Proceedings of the 1995 Winter Simulation Conference
ed. C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman

METAMODEL APPLICATIONS USING TERSM

Michael A. Zeimer

ISQ Department
P.O. Box 6696
Consolidated Freightways, Inc.
Portland, Oregon 97208-6696, U.S.A.

Jeffrey D. Tew

Department of Management
The Oregon Graduate Institute
Portland, Oregon 97291, U.S.A.

ABSTRACT

Tactical simulation models are often used to assess vulnerabilities and capabilities of combat systems and doctrines. Due to the complexity of tactical simulation models, it is often difficult to assess the relationship between input factors and the performance of the simulation model. To facilitate this type of assessment, simulation analysts often use the simulation model to emirically construct a *black-box* approximation of the causal and time dependent behavior of the simulation model. This type of approximation is known as a *metamodel* and can be viewed as a summary of the behavior of the simulation model. We demonstrate this technique in the context of an example using TERSM (Tactical Electronic Reconnaissance Simulation Model). The results indicate that metamodeling is applicable to tactical simulation models and that the technique has a wide range of uses.

1 INTRODUCTION

Tactical simulation models are often employed by the Department of Defense to assess the capabilities and vulnerabilites of various combat systems and doctrines. These simulation models are usually highly complex and of relatively high dimensionality. That is, the performance of the simulation model is dependent on a large number of parameters or input factors that act and interact in a complex manner. Thus, it is often difficult to assess the relationship of individual input factors to the performance of the simulation model. Recently, a technique known as *metamodeling* has generated interest in the simulation community for its ability to facilitate this type of assessment.

A metamodel is a mathematical approximation of the relationship between a set of input factors and one or more responses. Metamodels are usually estimated empirically via experimentation with a simulation model, and thus, metamodels are models of models. With respect to a given response, a metamodel is *black-box* approximations of the causal (mechanistic) and time dependent behavior of a simulation model. Figure 1 depicts the relationships among the real system, the simulation model, and the metamodel.

In this report, we introduce metamodeling and illustrate its applicability to the analysis of tactical simulations. In Section 2, we summarize the mathematical and statistical concepts and notation of metamodeling. In Section 3, we present an example using TERSM (Tactical Electronic Reconnaissance Simulation Model). In Section 4, we present some conclusions.

2 METAMODELS

Metamodels can have various forms, but we restrict our attention to the most commonly used class of models: least squares models. To simplify the discussion, we focus on polynomial and simple transformed response polynomial models of the forms

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

and

$$\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2}$$

where $\mathbf{y}$ is an $n \times 1$ vector of responses, $\mathbf{X}$ is an $n \times p$ data matrix containing the levels of the input factors, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown metamodel coefficients, $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of error terms, and $\mathbf{y}^*$ is a vector of transformed responses. The transformation on y can be any real function over the range of the untransformed response. Functions such as the square root and natural logarithm are often used to linearize sets of observations in order to obtain simpler and/or better approximations of system behavior.

For example, the relationship between a pair of factors, x_1 and x_2, and a response, y, may have the polynomial form

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon \quad (3)$$

or the transformed response polynomial form

$$\sqrt{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_1 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon. \quad (4)$$

Both metamodels are said to be linear models since all coefficients have power one.

The type of metamodel to use is often dictated by the purpose of the metamodel and by properties of the system. Metamodels are usually employed for one or more of the following purposes:

1. studying system behavior,
2. predicting responses,
3. sensitivity analysis, or
4. optimization.

Depending on the purpose of the metamodel, the form and the fineness of the approximation may vary greatly. For example, a simple linear approximation is often adequate for studying some elements of system behavior such as the degree to which certain factors affect the response while nonlinear approximations may be more appropriate for prediction.

Some of the important properties of the system that influence the type of model used include:

1. characteristics of the response (discrete or continuous, qualitative or quantitative, random or deterministic, etc.),
2. characteristics of the input factors (discrete or continuous, qualitative or quantitative, random or deterministic, etc.), and
3. dimensions of the experimental region.

In this paper, we restrict our attention to systems with quantitative, continuous responses; and quantitative, deterministic input factors. Metamodels for these systems can be estimated using the *method of least squares*. We consider both random and deterministic response cases. Metamodels can also be obtained using more advanced techniques which are beyond the scope of this report. The techniques outlined in the following subsections are applicable to random responses in general, but a subset of the outlined techniques are applicable to the deterministic response case as well. Thus, we will explain all the techniques in terms of the random response case and note exceptions for the deterministic response case.

In Section 2.1, we discuss least squares model estimation. In Section 2.2, we briefly summarize a pair of statistical analysis tools called *analysis of variance* and *statistical inference*. In Section 2.3, we briefly discuss some measures and methods for determining the validity of metamodels. In Section 2.4, we introduce and briefly discuss some techniques for efficiently designing experiments. Finally, in Section 2.5, we add some perspective to material in Sections 2.1-2.4 by outlining a general metamodeling process.

2.1 Least Squares Metamodel Estimation

To illustrate the method of least squares, consider a set of observations (y_i, $i = 1, 2, \ldots, n$) and corresponding set of factor levels (x_{ij}, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, p$) given by

$$\begin{array}{ccccc} y_1 & x_{11} & x_{12} & \cdots & x_{1p} \\ y_2 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ y_n & x_{n1} & x_{n2} & \cdots & x_{np} \end{array}.$$

Suppose we postulate a model given by

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \varepsilon. \quad (5)$$

In this case the number of parameters, p, is equal to five, corresponding to the number of coefficients in the postulated model. This same model can be written in general vector notation as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Now suppose that we obtain an estimated model given by

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}, \quad (6)$$

where $\hat{\mathbf{y}}$ is an $n \times 1$ vector of estimated responses, and $\mathbf{b}$ is a $p \times 1$ vector of estimated model coefficients. The $n \times 1$ vector of deviations of the observations about the fitted model, called the vector of residuals, is given by

$$\mathbf{e} = \hat{\mathbf{y}} - \mathbf{y}. \quad (7)$$

The least squares estimator of $\boldsymbol{\beta}$ is obtained by setting the derivative with respect to $\mathbf{b}$ of the sum of squared residuals equal to zero, such that

$$\begin{aligned} \frac{\delta}{\delta \mathbf{b}}[\mathbf{e}'\mathbf{e}] &= \mathbf{0} \quad (8) \\ \frac{\delta}{\delta \mathbf{b}}[(\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b})] &= \mathbf{0} \quad (9) \\ -2\mathbf{X}'\mathbf{y} + 2(\mathbf{X}'\mathbf{X})\mathbf{b} &= \mathbf{0}. \quad (10) \end{aligned}$$

Simplification leads to the least squares estimator

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{11}$$

(see Myers 1990, p. 88). Thus, least squares estimates are estimates for which the unweighted sum of squared residuals is minimized.

For example, consider the estimated simple linear regression model given by

$$\hat{y} = b_0 + b_1 x.$$

The ith predicted response, observed response, and residual are denoted by $\hat{y}_i$, y_i, and e_i respectively. The given model is a least squares model if and only if the sum of the squared vertical distances from each observation to the fitted model is minimized.

2.2 Analysis of Variance and Statistical Inference

Analysis of variance (ANOVA) and *statistical inference* are statistical methods that are commonly employed to quantify the importance of factors with respect to a given response. They are extremely useful and powerful because they allow the analyst to make statements concerning the statistical significance of various factors. For example, these methods could be used to determine if it is *likely* that x_1x_2 in the metamodel given by equation (5) affects the response. Without ANOVA and statistical inference, randomness and/or lack of fit between the metamodel and the observed data make such conclusions difficult to reach.

The principal drawback of ANOVA and statistical inference is that they require certain assumptions concerning the behavior of ε. In particular, the assumption that ε be normally and independently distributed with homogeneous variance is required. This assumption implies that there must be some random *noise* in the response which causes observed responses to be normally distributed with equal dispersion about the estimated model, independent of the location in x-space. These methods can still be employed in violation of the assumptions, but the results are unpredictable, undependable, and should be treated with suspicion especially when the response is not random (i.e. the deterministic response case).

To illustrate ANOVA and statistical inference, consider the ANOVA table for the metamodel given by equation (5), shown in Table 1. The purpose of ANOVA table construction is to break-down variability in the response and assign portions of the variability to sources of variation based on the observed contribution of each source. Contributions to variability are measured using sums of squares, and the corresponding degrees of freedom represent restriction on the calculation of sums of squares.

Table 1

ANOVA Table for the Metamodel Given by Equation (5)

Source of Variation	Degrees of Freedom	Sum of Squares
Metamodel	5	$\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{y}$
Error	$n-5$	$\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{y}$
Total	n	$\mathbf{y}'\mathbf{y}$

In addition to the basic ANOVA table, it is also possible to subdivide the metamodel sum of squares given in Table 1 in order to account for the variability due to individual model terms. For the metamodel given by equation (5), the variability due to x_1x_2 is given by

$$\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{y} - \mathbf{y}'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2\mathbf{y}, \tag{12}$$

where $\mathbf{X}_2$ is the data matrix without the fourth column (which corresponds to x_1x_2).

Sums of squares are used to measure the variability in the response because they possess useful distributional properties when our assumptions concerning ε hold. We can take advantage of these distributional properties to conduct statistical hypothesis tests (see Myers 1990, pp. 95-125). An hypothesis test is a formal means of quantifying the probability that an assertion is incorrect. For example, consider the hypothesis that $\beta_{12} = 0$, or in other words, that the interaction between x_1 and x_2 is insignificant. In the formal notation of hypothesis testing, this can be stated as

$$H_0 : \beta_{12} = 0 \text{ versus } H_1 : \beta_{12} \neq 0.$$

If the null hypothesis is correct, then under the given assumptions

$$\frac{\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{y} - \mathbf{y}'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2\mathbf{y}}{s^2} = F_{\alpha,1,n-p}, \tag{13}$$

where $F_{\alpha,1,n-p}$ is α point of an F-distribution with 1 and $n-p$ degrees of freedom (see Myers and Milton 1991, p. 116). The value of α for which equation (13) holds is called the *p-value* and is the probability that H_0 is true. Thus, a very low p-value for the given test indicates that it is highly likely that β_{12} explains a significant portion of the variability in the response. Such coefficients are said to be statistically significant.

2.3 Validation

The validity of a metamodel indicates the degree to which the specified purpose of the metamodel can be accomplished Sargent (1991a). For example, a simple linear approximation may be valid for studying some elements of system behavior but completely invalid as a means of making predictions. Validity is also specific to the experimental region used to develop the metamodel. In other words, the metamodel is expected to be valid for a specific purpose over the experimental region.

Validity can be measured using many available diagnostics. For a complete discussion of diagnostics for the random response case see Myers (1990, Chapter 4). Diagnostics for the deterministic response case are discussed in Kleijnen (1987). In this section, we simply discuss the diagnostics used in the example in Section 3.

One diagnostic that is appropriate for both deterministic and random responses is the squared coefficient of determination, R^2, which is given by

$$R^2 = \frac{\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}}. \tag{14}$$

Note that the numerator is the sum of squares for the metamodel and the denominator is the total sum of squares. Thus, R^2 measures the proportion of the total variability in the response explained by the metamodel. The higher R^2 the better the metamodel fits the given data. While R^2 provides a good, general measure of fit it does not measure the uniformity of fit. In other words, a metamodel with a high R^2 may have some areas of very poor fit as long as there are relatively large areas with very good fits. Also, R^2 only measures the degree to which the estimated metamodel fits the data that is used to estimate the metamodel. Thus, R^2 does not account for fit in areas where there is no data. In addition, for the random responses case, the use of R^2 by itself gravitates the model selection to an overfit model (one which tracks random error). This has the detrimental effect of reducing the prediction capability of the metamodel (see Myers 1990, pp. 179-180).

In order to test the validity of a metamodel across the entire experimental region, analysts often advocate a technique known as data splitting (see Myers 1990, pp. 169-170). Data splitting is applied by using some observations to fit a model, and a separate set of observations to measure the validity of the model. This allows the validity of the metamodel to be tested independently of the data used to fit the model. In cases where data is expensive and/or difficult to acquire, this may be impractical. Often data that is not used to fit the metamodel is supplemented with data that is used to fit the metamodel in order to measure validity. This can result in misleading measures of validity.

Another approach to validation involves the use of a diagnostic known as the PRESS statistic (see Myers 1990, pp. 170-178). To calculate the PRESS statistic, a set of PRESS residuals is calculated by mathematically *factoring out* the dependence of each observed residual on the data used to estimate the metamodel. This method eliminates the data splitting problems and is applicable for both the random and deterministic response case. However, the details of the PRESS statistic are beyond the scope of this report.

Two diagnostics that are appropriate for data splitting in the deterministic response case are the maximum absolute error (MAE) and the average absolute relative error (AARE). MAE is simply the absolute value of the largest residual. By basing model validation on MAE, the model selection is gravitated to a uniform but not necessarily good fit.

The AARE is given by

$$AARE = \frac{\sum_{i=1}^{n} |e_i/y_i|}{n}. \tag{15}$$

This error is similar to R^2 in that it provides a good, general measure of fit, but it has the same drawback as R^2 in that use of the AARE does not insure uniformity of fit.

Note that the difficulties with the individual diagnostics can be overcome by using them in combination. This is done in the example in Section 3.

2.4 Design of Experiments

The purpose of experimental design is to obtain *better* estimates and predictive models with fewer observations by carefully constructing $\mathbf{X}$. This is done by preselecting certain values for the k factors in the experiment. Assuming that we are using an unbiased estimator such as a least squares estimator, the quality of experimental designs is usually measured with the variance of prediction. The variance of prediction is the variance of the true population about the fitted model at some arbitrary location in x-space, $\mathbf{x}_0$ and is given by

$$\mathrm{var}[\hat{y}(\mathbf{x}_0)] = \sigma^2 \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0. \tag{16}$$

For any set of similarly *scaled* competing designs with the same number of observations, it can be shown that $\mathrm{var}[\hat{y}(\mathbf{x}_0)]$ is minimized when $\mathbf{X}'\mathbf{X} = nI$ (see Myers 1976, p. 109). In this case, $\mathbf{X}$ is said to be orthogonal. The minimization of the variance of prediction for orthogonal designs is due in part to the fact that orthogonal designs result in models with uncorrelated coefficients. This is not true of any other designs. Thus, orthogonal or near-orthogonal designs should be used whenever possible.

We consider two basic designs in this report: (1) 2^k factorials and (2) central composite designs (CCD). A 2^k factorial experiment consists of k factors each at two levels arranged in all possible factor/level combinations. To obtain an orthogonal design, the levels of the input factors are usually centered and scaled such that high level of the factor appears as a 1 and the low level of the factor appears as a -1. The centering and scaling transformation is given by

$$x_i = 2\left(\frac{\xi_i - \bar{\xi}}{d_i}\right), \tag{17}$$

where ξ_i is the level of the ith input factor, $\bar{\xi_i}$ is the average of the low and high levels of ξ_i, x_i is the centered and scaled level of the ith input factor, and d_i is the spacing between the low and high levels of ξ_i. While centering and scaling input factors in designed experiments usually results in better metamodels, it may require the analyst to perform some extra work in order to analyze the model. For example, if an analyst needs to predict a response at some point $\boldsymbol{\xi}_0$ using a metamodel for centered and scaled input factors, then he must rescale $\boldsymbol{\xi}_0$ to $\mathbf{x}_0$ using the same centering and scaling formula used in the experiment.

To illustrate, consider a 2^2 factorial experiment replicated r times for the purpose of estimating the regression model given by

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon. \tag{18}$$

The corresponding design matrix is given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r \\ \mathbf{1}_r & -\mathbf{1}_r & -\mathbf{1}_r & \mathbf{1}_r \\ \mathbf{1}_r & \mathbf{1}_r & -\mathbf{1}_r & -\mathbf{1}_r \\ \mathbf{1}_r & -\mathbf{1}_r & \mathbf{1}_r & -\mathbf{1}_r \end{bmatrix} \tag{19}$$

where r is the number of replications of the experiment and $\mathbf{1}_r$ is an $r \times 1$ column vector of ones.

$$\mathbf{b} = \frac{1}{r}\mathbf{X}'\mathbf{y}. \tag{20}$$

An important advantage of factorial experiments over one-variable-at-a-time experimentation is that we can estimate the interaction effects of the factors on the response. Further details on factorial experiments can be found in Box, Hunter, and Hunter (1978, Chapter 10).

A central composite design (CCD) consists of a 2^k factorial design augmented with $2k+1$ extra design points to allow the estimation of second order models. To illustrate, consider a CCD replicated r times for the purpose of estimating the regression model given by

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon. \tag{21}$$

The corresponding design matrix is given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r \\ \mathbf{1}_r & -\mathbf{1}_r & -\mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r \\ \mathbf{1}_r & \mathbf{1}_r & -\mathbf{1}_r & -\mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r \\ \mathbf{1}_r & -\mathbf{1}_r & \mathbf{1}_r & -\mathbf{1}_r & \mathbf{1}_r & \mathbf{1}_r \\ \mathbf{1}_r & \boldsymbol{\alpha}_r & \mathbf{0}_r & \mathbf{0}_r & \boldsymbol{\alpha^2}_r & \mathbf{0}_r \\ \mathbf{1}_r & -\boldsymbol{\alpha}_r & \mathbf{0}_r & \mathbf{0}_r & \boldsymbol{\alpha^2}_r & \mathbf{0}_r \\ \mathbf{1}_r & \mathbf{0}_r & \boldsymbol{\alpha}_r & \mathbf{0}_r & \mathbf{0}_r & \boldsymbol{\alpha^2}_r \\ \mathbf{1}_r & \mathbf{0}_r & -\boldsymbol{\alpha}_r & \mathbf{0}_r & \mathbf{0}_r & \boldsymbol{\alpha^2}_r \end{bmatrix}. \tag{22}$$

Note that when $\alpha = 1$, the CCD design is near-orthogonal and allows for the estimation quadratic curvature. A near orthogonal design for, higher order models can be obtained by layering CCDs. This concept is demonstrated in the example in Section 3. Further details on CCDs can be found in Myers (1976, pp. 127-134).

3 TERSM EXAMPLE

The Tactical Electronic Reconnaissance Simulation Model (TERSM) was built in 1969 by the Rand Corporation, for the purpose of making comparative performance evaluations of a variety of airborne direction-finding systems. Simulating a reconnaissance mission through a pulsed radar environment, its primary output is a lower bound on the emitter location accuracy attainable by accumulation and processing of bearing measurements. These measures of emitter location accuracy are known as Circular Error Probabilities, of CEPs. In essence, it is the imaginary circle or ellipse around an emitter, of such size that the probability of that emitter's actually falling in the circle is 50%. Obviously, the smaller the CEP, the more accurate the associated location estimate.

The model was designed to simulate a reconnaissance mission in sufficient detail to assess the influence of variations of system design parameters and input factors on overall system performance. Thus, by altering input factors and parameters, analysts can use TERSM to compare and contrast proposed airborne direction-finding systems and tactics.

To demonstrate the usefulness of metamodeling in this type of assessment and to illustrate the metamodeling process, we elected to perform a system behavior study. In particular, we wanted a reasonably accurate approximation of the relationship between the number of emitters located on the *test* mission within five nautical miles or less CEPs, and four input factors: (1) altitude in feet, (2) velocity in knots, (3) azimuth angle in degrees, and (4) channel capacity in number of channels. The test mission consists of a set of hostile and friendly radar emitters of various types arranged in set locations on a hypothetical battlefield. Historically, the test mission has been the standard scenario used to compare competing systems. We selected the following experimental region: (1) altitude from 5000 to 40000 feet, (2) velocity from 186 to 1150 knots, (3) azimuth angle from 60 to 150 degrees, and (4) and channel capacity from 4 to 30 channels. The selected factors and experimental region were chosen based on previous studies with TERSM. Note that the response is continuous and deterministic and that all of the input factors are continuous and deterministic except for channel capacity, which is discrete and deterministic. To accomplish our stated purpose, we specified the following goals for the validity of the model: (1) R^2 of at least 95%, (2) MAE less than 100, and (3) AARE less than 5%. All 49 observations given in Appendix I were used to validate the model.

To obtain a satisfactory metamodel, we made seven model fitting iterations. The observations used to estimate the metamodels were drawn from the validation test set. Although the stated purpose of the metamodel does not include simplicity, all terms with p-values greater 50% were eliminated sequentially from each model in order to obtain models of manageable size. Usually p-values of around 10-30% are used to eliminate insignificant factors, but since we were dealing with an obvious violation of assumptions (i.e. a deterministic response), 50% was used for conservatism. All models were estimated for centered and scaled factors levels. The correspondence between the actual input factors used in the experiment and the centered and scaled input factor levels used to fit the metamodels is given in Table 2a and Table 2b.

Table 2a

Correspondence Between Actual Input Factor Levels and Centered and Scaled Input Factor Levels

Input Factor	Variable	-1	-0.5
Altitude	x_1	5000	13750
Velocity	x_2	186	427
Azimuth	x_3	60	82
Channel Cap.	x_4	4	10

Table 2b

Correspondence Between Actual Input Factor Levels and Centered and Scaled Input Factor Levels

Input Factor	Variable	0	0.5	1
Altitude	x_1	22500	31250	40000
Velocity	x_2	668	909	1150
Azimuth	x_3	105	128	150
Channel Cap.	x_4	17	24	30

The following is the sequence of postulated and fitted metamodels:

Metamodel 1

$$\begin{aligned} y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\ & \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \\ & \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \\ & \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \beta_{234} x_2 x_3 x_4 + \\ & \beta_{1234} x_1 x_2 x_3 x_4 + \varepsilon \end{aligned}$$

$$\begin{aligned} \hat{y} = {} & 224.118 + 85.750 x_2 + 57.750 x_3 + 89.000 x_4 + \\ & 27.750 x_1 x_4 + 21.000 x_2 x_3 + 47.250 x_2 x_4, \end{aligned}$$

Metamodel 2

$$\begin{aligned} y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\ & \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \\ & + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \\ & \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \beta_{234} x_2 x_3 x_4 + \\ & \beta_{1234} x_1 x_2 x_3 x_4 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \\ & \beta_{33} x_3^2 + \beta_{44} x_4^2 + \varepsilon, \end{aligned}$$

$$\begin{aligned} \hat{y} = {} & 532.633 + 11.944 x_1 + 91.944 x_2 + 61.778 x_3 + \\ & 102.333 x_4 + 16.000 x_1 x_3 + 27.75 x_1 x_4 + \\ & 21.000 x_2 x_3 + 47.250 x_2 x_4 + 19.750 x_3 x_4 + \\ & 12.250 x_1 x_2 x_3 + 12.500 x_1 x_3 x_4 - \\ & 30.608 x_1^2 - 212.608 x_2^2 - 86.108 x_3^2, \end{aligned}$$

Metamodel 3

$$\begin{aligned} \ln y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\ & \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \\ & \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \\ & \beta_{123} x_1 x_2 x_3 + \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \\ & \beta_{234} x_2 x_3 x_4 + \beta_{1234} x_1 x_2 x_3 x_4 + \beta_{11} x_1^2 + \\ & \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{44} x_4^2 + \varepsilon, \end{aligned}$$

$$\begin{aligned} \widehat{\ln y} = {} & 6.346 - 0.047 x_1 + 0.417 x_2 + 0.306 x_3 + \\ & 0.467 x_4 - 0.025 x_1 x_2 + 0.075 x_1 x_3 + \\ & 0.170 x_1 x_4 + 0.098 x_2 x_4 - 0.028 x_3 x_4 + \\ & 0.066 x_1 x_2 x_3 - 0.078 x_2 x_3 x_4 - 0.049 x_1 x_2 x_3 x_4 - \\ & 0.133 x_1^2 - 0.679 x_2^2 - 0.125 x_3^2 - 0.363 x_4^2, \end{aligned}$$

Metamodel 4

$$\begin{aligned}
\sqrt{y} = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\
& \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \\
& \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \\
& \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \beta_{234} x_2 x_3 x_4 + \\
& \beta_{1234} x_1 x_2 x_3 x_4 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \\
& \beta_{33} x_3^2 + \beta_{44} x_4^2 + \varepsilon,
\end{aligned}$$

$$\begin{aligned}
\widehat{\sqrt{y}} = {} & 23.319 + 2.994 x_2 + 2.042 x_3 + 3.288 x_4 + \\
& 0.488 x_1 x_3 + 1.006 x_1 x_4 + 0.407 x_2 x_3 + \\
& 1.155 x_2 x_4 + 0.257 x_3 x_4 - 0.400 x_1 x_2 x_3 - \\
& 0.288 x_2 x_3 x_4 - 0.841 x_1^2 - 5.757 x_2^2 - 0.701 x_3^2 - \\
& 2.683 x_4^2,
\end{aligned}$$

Metamodel 5

$$\begin{aligned}
y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\
& \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \\
& \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \\
& \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \beta_{234} x_2 x_3 x_4 + \\
& \beta_{1234} x_1 x_2 x_3 x_4 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \\
& \beta_{44} x_4^2 + \beta_{111} x_1^3 + \beta_{222} x_2^3 + \beta_{333} x_3^3 + \\
& \beta_{444} x_4^3 + \beta_{1111} x_1^4 + \beta_{2222} x_2^4 + \\
& \beta_{3333} x_3^4 + \beta_{4444} x_4^4 + \varepsilon,
\end{aligned}$$

$$\begin{aligned}
\hat{y} = {} & 549.756 - 32.722 x_1 - 111.537 x_2 - 62.778 x_3 + \\
& 180.556 x_4 + 16.441 x_1 x_3 + 26.647 x_1 x_4 + \\
& 20.882 x_2 x_3 + 46.971 x_2 x_4 + 18.676 x_3 x_4 + \\
& 12.508 x_1 x_2 x_3 + 9.785 x_1 x_2 x_4 + 12.892 x_1 x_3 x_4 + \\
& 61.972 x_1^2 - 22.962 x_3^2 - 86.491 x_4^2 + \\
& 44.667 x_1^3 + 203.481 x_2^3 - 78.222 x_4^3 - 90.454 x_1^4 - \\
& 210.918 x_2^4,
\end{aligned}$$

Metamodel 6

$$\begin{aligned}
\ln y = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\
& \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 \\
& + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \\
& \beta_{123} x_1 x_2 x_3 + \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \\
& \beta_{234} x_2 x_3 x_4 + \beta_{1234} x_1 x_2 x_3 x_4 + \\
& \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{44} x_4^2 + \\
& + \beta_{111} x_1^3 + \beta_{222} x_2^3 + \beta_{333} x_3^3 + \beta_{444} x_4^3 + \\
& \beta_{1111} x_1^4 + \beta_{2222} x_2^4 + \beta_{3333} x_3^4 + \beta_{4444} x_4^4 + \varepsilon,
\end{aligned}$$

$$\begin{aligned}
\widehat{\ln y} = {} & 6.350 - 0.047 x_1 - 0.308 x_2 + 0.075 x_3 + \\
& 0.267 x_4 - 0.022 x_1 x_2 + 0.073 x_1 x_3 + \\
& 0.162 x_1 x_4 + 0.021 x_2 x_3 + 0.099 x_2 x_4 - \\
& 0.029 x_3 x_4 + 0.066 x_1 x_2 x_3 + 0.014 x_1 x_2 x_4 - \\
& 0.078 x_2 x_3 x_4 - 0.049 x_1 x_2 x_3 x_4 + 0.260 x_1^2 - \\
& 0.125 x_3^2 - 0.359 x_4^2 + 0.725 x_2^3 + 0.230 x_3^3 + \\
& 0.199 x_4^3 - 0.398 x_1^4 - 0.682 x_2^4,
\end{aligned}$$

Metamodel 7

$$\begin{aligned}
\sqrt{y} = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \\
& \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \\
& + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \\
& \beta_{123} x_1 x_2 x_3 + \beta_{124} x_1 x_2 x_4 + \beta_{134} x_1 x_3 x_4 + \\
& \beta_{234} x_2 x_3 x_4 + \beta_{1234} x_1 x_2 x_3 x_4 + \beta_{11} x_1^2 + \\
& \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{44} x_4^2 + \\
& + \beta_{111} x_1^3 + \beta_{222} x_2^3 + \beta_{333} x_3^3 + \beta_{444} x_4^3 + \\
& \beta_{1111} x_1^4 + \beta_{2222} x_2^4 + \beta_{3333} x_3^4 + \beta_{4444} x_4^4 + \varepsilon,
\end{aligned}$$

$$\begin{aligned}
\widehat{\sqrt{y}} = {} & 23.567 - 0.669 x_1 - 2.842 x_2 + 1.298 x_3 + \\
& 3.344 x_4 - 0.491 x_1 x_3 + 0.963 x_1 x_4 + \\
& 0.414 x_2 x_3 + 1.155 x_2 x_4 + 0.231 x_3 x_4 + \\
& 0.404 x_1 x_2 x_3 + 0.198 x_1 x_2 x_4 + 0.201 x_1 x_3 x_4 - \\
& 0.285 x_2 x_3 x_4 + 2.037 x_1^2 - 0.788 x_3^2 \\
& - 2.743 x_4^2 + 0.714 x_1^3 + 5.836 x_2^3 + \\
& 0.744 x_3^3 - 2.947 x_1^4 - 5.823 x_2^4.
\end{aligned}$$

Information on the designs used to estimate these models and the resulting validity measures are given in Table 3a and Table 3b.

Table 3a

Summary and Results of the Model Fitting Procedure

Model	R^2	MAE	AARE
1	71.7%	421.6	44.3%
2	95.5%	200.6	14.2%
3	99.2%	256.3	12.3%
4	98.2%	225.4	11.4%
5	97.4%	120.9	8.3%
6	99.4%	94.3	4.0%
7	98.9%	73.51	4.7%

Table 3b

Summary and Results of the Model Fitting Procedure

Model	Design	Obs.
1	2^k factorial w/ 1 center run	1-17
2	CCD w/ $\alpha = 1$	1-25
3	CCD w/ $\alpha = 1$	1-25
4	CCD w/ $\alpha = 1$	1-25
5	CCD w/ $\alpha_1 = 1$ and $\alpha_2 = 0.5$	1-49
6	2 layer CCD w/ $\alpha_1 = 1$ and $\alpha_2 = 0.5$	1-49
7	2 layer CCD w/ $\alpha_1 = 1$ and $\alpha_2 = 0.5$	1-49

Note the improvement in the validity as the complexity of the metamodel increases. Both models (6) and (7) satisfied our validity requirement and we stopped here. However, note that the validity test set is the same as the data set used to fit these models. Thus, as discussed in Section 2.3, these models are probably somewhat less valid than the validity measures indicate. If the process was continued, the next step would be to expand the validity test set and reevaluate the validity of the selected models.

Figures 3-6 from Zeimer, Tew, Sargent and Sisti (1993) contain surface and contour plots of models (1), (2), (5), and (7) respectively with CS azimuth angle and CS channel capacity held at zero. The prefix CS indicates the centered and scaled input factor. These graphs allow the estimated relationship between the response, CS altitude, and CS velocity to be observed independently of CS azimuth angle and CS channel capacity. They also suggest that the increasing complexity of the model manifests itself in increasingly refined shapes and the *optimum* value of the response changes as the metamodels become more accurate.

4 SUMMARY AND CONCLUSIONS

In this paper, we have introduced and successfully demonstrated a possible application of metamodeling to analysis of tactical simulations. While metamodeling is not appropriate for all simulation analysis problems, it does have a wide range of possible applications. In particular, metamodeling may provide a means of adequately aggregating simulation model behavior in hierarchical modeling schemes (see Sargent 1986, and Sisti 1989 and 1992).

REFERENCES

Box, G. E. P., Hunter, W. G., and J. S. Hunter. 1978. *Statistics for Experimenters.* John Wiley, New York.

Kleijnen, J. P. C. 1987. *Statistical Tools for Simulation Practitioners.* Marcel Dekker, New York.

Myers, R. H. 1976. *Response Surface Methodology.* Edwards Brothers, Ann Arbor, Michigan.

Myers, R. H. 1990. *Classical and Modern Regression with Applications.* PWS-KENT, Boston.

Myers, R. H. and J. S. Milton. 1991. *A First Course in the Theory of Linear Statistical Models.* PWS-KENT, Boston.

Sisti, A. F., 1989. A Model Integration Approach to Electronic Combat Effectiveness Evaluation, RADC-TR-89-183.

Sisti, A. F., 1992. Large-Scale Battlefield Simulation Using Multi-Level Model Integration Methodology, RL-TR-92-69.

Sargent, R. G. 1986. Joining Existing Simulation Programs. *Proceedings of the 1986 Winter Simulation Conference*, J. Wilson, J. Henrikson, and S. Roberts, eds., 512-516.

Sargent, R. G. 1991a. Simulation Model Verification and Validation. *Proceedings of the 1991 Winter Simulation Conference*, B. L. Nelson, W. D. Kelton, and G. M. Clark, eds., 34-47.

Sargent, R. G. 1991b. Research Issues in Metamodeling. *Proceedings of the 1991 Winter Simulation Conference*, B. L. Nelson, W. D. Kelton, and G. M. Clark, eds., 888-893.

Zeimer, M., Tew, J. D., Sargent, R., and A. F. Sisti. 1993. Metamodeling Procedures for Engagement Simulation Models. RL-TR-93.

AUTHOR BIOGRAPHIES

MICHEAL A. ZEIMER is a Systems Analyst with the Consolidated Freightways, Inc. Prior to this appointment he was a Product Analyst with Manugistics, Inc. He received his B.S. in Industrial Engineering and Operations Research in 1990 and his M.S. in Industrial and Systems Engineering in 1992, both from Virginia Polytechnic Institute and State University. His current research interests include metamodeling, hierarchical simulation modeling, experimental design, and simulation optimization.

JEFFREY D. TEW is a Senior Systems Engineer at Consolidated Freightways, Inc. and an Adjunct Associate Professor of Computer Simulation at the Oregon Graduate Institute. Prior to these positions he was on the facluty in the Industrial and Systems Engineering Department at Virginia Polytechnic Institute and State University. He has published numerous articles in many archival journals such as *EJOR*, *JSCS*, *Management Science*, *Operations Research*, and *SCS Transactions* among others. He is the *1994 WSC Proceedings* Editor as well as the Vice President/President Elect of the TIMS College on Simulation. He received a B.S. in mathematics from Purdue University in 1979, a M.S. in statistics from Purdue University in 1981, and a Ph.D. in industrial engineering from Purdue University in 1986. His current research interests include variance reduction techniques, simulation optimization, and the design of simulation experiments. He is a member of Alpha Pi Mu, ACM, ASA, IIE, IMS, INFORMS, SCS, and Sigma Xi.