strategies – replication, randomization, and blocking – useless. The choice of modeling fitting technique is another interesting topic due to the complexity of the response surface in a deterministic model. The design problem, the goal of the model, and the knowledge of the analyst all help guide in the model fitting technique selection.

Stochastic computer simulations, which contain randomness unlike deterministic models, are connected to a separate body of literature. The literature for design and analysis techniques rarely overlaps the deterministic computer simulation model literature. Design techniques associated with these models are often generated using traditional experimental design techniques such as factorial designs or using sequential design methods. Traditional methods of design are based on physical experimentation in which the response is a stochastic variable, which warrants the assumption that these methods can also apply to stochastic simulation. Randomization, replication, and blocking – techniques used in the physical experimental domain – were developed to increase the validity of the experiment and can be somewhat applicable. There are still slight differences between the stochastic computer model and the physical domain. One difference is the ability for the simulation to have some control over the randomization in the experiment. Another difference is

the lack of noise variables or the ability to control them via random distributions. Because of the similarities between physical and stochastic systems, designs such as factorial designs are often used on the stochastic model. Additionally, design optimality (alphabetic optimality criterion) is used in the creation of designs for stochastic computer simulation models. Modeling the response variables for stochastic simulation is also similar to the modeling of responses from a physical experiment. There are numerous techniques recommended for the analysis of the stochastic model outputs. Some examples include polynomial regression models, nonlinear regression models, and knowledge driven metamodels.

While both deterministic and stochastic computer simulation models have a wide variety of designs and model fitting techniques to choose from, there is little known about which designs or models should be chosen given specific situations. Measuring “goodness” or “quality” of the design is important in determining which design should be used to carry out the experiments. This research introduces a method known as Fraction of Design Space (FDS) plots to evaluate experimental designs for computer simulations. Section 2 introduces FDS plots. Section 3 presents a number of designs of interest for computer simulation models. Section 4 contains comparative plots and a discussion of general findings. This is followed by the conclusions in Section 5, which summarize the findings.

2 COMPARISON METHODOLOGY

Comparing designs based on their prediction variance is one way of evaluating the performance of designs. If the results will be used to make prediction about untried design point locations, it is important to understand how well a given design will perform at these untried location. This method is often used when comparing designs intended for use in a physical experiment, but not often employed for the study of designs for computer simulation experiments. The value of comparing designs based on their predictive capabilities is further enhanced when comparisons can be based on a graphical representation. Three powerful graphical comparison techniques are presented in the literature. Giovannitti-Jensen and Myers (1989) introduced variance dispersion graphs (VDGs), which plot the prediction variance at increasing distances from the center of the design. VDGs are used to assess the prediction capability of a response surface design. Khuri et al. (1996) illustrate the use of quantile plots for describing the distribution of the scaled prediction variance. Zahran et al. (2003) introduce fraction of design space plots that detail the scaled prediction variance (SPV) over increasing fractions of the design. These FDS plots are also used in the assessment of prediction capability for response surface designs, where the form of the model is

assumed to be a linear polynomial model. FDS plots are the most powerful graphical technique of the three because they allow a single graphical representation for each design, where the VDGs and quantile plots require multiple graphs to analyze the performance of a single design.

This research utilizes FDS plots, which plot the estimation of scaled prediction variance on the y-axis by increasing fractions of the design space on the x-axis. This is accomplished by generating 10,000 uniformly selected points within the design region and calculating the scaled prediction variance for each of these points. The variances are then sorted smallest to largest and plotted against the design volume, $[0,1]$, where 1 represents the entire design region. Figure 2 illustrates an example of an eight run I-optimal design created for a second order linear regression model in two variables.

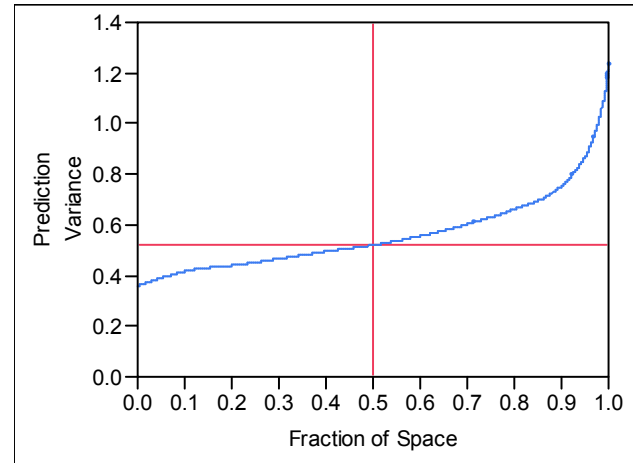


Figure 2: Fraction of Design Space Plot for an eight run I-optimal design.

To obtain an estimate of the scaled prediction variance, the assumption of a meta-model is required. We assume the analyst is interested in fitting either a second order linear regression meta-model or a Gaussian Process (GASP) meta-model. Subsections 2.1 and 2.2 will describe the prediction variance calculations for the polynomial and GASP meta-models, respectively.

2.1 Prediction Variance for the Linear Regression Model

The general form of the linear regression model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$. Thus the errors are uncorrelated with zero mean and variance, σ^2 . In equation (1), \mathbf{X} represents the design matrix and $\boldsymbol{\beta}$ represents the vector of unknown model parameters. The scaled prediction va-

variance for the linear regression model given in (1) is calculated as

$$\frac{N \text{Var}[\hat{y}(x)]}{\sigma^2} = N \mathbf{x}^{(m)'} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{x}^{(m)}, \quad (2)$$

where $\mathbf{x}^{(m)}$ is the prediction location in design variable space. The vectors represents the model of interest. The division by the variance, σ^2 , makes the quantity scale-free. The multiplication by N , the number of runs in the design, allows the comparison of designs with unequal number of design runs. The prediction variance quantity is penalized for having a larger sample size. The multiplication by N can be left out of the equation as well. Doing this can give insight into the behavior of the prediction variance with respect to adding additional design points (Myers and Montgomery (2002)).

2.2 Prediction Variance for the GASP Model

The Gaussian process (GASP) model treats the response, $y(\mathbf{x})$, as a realization of a multivariate normal distribution. Interestingly, this model is applied mostly to deterministic simulation model outputs, where there is no random error. The appeal for its use in deterministic model fitting is its ability to act as an interpolator and fit through each point perfectly. The GASP model has been shown to do an excellent job with fitting the response data from a simulation – both deterministic and stochastic – and providing fits with excellent prediction capabilities (Sacks et al. (1989) and van Beers and Kleijnen (2008)).

The GASP output response is represented as an $n \times 1$ data vector $\mathbf{y}(\mathbf{x})$, where $\mathbf{y}(\mathbf{x}) \sim N(\mu \mathbf{1}_n, \sigma^2 R(\mathbf{X}, \boldsymbol{\theta}))$. The $n \times n$ correlation matrix, $R(\mathbf{X}, \boldsymbol{\theta})$, can be represented by one of several forms (see Sacks et al., 1989). The form of the correlation function used throughout this paper is

$$R_{ij}(\mathbf{X}, \boldsymbol{\theta}) = \exp\left(-\sum_{k=1}^p \theta_k (x_{ik} - x_{jk})^2\right)$$

where $\theta_k \geq 0$. If θ_k corresponds to the correlation in the k^{th} factor. When this number is close to 0 the fitted surface in the k^{th} direction will be relatively flat, whereas a large value for θ_k corresponds to low correlation in the k^{th} factor and the fitted surface will be rough (or bumpy) in the direction of the k^{th} variable.

The fitted GASP equation is

$$\hat{y}(\mathbf{x}) = \hat{\mu} + r'(\mathbf{x}, \hat{\boldsymbol{\theta}}) R^{-1}(\mathbf{X}, \hat{\boldsymbol{\theta}}) (\mathbf{y} - \hat{\mu} \mathbf{1}_n)$$

where $\hat{\mu}$, $\hat{\sigma}$, and $\hat{\boldsymbol{\theta}}$ represent the fitted mean, variance, and theta values, respectively. Maximum likelihood is used to fit these parameters. In the fitted equation, $r_i(\mathbf{x}, \hat{\boldsymbol{\theta}})$ is an $n \times 1$ vector of estimated correlations of the unobserved $y(\mathbf{x})$ at a new value of the explanatory variables with the

observations in the data, $y(\mathbf{x})$. The form used for $r_i(\mathbf{x}, \hat{\boldsymbol{\theta}})$ in this paper is

$$r_i(\mathbf{x}, \hat{\boldsymbol{\theta}}) = \exp\left\{-\sum_{k=1}^p \theta_k (x_k - x_{jk})^2\right\}$$

Thus $\hat{y}(\mathbf{x})$ interpolates the data. Under this model, the scaled prediction variance is calculated as

$$\frac{\text{Var}(\hat{y}(\mathbf{x}))}{\sigma^2} = 1 - r'(\mathbf{x}, \hat{\boldsymbol{\theta}}) R^{-1}(\mathbf{X}, \hat{\boldsymbol{\theta}}) r(\mathbf{x}, \hat{\boldsymbol{\theta}}) + \frac{(1 - \mathbf{1}' R^{-1}(\mathbf{X}, \hat{\boldsymbol{\theta}}) r(\mathbf{x}, \hat{\boldsymbol{\theta}}))^2}{\mathbf{1}' R^{-1}(\mathbf{X}, \hat{\boldsymbol{\theta}}) \mathbf{1}} \quad (3)$$

3 DESIGN OF EXPERIMENTS

The introduction of this paper briefly described the differences in experimental design strategies for deterministic and stochastic simulation models. From our survey of the literature, we have discovered that when experimenting on a deterministic model, space-filling designs are preferred and when using a stochastic model RSM designs such as central composite designs, factorial designs, or optimal designs are preferred. In this paper we explore the performance of space-filling designs and optimal designs for linear regression models (specifically high order polynomial models) and space-filling designs for GASP models. Comparisons of these designs can be made with respect to the model used and later we will discuss potential uses for the combinations of the designs and models to a particular simulation type (deterministic or stochastic).

In this section, we present the designs that will be used in the comparison. Note that while we use only several designs, the technique can be applied to any design of interest to the reader. Section 3.1 presents the space-filling designs and section 3.2 present the optimal designs.

3.1 Space-Filling Designs

These inherent differences between computer and physical experiments have led to the development of a series of experimental designs for use in computer modeling. We refer to these designs as space-filling designs. There are numerous space-filling designs that have been proposed in the last 30 years. The space-filling designs compared in this paper are the sphere packing design, the Latin Hypercube design, the uniform design, the maximum entropy design, and the GASP Integrated Mean Square Error (IMSE) design. These designs were chosen because of their popularity in the literature and because they can be created with commercially available software packages. Two dimensional plots of these five space-filling designs can be found in Johnson et al. (2008) and Jones and Johnson (2008). The space-filling designs can be created using

several commercial software packages. A brief description of each design follows.

The sphere packing design, also known as the maximin design, was developed in Johnson et al. (1990). This design maximizes the minimum distance between pairs of designs points. The Latin hypercube design was developed by McKay et al. (1979) and this design, represented as an $n \times s$ matrix consists of a random permutation of the columns $\{1, \dots, n\}$. In this paper we use the maximin LHD. The uniform design was created by Fang (1980) and Wang and Fang (1981) and the goal of this design is to generate a set of point in the design space to be uniformly scattered, as in the uniform distribution. Shewry and Wynn (1987) developed the maximum entropy design, which uses entropy as the optimality criterion where entropy is a measure of the amount of information contained in the distribution of a data set. This design is the GASP models equivalent to the D – optimal design for the linear regression model. The GASP IMSE, created by Sacks et al. (1989) is the GASP models corresponding design to the I–optimal design for linear regression models.

3.2 Optimal Designs

Optimal designs seek to optimize a specific criterion. They were developed originally for the linear model, but as seen in the space-filling design section, there are also optimal designs (such as the GASP IMSE design) that are developed for nonlinear or nonparametric models. The optimal designs explore in this paper are I-optimal designs and D-optimal designs. I-optimal designs, or integrated variance designs, minimize the average scaled prediction variance over the design region. That is, the I-optimality criteria seeks to minimize the average value of equation (2). D-optimal designs are ones that minimize the generalized variance of the model coefficients. This is done by creating a design that maximizes $|\mathbf{X}'\mathbf{X}|$ (Myers and Montgomery (2002).

4 RESULTS

The goal of this research is to measure the “goodness” of several designs with respect to the second order linear regression model and the GASP model. The next two subsections will present results for experimental design/model fitting combinations.

4.1 Comparing Designs for the Second Order Linear Regression Model

In order to test the predictive capabilities of space-filling designs and optimal designs when fitting a second order polynomial, designs ranging from 2 – 5 factors were generated. When generating a space-filling design there is no

model specification is necessary and only the number of runs is needed. To generate an optimal design, a model is specified as well as the number of design points. For the second order linear regression model we compared the following designs: sphere packing, Latin hypercube, uniform, maximum entropy, I-optimal and D- optimal. Table 1 illustrates the number of parameters in a second order polynomial for designs with 2 – 5 factors.

Table 1: Number of model parameters for a second order linear regression model with various factors.

Factors	Number of Parameters (p)
2	6
3	10
4	15
5	21

For each case, 2 – 5 factors, we tested four separate designs with increasing number of runs. This allowed us to not only compare the design, but also study the effect sample size has on the results. We tested four different run scenarios: one design contained a minimum number of design points (shown in Table 1), the second design contained the minimum design points plus two additional points, the third design contained the minimum design points plus four additional points, and finally the fourth design contained double the number of minimum points needed. Table 2 illustrates all of the designs generated with their respective number of runs (p is equal to the number of parameters in the design as shown in Table 1).

Table 2: Number of runs required for each tested scenario.

Factors	Runs			
	p	p+2	p+4	2p
2	6	8	10	12
3	10	12	14	20
4	15	17	19	30
5	21	23	25	42

For each of the designs illustrated in Table 2, FDS plots were generated. Figure 3 illustrates FDS plots for each of the designs evaluated for a 2nd order model with two factors and 10 runs. Figure 3 shows that the I-optimal design dominates the other design by having the lowest prediction variance across 99.9% of the design region. The I-optimal design is followed by the D-optimal and sphere packing design which have equivalent prediction variance performance in this example. The worst design in the example is the maximum entropy design.

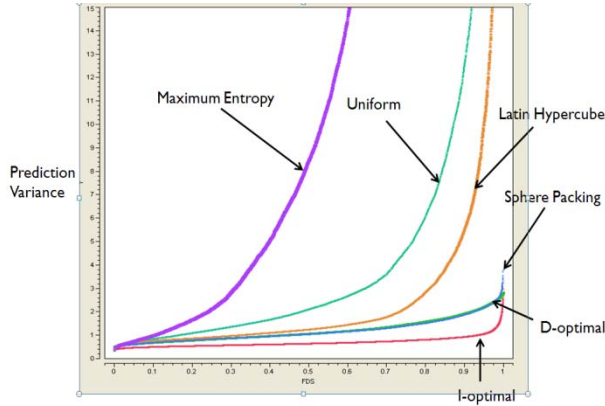


Figure 3: FDS plots for experimental designs used to fit a second order polynomial regression model.

In addition to generating the FDS plots, percentile of prediction variance for all of the designs were generated. This allowed for a tabular comparison of the designs. The results for all of these cases, as well as third order – fifth order models can be found in Johnson et al. (2008a). The results are summarized as follows:

- The I-optimal design had the best prediction variance properties of any design
- The Sphere packing designs were generally the best space-filling designs in terms of lowest prediction variance across the design region
- The space-filling designs exhibited high variability with respect to prediction variance performance

4.2 Comparing Designs for the GASP Model

For the GASP model fitting, we compared the following designs: sphere packing, Latin hypercube, uniform, maximum entropy, and Gaussian Process Integrated Mean Square Error (GASP IMSE). To compare design for GASP model based on the prediction variance requires the specification of the design, sample size, dimension (number of input variables or factors), and value of the unknown thetas (one theta for each dimension). This situation requires the use of a designed experiment to study the effect that these factors have on the prediction variance. The design of experiments is currently being conducted and will be presented in Johnson et al. (2008b). We will present some preliminary results here.

Initial findings indicate a clear ordering for most of the scenarios tested. That is, there seems to be a pattern for the dominating designs – design is a significant factor with respect to prediction variance. The design with the lowest prediction variance is generally the GASP IMSE design, followed in performance by the Latin hypercube and the uniform, which have similar performance. The sphere packing and maximum entropy designs seem to

exhibit the worst performance. Figure 4 displays an example of the FDS plots for each of the 5 designs. The plots in Figure 4 are based on a four variable, 100 run design, with all of the thetas equal to 2.75.

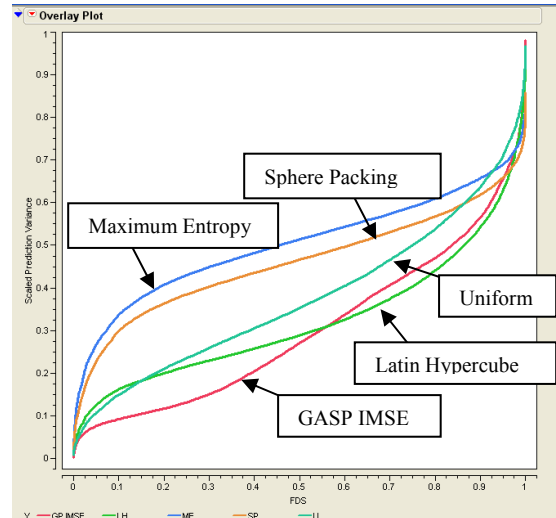


Figure 4: FDS plots for each of 5 space-filling designs with 4 variables, 100 runs, and thetas equal to 2.75.

In addition to design being a significant factor, initial results also indicate that the value of theta has a significant effect on the prediction variance properties of a design. Figure 5 illustrates three FDS plots based on three GASP IMSE designs. The designs used to create these FDS plots all had two variables and 20 runs. The difference between the designs is the estimate of the unknown theta parameter. The design with the lowest prediction variance has theta estimate of 1, the middle plot has theta estimates of 5, and the plot with the highest prediction variance across the design space has theta estimates equal to 15.

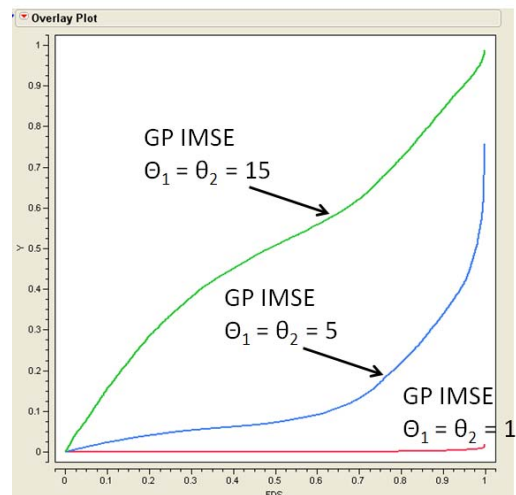


Figure 5: FDS plots for three GP IMSE designs with two variables, 20 runs, and varying estimates of theta.

5 CONCLUSIONS

Deterministic and stochastic simulation models both require careful planning and execution of experimental design strategies. Based on the type of computer simulation – deterministic or stochastic – a researcher must make choices with respect to choice of design (i.e. space-filling or optimal) and choice of model fitting technique (i.e. linear model or GASP model). This research demonstrated a new way of comparing designs for computer simulation experiments. Theoretical prediction performance of space-filling and optimal designs with respect to the second order polynomial model and the GASP model were illustrated. The theoretical results showed that there was a dominate design in both cases. The I-optimal design dominated the linear model and the GASP IMSE design dominated the GASP model. These results are somewhat intuitive as both of these designs are intended to minimize the variance in the design region with respect to the specific model. While the FDS graphical strategy is useful for comparing design types, the FDS plotting capabilities also allow the assessment of the other effects on prediction variance such as sample size, dimension, and unknown theta parameters (in the case of the GASP model).

REFERENCES

- Allen, T.T., Bernshteyn, M.A., and Kabiri-Bamoradian, K. 2003. Constructing Meta-Models for Computer Experiments. *Journal of Quality Technology* 35(3): 264 – 274.
- Currin, C., Mitchell, T. J, Morris, M. D., and Ylvisaker, D. 1988. A Bayesian Approach to the Design and Analysis of Computer Experiments. ORNL-6498, available from National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.
- Currin, C., Mitchell, T. J, Morris, M. D., and Ylvisaker, D. 1991. Bayesian Prediction of Deterministic Functions, With Applications to the Design and Analysis of Computer Experiments. *Journal of the American Statistical Association* 86:953 – 963.
- Fang, K.T. 1980. The Uniform Design: Application of Number-Theoretic Methods in Experimental Design. *Acta Math. Appl. Sinica.* 3:363 – 372.
- Giovannitti-Jensen, A. and Myers, R. H. 1989. Graphical Assessment of the Prediction Capability of Response Surface Designs. *Technometrics* 19:117 – 130.
- Iman, R.L., Johnson, M.E., and Watson, C.C. 2006. Statistical Aspects of Forecast Planning for Hurricanes. *The American Statistician* 60(2):1713 – 1726.
- Johnson, M.E., Moore, L.M. and Ylvisaker, D. 1990. Minimax and maxmin distance design. *Journal for Statistical Planning and Inference* 26:131 – 148.
- Johnson, R.T., Fowler, J.W, and Mackulak, G.T. 2005. A Discrete Event Simulation Model Simplification Technique. In *Proceedings of the 2005 Winter Simulation Conference*:2172 – 2176.
- Johnson, R.T, Leach, S., Fowler, J.W, and Mackulak, G.T. 2004a. Variance-based Sampling for Cycle Time – Throughput Confidence Intervals. In *Proceedings of the 2004 Winter Simulation Conference*:716 – 720.
- Johnson, R.T., Montgomery, D.C., Jones, B., and Peter, P.A. 2008a. Comparing Computer Experiments for Fitting High Order Polynomials. submitted to *Journal of Quality Technology*.
- Johnson, R.T., Montgomery, D.C., Jones, B., and Fowler, J.W. 2008b. A Comparison of Computer Experiments for the Gaussian Process Model. Working paper.
- Johnson, R.T., Yang, F., Ankenman, B.E. and Nelson, B.L. 2004b. Nonlinear regression fits for simulated cycle time vs. throughput curves for semiconductor manufacturing. In *Proceedings of the 2004 Winter Simulation Conference*:1951-1955.
- Jones, B. and Johnson, R.T. 2008. The Design and Analysis of the Gaussian Process Model. Submitted to *Quality and Reliability Engineering International*.
- Khuri, A., Kim, H. J., and Um, Y. 1996. Quantile Plots of the Prediction Variance for Response Surface Designs. *Computational Statistics & Data Analysis* 22:395 – 407.
- McKay, N. D., Conover, W. J., Beckman, R. J. 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21:239 – 245.
- Myers, R. H. and Montgomery, D. C. 2002. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. Wiley, New York, NY.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. 1989. Design and Analysis of Computer Experiments. *Statistical Science* 4(4):409 – 423.
- Shewry, M.C. and Wynn, H.P. 1987. Maximum entropy sampling. *Journal of Applied Statistics* 14:898 – 914.
- Van Beers, W.C.M. and Kleijnen, J.P.C. 2008. Customized Sequential Designs for Random Simulation Experiments: Kriging metamodeling and Bootstrapping. *European Journal of Operations Research* 186:1099 – 1113.
- Wang, Y. and Fang, K.T. 1981. A Note on Uniform Distribution and Experimental Design. *KeXue TongBao* 26: 485 – 489.
- Watson, C.C. and Johnson, M.E. 2004. Hurricane Loss Estimation Models: Opportunities for Improving the State of the Art. *Bulletin of the American Meteorological Society* 85:1713 – 1726.
- Xiao, X., Edwards, J.R., Hassan, H.A., and Cutler, A. 2006. Variable Turbulent Schmidt-Number Formulation for Scramjet Applications. *AIAA Journal* 44(3):593 – 599.

Zahran, A., Anderson-Cook, C. M., Myers, R. H. 2003. Fraction of Design Space to Assess Prediction Capability of Response Surface Designs. *Journal of Quality Technology* 35(4):377 – 386.

AUTHOR BIOGRAPHIES

RACHEL T. JOHNSON is a PhD student in the Department of Industrial Engineering at Arizona State University. Her research areas of interest are the design and analysis of computer simulation experiments and simulation methodology. Her research is funded by the National Aeronautics and Space Administration (NASA).

DOUGLAS C. MONTGOMERY is a Regents' Professor and the ASU Foundation Professor of Engineering at Arizona State University. He is co-director of the ASU Committee on Statistics. His research interests are in industrial statistics. He is a recipient of the Shewhart Medal, the Brumbaugh Award, the Lloyd S. Nelson Award, the Hunter Award, and the Shewell Award (twice), all from the American Society for Quality, the Deming Lecture Award from the American Statistical Association, and the Ellis R. Ott Award. He is the one of the Chief Editors of *Quality & Reliability Engineering International*, a former Editor of the *Journal of Quality Technology*, and a member of several other editorial boards. Professor Montgomery is a Fellow of the American Statistical Association, a Fellow of the American Society for Quality, A Fellow of the Royal Statistical Society, a Fellow of the Institute of Industrial Engineers, an Elected Member of the International Statistical Institute, and an Academician of the International Academy for Quality.

BRADLEY JONES is the Director of Statistical Research & Development in the JMP Division of SAS Institute. He is the inventor of the prediction profile plot - an interactive graph for exploring multivariate response surfaces. He holds a patent on the use of designed experiments for minimizing registration error in multi-layer laminated circuit boards. He is a fellow of the American Statistical Association and serves as an Associate Editor of *Technometrics* and *Quality Engineering*.

JOHN W. FOWLER is a Professor of Industrial Engineering at Arizona State University (ASU). His research interests include modeling, analysis, and control of manufacturing and service systems. He is the Co-Director of the Modeling and Analysis of Semiconductor Manufacturing Laboratory at ASU which has done research for the National Science Foundation, the Semi-conductor Research Corp., International SEMATECH, Asyst, IBM, Intel, Motorola, Infineon Technologies, and ST Microelectronics. Dr. Fowler is an author on over 65 journal

publications, 90 conference papers, and 10 book chapters. He is an Area Editor for *SIMULATION: Transactions of the Society for Modeling and Simulation International* and for *Computers and Industrial Engineering*, an Associate Editor for *IEEE Transactions on Semiconductor Manufacturing*, and on the Editorial Board for *IIE Transactions* and the *Journal of Simulation*. He is an IIE Fellow, the INFORMS Vice President for Chapters/Fora, and is on the Winter Simulation Conference Board of Directors.