



Available online at www.sciencedirect.com



Applied Mathematical Modelling 31 (2007) 2095–2110

APPLIED
MATHEMATICAL
MODELLING

www.elsevier.com/locate/apm

Global and local optimization using radial basis function response surface models

Dale B. McDonald ^{a,*}, Walter J. Grantham ^a, Wayne L. Tabor ^b,
 Michael J. Murphy ^c

^a School of Mechanical and Materials Engineering, Washington State University, P.O. Box 99164-2920, Pullman, WA, USA

^b Department of Mathematics, Marshall University, Huntington, WV 25755, USA

^c Energetic Materials Center, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

Received 1 April 2001; received in revised form 1 June 2006; accepted 30 August 2006

Available online 31 October 2006

Abstract

The focus of this paper is the optimization of complex multi-parameter systems. We consider systems in which the objective function is not known explicitly, and can only be evaluated through computationally intensive numerical simulation or through costly physical experiments. The objective function may also contain many local extrema which may be of interest. Given objective function values at a scattered set of parameter values, we develop a response surface model that can dramatically reduce the required computation time for parameter optimization runs. The response surface model is developed using radial basis functions, producing a model whose objective function values match those of the original system at all sampled data points. Interpolation to any other point is easily accomplished and generates a model which represents the system over the entire parameter space. This paper presents the details of the use of radial basis functions to transform scattered data points, obtained from a complex continuum mechanics simulation of explosive materials, into a response surface model of a function over the given parameter space. Response surface methodology and radial basis functions are discussed in general and are applied to a global optimization problem for an explosive oil well penetrator.

© 2006 Elsevier Inc. All rights reserved.

1. Introduction

We are interested in the optimal design of complex multi-parameter scientific and engineering systems. We seek the global maximum of a scalar-valued objective function $F(\mathbf{x})$, $\mathbf{x} \in R^n$. Typically, the objective function $F(\cdot)$ is not known explicitly and can only be evaluated experimentally or through a computationally intensive numerical simulation. The results of such endeavors are assumed to have yielded a set of m data points with objective function values $y_i = F(\mathbf{x}_i)$, $i = 1, \dots, m$ where $\mathbf{x}_i = [x_{1i} \dots x_{ni}]^T$ and $(\cdot)^T$ denotes

* Corresponding author.

E-mail address: yukonjack@wsu.edu (D.B. McDonald).

transpose. Furthermore, we assume that the objective function may possess multiple local maximum points which may be of interest in seeking the global optimum. In many instances the required computation time for parameter optimization of such complex systems is too great to be considered feasible because of the large number of optimization and simulation runs required.

One class of such systems, which will be considered later, involves the design of shaped-charge explosives [1], such as an oil well penetrator. The simulation of the formation and propagation of the shaped-charge explosion requires the use of complex continuum mechanics multi-physics simulation codes, often referred to as hydrocodes. The CPU time required to analyze “simple” systems of dimension two or more may be prohibitive. The data used in this paper consisted of objective function evaluations at 480 random data points in a space of seven independent variables. The multi-physics hydrocode simulation used to evaluate the objective function at these points required 10 CPU days on a fast workstation to produce the final data set.

To address these difficulties, response surface methodology will be used to develop composite global approximation models for parameter optimization of such complex systems. We will conclude by discussing how this methodology may be applied to the problem of real time or “on-line” control of systems where the objective function is not known explicitly.

2. Response surface methodology

Response surface methodology is a collection of mathematical and statistical techniques that can be used to approximate and optimize a system [2]. Specifically, we are interested in modelling scattered data obtained from experimentation or numerical simulation. Furthermore, we assume that the objective function may possess many local maxima or minima. Approaches such as linear or quadratic least squares models of an objective function may provide useful information, but would not adequately model multiple maxima and minima. For these reasons radial basis functions [3] will be used to generate response surface models. In addition, we employ models based on several different radial basis functions in order to study the approximate location of local maxima for the underlying system.

Radial basis functions have the form

$$F(\mathbf{x}) = \sum_{i=1}^m \beta_i \phi(r_i, c), \quad (1)$$

where $r_i(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_i\|$ is the distance of the point \mathbf{x} from the i th data point \mathbf{x}_i in the parameter space, $\|\cdot\|$ denotes the Euclidean norm, $\phi(\cdot)$ is a suitably chosen radial basis function, c is a user-defined constant which is usually required to be non-negative, and β_i is the radial basis coefficient corresponding to the i th data point.

2.1. Response surface modelling

We are concerned with the analysis of a complex system which involves a response

$$y = F(\mathbf{x}) \quad (2)$$

that depends on the parametric design variables $\mathbf{x} = [x_1 \dots x_n]^T$. The function $F(\cdot)$ is assumed to be unknown and may be quite complicated. The system (2) is represented by scattered data $y_i = F(\mathbf{x}_i)$, $i = 1, \dots, m$ where $\mathbf{x}_i = [x_{1i} \dots x_{ni}]^T$, and we are interested in finding the set of all local maximizers which, of course, includes the global maximizer. To study such a system, at reasonable cost, a suitable approximation to the function $F(\mathbf{x})$ will be developed. This approximation constitutes a response surface model, and can be used for parameter optimization runs to obtain information such as objective function values, gradient and Hessian information, along with the locations of extrema. The response surface model can reduce computation and processing times for complex systems from days to minutes on a typical workstation. Since many interesting engineering and scientific systems are naturally quite complex, the processing time may be extensive. The usefulness of a model which can greatly reduce such processing time is quite clear.

The implicit nature of the collected data, along with the expected or assumed character of the unknown objective function determines the type of methods used to generate a response surface model [4]. A number

of techniques exist; for example piecewise polynomial and multiple linear regression models have been used to create empirical models of systems based on the observed data from the system or process of interest [2].

As an example, a linear model (with $\mathbf{z} = \mathbf{x}$ and $N = n$)

$$y = \beta_0 + \beta_1 z_1 + \cdots + \beta_N z_N, \quad (3)$$

based on the development of the least squares problem [2] reduces to a system of the form

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (4)$$

where $\mathbf{y} = [y_1 \dots y_m]^T$ is a vector of measured function evaluations, $\boldsymbol{\varepsilon}$ is the corresponding vector of errors between the measured y_i and the modeled $y(\mathbf{z}_i)$ function values from (3), and

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_N \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} 1 & z_{11} & \cdots & z_{N1} \\ \vdots & \ddots & & \\ 1 & z_{1m} & \cdots & z_{Nm} \end{bmatrix}. \quad (5)$$

The vector of least squares estimators, $\boldsymbol{\beta}$, is sought that minimizes the sum of the square errors

$$L = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = [\mathbf{y} - \mathbf{Z}\boldsymbol{\beta}]^T [\mathbf{y} - \mathbf{Z}\boldsymbol{\beta}]. \quad (6)$$

The least squares estimators must satisfy

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = -2[\mathbf{y} - \mathbf{Z}\boldsymbol{\beta}]^T \mathbf{Z} = \mathbf{0}^T, \quad (7)$$

which simplifies to

$$\mathbf{Z}^T \mathbf{Z} \boldsymbol{\beta} = \mathbf{Z}^T \mathbf{y}.$$

Thus, we have the least squares estimator of $\boldsymbol{\beta}$ as

$$\boldsymbol{\beta} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y}. \quad (8)$$

Clearly, a linear model of the form

$$y = a + \mathbf{b}^T \mathbf{x},$$

would not be very useful for estimating maxima or minima. But the same technique could be used for, say, a quadratic model

$$y = a + \mathbf{b}^T \mathbf{x} + \mathbf{x}^T \mathbf{C} \mathbf{x}, \quad \mathbf{C}^T = \mathbf{C}, \quad (9)$$

with $\boldsymbol{\beta} = [a \ b_1 \dots b_n \ c_{11} \dots c_{mn} \ 2c_{12} \dots 2c_{n-1,n}]^T$ and $\mathbf{Z} = [x_1 \dots x_n x_1^2 \dots x_n^2 x_1 x_2 \dots x_{n-1,n}]^T$.

A least squares quadratic model will yield (at most) one interior maximum, minimum, or saddle point [5,6]. Such a model would not be suitable for approximating an objective function that has multiple extrema. To accommodate this case, we introduce radial basis functions.

2.2. Radial basis function approximation

The radial basis function method is concerned with finding a function $f(\mathbf{x})$, $\mathbf{x} \in R^n$ that approximates a real-valued function $F(\mathbf{x})$ of n variables. This method uses basis functions $\phi(r)$ that depend on the radial distance r between \mathbf{x} and each data point \mathbf{x}_i . Some consistent terminology is needed at this point. The term “objective function” refers to an unknown function $F(\mathbf{x})$ from which we have been given sampled data $y_i = F(\mathbf{x}_i)$, $i = 1, \dots, m$. The explicit form of the function $F(\mathbf{x})$ is unknown, and our goal is to accurately model this function so that we may find its global maximum. The term “model function” refers to the approximation $f(\mathbf{x}, c)$ to the objective function which we have created. Through the radial basis function method we are able to gather all types of information about this model which can then be applied to our unknown objective function.

As an introduction to radial basis functions, consider the extension to several variables of spline functions of one variable [3]. Let the values $F(\mathbf{x}_i) : i = 1, 2, \dots, m$ be given, with the sequence of points $\mathbf{x}_i : i = 1, 2, \dots, m$ strictly increasing. The linear spline that interpolates the data is composed of line segments

$$f(x) = \frac{(x_{i+1} - x)F(x_i) + (x - x_i)F(x_{i+1})}{x_{i+1} - x_i},$$

for $x_i \leq x \leq x_{i+1}$, $i \in \{1, 2, \dots, m - 1\}$. One can easily show that $f_1(x)$ is equivalent to

$$f(x) = \sum_{i=1}^m \beta_i |x - x_i|, \quad x_1 \leq x \leq x_n,$$

where β_i are radial basis coefficients chosen so that

$$f(x_i) = F(x_i), \quad i = 1, 2, \dots, m.$$

For functions of several variables, the general form of the radial basis approximation is [3]

$$f(\mathbf{x}) = \sum_{i=1}^m \beta_i \phi(\|\mathbf{x} - \mathbf{x}_i\|), \quad (10)$$

with $\boldsymbol{\beta} = [\beta_1 \dots \beta_m]^\top$ chosen to satisfy

$$f(\mathbf{x}_i) = F(\mathbf{x}_i), \quad i = 1, 2, \dots, m, \quad (11)$$

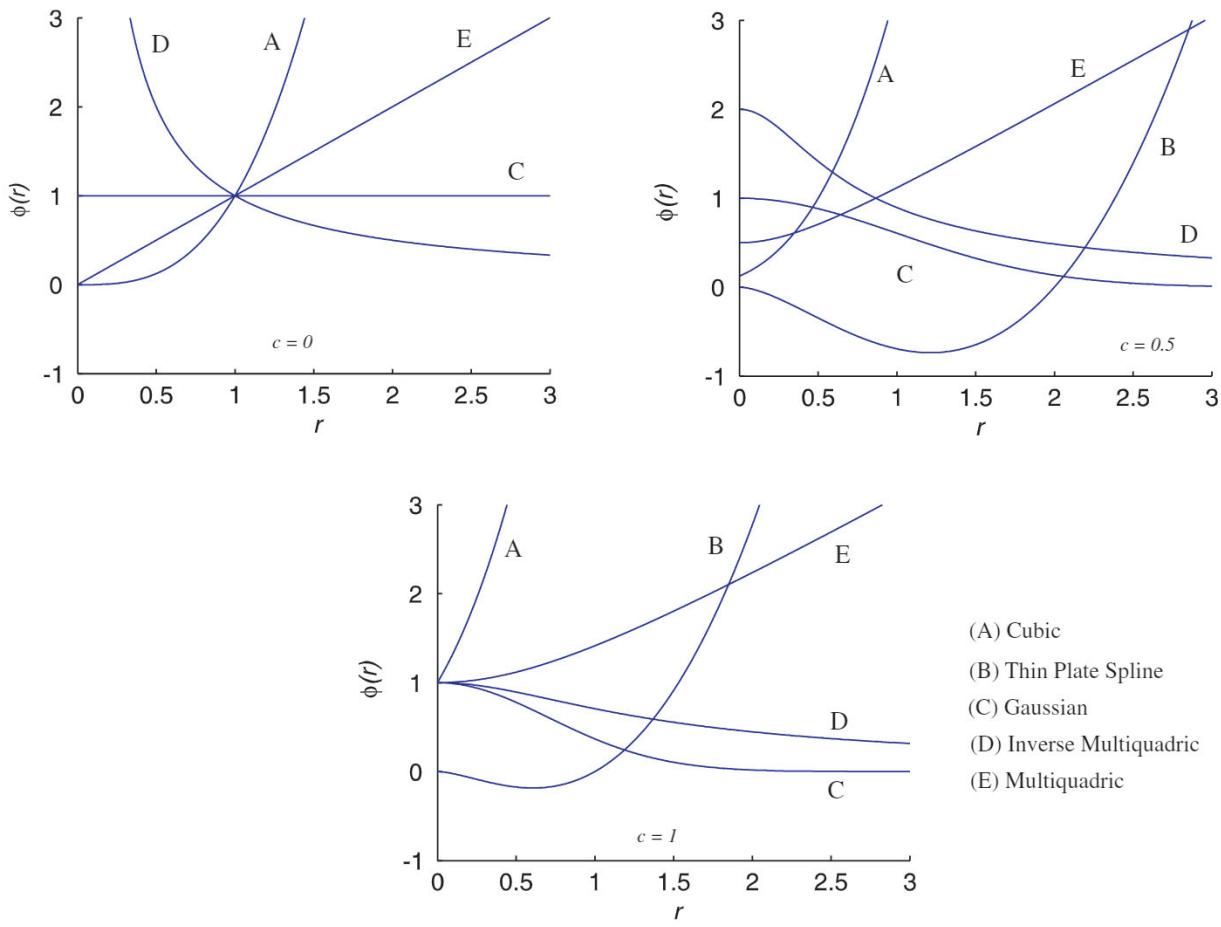


Fig. 1. Radial basis functions.

that is,

$$\mathbf{A}\beta = \mathbf{F}, \quad (12)$$

where $\mathbf{F} = [F(\mathbf{x}_1) \dots F(\mathbf{x}_m)]^\top$ and

$$\mathbf{A} = \begin{bmatrix} \phi(\|\mathbf{x}_1 - \mathbf{x}_1\|) & \cdots & \phi(\|\mathbf{x}_1 - \mathbf{x}_m\|) \\ \vdots & \ddots & \vdots \\ \phi(\|\mathbf{x}_m - \mathbf{x}_1\|) & \cdots & \phi(\|\mathbf{x}_m - \mathbf{x}_m\|) \end{bmatrix}. \quad (13)$$

The five classes of radial basis functions considered in this paper (adapted from [3]) are:

- (1) Cubic $\phi(r, c) = (r + c)^3$.
 - (2) Thin plate spline $\phi(r, c) = r^2 \ln(cr)$.
 - (3) Gaussian $\phi(r, c) = \exp(-cr^2)$.
 - (4) Inverse multiquadric $\phi(r, c) = \frac{1}{\sqrt{r^2 + c^2}}$.
 - (5) Multiquadric $\phi(r, c) = \sqrt{r^2 + c^2}$.
- (14)

Each radial basis function can be used to create a model $f(\mathbf{x})$ that approximates the true response function $F(\mathbf{x})$ [7]. The construction of such a model guarantees that the function value $f(\mathbf{x})$ of the model will be equal to the objective function value $F(\mathbf{x})$ at all data points, as seen from (11)–(13). Fig. 1 illustrates the radial behavior of each function in (14) for various values of the parameter c . Selection of an appropriate radial basis function certainly may depend qualitatively on the nature of the data and expected system response.

3. Construction of a response surface model using radial basis functions

The radial basis function method allows the user to create, for any appropriate choice of a basis function $\phi(r, c)$, a model which matches the objective function $F(\mathbf{x})$ at all data points. To accomplish this, a Response Surface Model module was created. RSM is a Fortran program which allows the user to construct a response surface model for a given data set and corresponding objective function values. RSM then determines local maxima for the model objective function over the parameter space determined by the data, and allows for objective function, gradient, and Hessian evaluations at any point. The procedure for creating such a response surface model is as follows:

1. A particular basis function $\phi(r, c)$ is chosen. In this paper we will be using functions from the five classes in (14). Within each class, a particular member is determined by the value assigned to the constant c .
2. The data points are scaled so each component of \mathbf{x} is in the range $-1 \leq x_i \leq 1$.
3. An $m \times m$ symmetric matrix $\mathbf{R} = [r_{ij}]$ is constructed, where m is the number of data points. Each entry r_{ij} of this matrix is the Euclidean distance between data points \mathbf{x}_i and \mathbf{x}_j .
4. The chosen basis function $\phi(r, c)$ is applied component-wise to the matrix \mathbf{R} , creating the matrix \mathbf{A} .
5. The matrix equation $\mathbf{A}\beta = \mathbf{F}$ is solved, where each component F_i of the vector $\mathbf{F} = [F_1, \dots, F_m]^\top$ is the objective function value at the corresponding data point \mathbf{x}_i . The resulting vector $\beta = [\beta_1, \dots, \beta_m]^\top$ is the vector of radial basis coefficients.
6. The model function value $f(\mathbf{x})$ at an arbitrary point \mathbf{x} within the parameter space is found in the following manner. A vector $\mathbf{g}(\mathbf{x}) = [g_1, \dots, g_m]^\top$ is constructed whose components g_i are obtained by the formula $g_i = \phi(r_i(\mathbf{x}), c)$, where $r_i(\mathbf{x})$ is the distance between \mathbf{x} and the i th data point \mathbf{x}_i . Then

$$f(\mathbf{x}) = \beta^\top \mathbf{g}(\mathbf{x}) = \sum_{i=1}^m \{\beta_i g_i\}. \quad (15)$$

Radial basis functions possess several characteristics that are quite appealing. For a given choice of basis function the resulting model $f(\mathbf{x}, c)$ is unique. In particular, in [3] it is shown that the symmetric matrix \mathbf{A} in

(13) is guaranteed nonsingular for all of the radial basis functions in (14) except for the thin plate spline, provided the data points \mathbf{x}_i are distinct and the number of data points m is finite. For the thin plate spline the locations of the data points used in this paper yield nonsingularity of \mathbf{A} . However, the value of c generally affects the condition number [8] of \mathbf{A} . Finally, since the radial basis functions in (14) are simple functions, it is easy to obtain function evaluation, gradient, and Hessian matrix information at any point necessary (other than at sampled data points), using analytic partial derivatives instead of numerical approximations to partial derivatives, such as forward differences. Due to the simple nature of radial basis functions, a thorough study of a complex system can be achieved in a very short period of time.

3.1. Modelling of scattered data with RSM

The Matlab “peaks” function

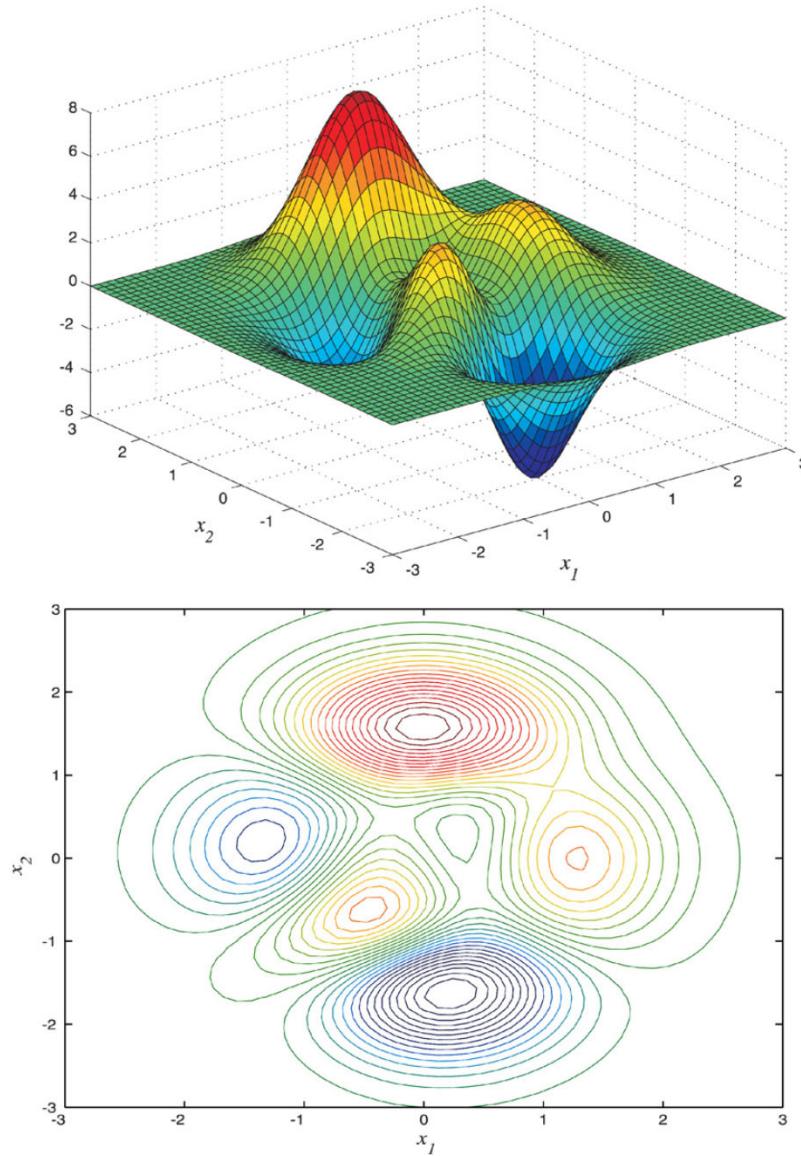
$$F(\mathbf{x}) = 3(1 - x_1)^2 \exp[-x_1^2 - (x_2 + 1)^2] - 10\left(\frac{1}{5}x_1 - x_1^3 - x_2^5\right) \exp[-x_1^2 - x_2^2] - \frac{1}{3} \exp[-(x_1 + 1)^2 - x_2^2], \quad (16)$$

provides not only an illustration of the modelling capabilities of the RSM Fortran code, but also a validation of the module itself. The peaks function consists of several maxima and minima distributed over the parameter space, as illustrated in Fig. 2. The peaks function $F(\mathbf{x})$ and its contours were plotted using a 51×51 grid (2601 plot points) in the region $-3 \leq x_j \leq 3$, $j = 1, 2$. Response surface models $f(\mathbf{x})$ of $F(\mathbf{x})$ were constructed both with random scattered data points and by sampling $m = k^2$ data points and their objective function values on a $k \times k$ grid. The results shown in Figs. 3 and 4 are based on the cubic radial basis function with $c = 0$; similar results are obtained for each of the basis functions in (14). Fig. 3 shows $f(\mathbf{x})$ and its contours constructed using an 11×11 modelling grid, but with the same 2601 point plotting grid as for the peaks function.

As previously noted, the objective function values at the data points will be reproduced by any RSM model of that function. This provides one check as to the validity of the model. As shown in Figs. 2 and 3, the system and model, represented by contour plots are similar. An additional check as to the validity of the model is to randomly sample points of the objective function at points not used to create the model function, then calculate the error between the two. Consider the cubic function, with $c = 0$; we used 49, 121, and 225 sampled data points to create three models. We then computed the square root of the squared difference between the cubic model value and the “actual” peaks function value at each of the 2601 data points. For the model created with 49 sampled data points, the cumulative error was 1310, or an average of .5037 per each of the 2601 points. For the model created with 121 sampled data points, the cumulative error was 181, or an average of .0696 per each of the 2601 points. For the model created with 225 sampled data points, the cumulative error was 77.1, or an average of .0296 per each of the 2601 points.

3.2. RSM model information

Once a satisfactory model has been created, several types of information can be obtained from it. This includes information regarding function evaluations, analytic gradients, and the Hessian matrix of second partial derivatives. Each radial basis function must be evaluated separately to determine in what cases the functions are differentiable and twice differentiable. The cases include, for what values of the constant c are the functions “good”, and also the nature of the functions themselves in terms of differentiability. This information is used to compare the models with the original system (objective function evaluations), as well as each other (function, gradient, Hessian evaluations). The relative merits of one or more models can be shown not only in comparison with the true response but with corroboration between the models themselves. For a given scattered data set, we may create several response surface models by selecting different radial basis functions and/or choosing different values for the parameter c . When two or more distinct models predict extrema in

Fig. 2. Peaks function ($51 \times 51 = 2601$ plot points).

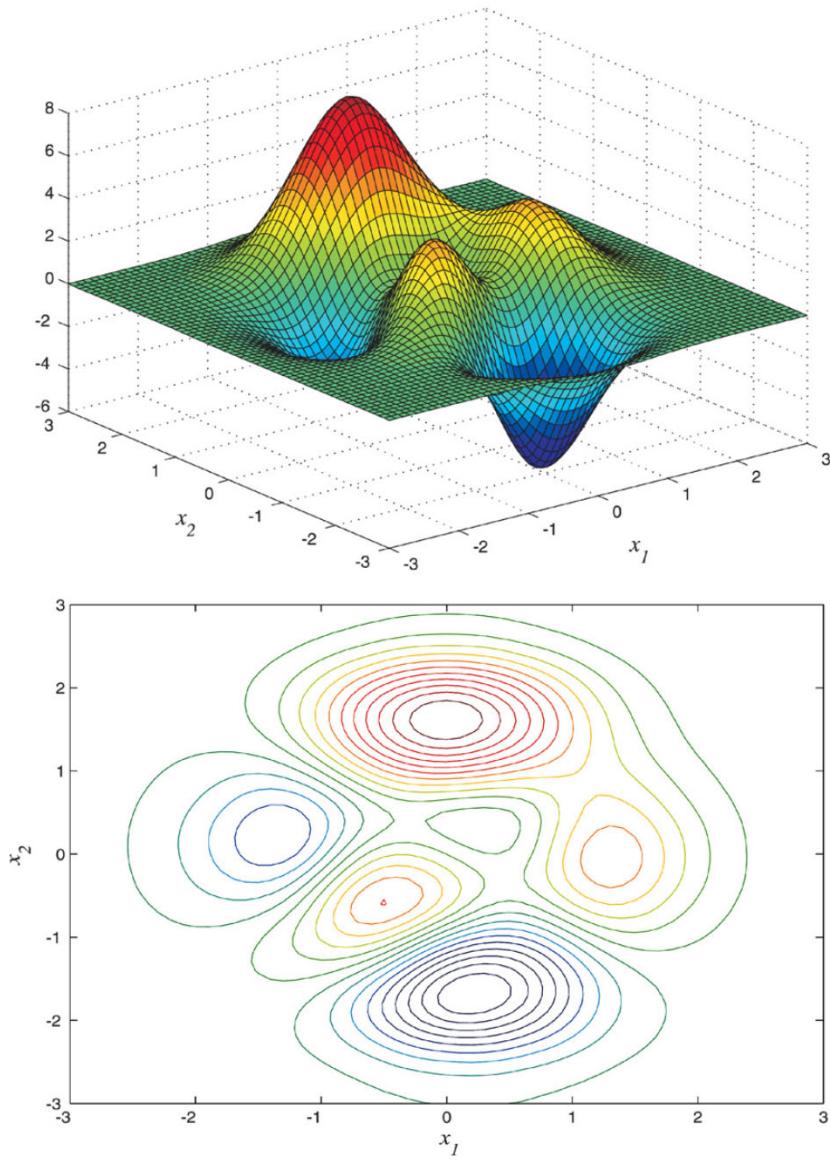
approximately the same location, we not only have a verification of the individual models themselves, but also a consensus on the locations of candidate optimal design points.

For an RSM model based on a particular radial basis function $\phi(r)$, gradients can be computed analytically from (11) and (6) as

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \boldsymbol{\beta}^\top \frac{\partial \mathbf{g}}{\partial \mathbf{x}} = \sum_{i=1}^m \beta_i \phi'(r_i) \frac{\partial r_i}{\partial \mathbf{x}}, \quad (17)$$

where $\phi'(r) = \partial \phi / \partial r$,

$$r_i(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_i\| = \sqrt{(\mathbf{x} - \mathbf{x}_i)^\top (\mathbf{x} - \mathbf{x}_i)}$$

Fig. 3. Cubic RSM model ($11 \times 11 = 121$ plot points).

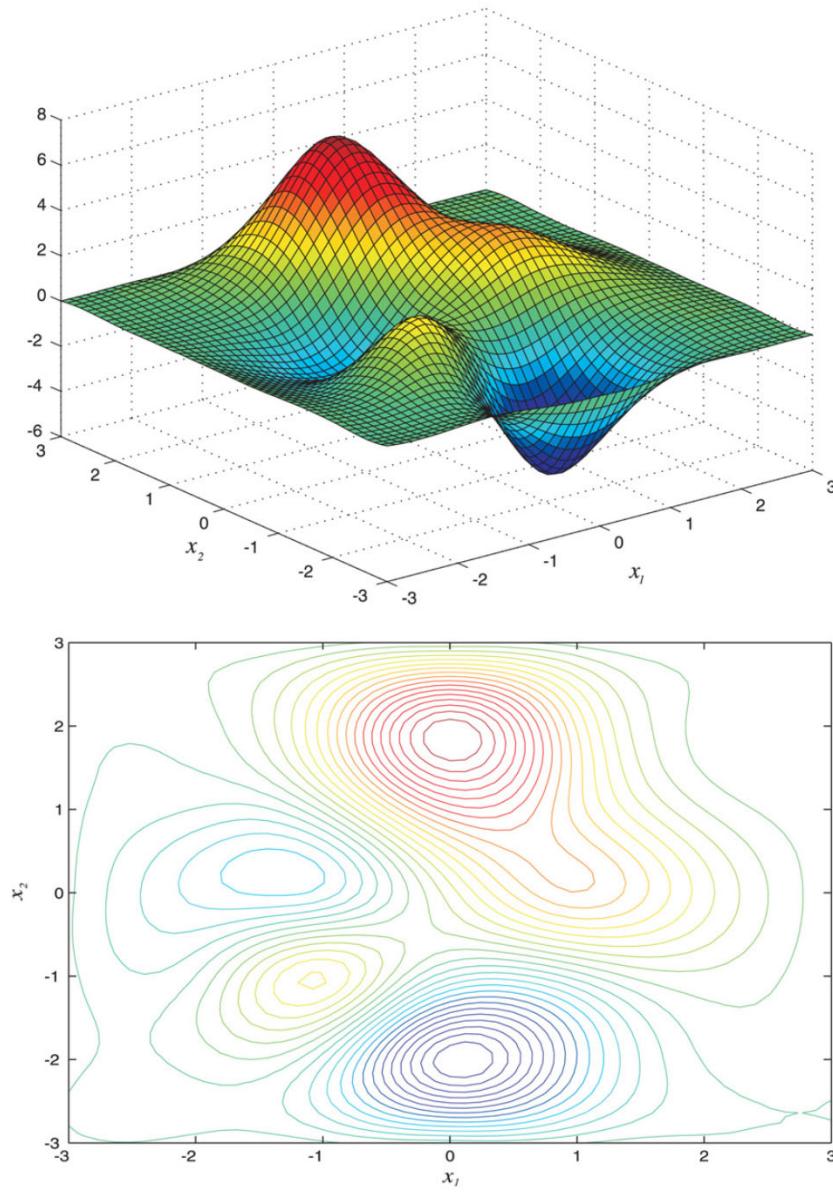
and

$$\frac{\partial r_i}{\partial \mathbf{x}} = \frac{1}{r_i(\mathbf{x})} (\mathbf{x} - \mathbf{x}_i)^\top. \quad (18)$$

Thus

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \sum_{i=1}^m \frac{\beta_i \phi'(r_i)}{r_i(\mathbf{x})} (\mathbf{x} - \mathbf{x}_i)^\top. \quad (19)$$

Table 1 contains formulas for $\phi'(r)$ and $\phi''(r)$ for the radial basis functions in (14). In this paper the symbols ϕ' and ϕ'' imply differentiation of ϕ with respect to r . Later we will differentiate ϕ (and its derivatives with respect to r) with respect to c and this operation will be denoted, for example, by $\partial\phi/\partial c$.

Fig. 4. Cubic RSM model ($7 \times 7 = 49$ data points).

We wish to examine the existence and continuity of partial derivatives of $f(\mathbf{x})$, including the limiting case where \mathbf{x} may be at a data point \mathbf{x}_j for some $j \in \{1, \dots, m\}$, yielding $r_j(\mathbf{x}_j) = 0$ in the denominator of (19). Consider the following form of the gradient:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \sum_{\substack{i=1 \\ i \neq j}}^m \left(\frac{\beta_i \phi'(r_i)}{r_i(\mathbf{x})} (\mathbf{x} - \mathbf{x}_i)^\top \right) + \frac{\beta_j \phi'(r_j)}{r_j(\mathbf{x})} (\mathbf{x} - \mathbf{x}_j)^\top, \quad (20)$$

where we are interested in the gradient at a data point \mathbf{x}_j . Let $\mathbf{x} = \mathbf{x}_j + \alpha \boldsymbol{\eta}$, where α is an arbitrary scaling factor, and $\boldsymbol{\eta}$ is any unit vector. The radius r_j is then

$$r_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_j\| = |\alpha| \sqrt{\boldsymbol{\eta}^\top \boldsymbol{\eta}} = |\alpha|.$$

Table 1
Radial basis functions and derivatives

#	$\phi(r)$	$\phi'(r)$	$\phi''(r)$
1	$(r+c)^3$	$3(r+c)^2$	$6(r+c)$
2	$r^2 \ln(cr)$	$2r \ln(cr) + r$	$2\ln(cr) + 3$
3	$e^{(-cr^2)}$	$-2cre^{(-cr^2)}$	$[4c^2r^2 - 2c]e^{(-cr^2)}$
4	$\frac{1}{\sqrt{r^2 + c^2}}$	$\frac{-r}{(r^2 + c^2)^{3/2}}$	$\frac{2r^2 - c^2}{(r^2 + c^2)^{5/2}}$
5	$\frac{1}{\sqrt{r^2 + c^2}}$	$\frac{r}{(r^2 + c^2)^{3/2}}$	$\frac{c^2}{(r^2 + c^2)^{5/2}}$

Let the term of interest within (20) be denoted as

$$\mathbf{h}_j^\top = \frac{\beta_j \phi'(r_j)}{r_j(\mathbf{x})} (\mathbf{x} - \mathbf{x}_j)^\top.$$

Using the above substitutions for \mathbf{x} and α , we have

$$\mathbf{h}_j = \frac{\beta_j \phi'(r_j)}{|\alpha|} (\mathbf{x}_j + \alpha \boldsymbol{\eta} - \mathbf{x}_j)^\top = \frac{\beta_j \phi'(r_j)}{|\alpha|} \alpha \boldsymbol{\eta}^\top = \beta_j \phi'(r_j) \operatorname{sgn}(\alpha) \boldsymbol{\eta}^\top.$$

The existence and continuity of partial derivatives of $f(\mathbf{x})$ depends upon this term, in the (right- and left-hand) limit as $\alpha \rightarrow 0$ and with $r_j(\mathbf{x}) \rightarrow 0$. With α greater than zero and decreasing we have,

$$\lim_{\alpha \downarrow 0} (\mathbf{h}_j) = \beta_j \phi'(0) \boldsymbol{\eta}^\top.$$

With $\alpha < 0$ and increasing we have

$$\lim_{\alpha \uparrow 0} (\mathbf{h}_j) = -\beta_j \phi'(0) \boldsymbol{\eta}^\top.$$

The right- and left-hand limits must converge to the same value for this limit to exist and this can happen only if $\phi'(0) = 0$. From this, we have a condition that determines whether the gradient at a data point exists and is continuous at a data point \mathbf{x}_j . Namely, for any basis function and choice of c , the derivative with respect to r evaluated at $r = 0$ must equal zero. Each basis function derivative in Table 1 must be examined to verify where this condition holds. The cubic function derivative must have a c value of zero for its gradient to exist at a data point. The gradient for the thin plate spline function is well defined for all choices of $c > 0$. The gradient of the Gaussian function is well defined for all choices of c . The multiquadric and inverse multiquadric gradients are well defined for all choices of $c \neq 0$. Clearly, the radial basis function models are continuously differentiable away from sampled data points.

The Hessian matrix can be calculated from (19) as

$$\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^2} = \sum_{i=1}^m \frac{\beta_i}{r_i(\mathbf{x})} \left\{ \phi'(r_i) \mathbf{I} + \left[\phi''(r_i) - \frac{\phi'(r_i)}{r_i(\mathbf{x})} \right] (\mathbf{x} - \mathbf{x}_i) \frac{\partial r_i}{\partial \mathbf{x}} \right\}. \quad (21)$$

All models created with the basis functions in (14) are twice continuously differentiable away from the sampled data points. We must examine the existence and continuity of the second partials at the individual data points. At a data point \mathbf{x}_j the corresponding term in (21) is

$$\mathbf{H}_j = \frac{\beta_j}{r_j(\mathbf{x})} \left\{ \phi'(r_j) \mathbf{I} + \left[\phi''(r_j) - \frac{\phi'(r_j)}{r_j(\mathbf{x})} \right] (\mathbf{x} - \mathbf{x}_j) \frac{\partial r_j}{\partial \mathbf{x}} \right\}. \quad (22)$$

With $\partial r_i / \partial \mathbf{x}$ as given in (17), (22) becomes

$$\mathbf{H}_j = \frac{\beta_j}{r_j(\mathbf{x})} \left\{ \phi'(r_j) \mathbf{I} + \left[\phi''(r_j) - \frac{\phi'(r_j)}{r_j(\mathbf{x})} \right] \frac{(\mathbf{x} - \mathbf{x}_j)(\mathbf{x} - \mathbf{x}_j)^\top}{r_j(\mathbf{x})} \right\}. \quad (23)$$

Now we let $\mathbf{x} = \mathbf{x}_j + \alpha\eta$ and with $\|\eta\| = 1$, then $r_j(\mathbf{x}) = \sqrt{\alpha^2\eta\eta^\top} = |\alpha|$. Substitution of these terms (and with $\phi'(0) = 0$) into (23) results in

$$\mathbf{H}_j = \beta_j \left\{ \frac{\phi'(|\alpha|)}{|\alpha|} [\mathbf{I} - \eta\eta^\top] + \phi''(|\alpha|)\eta\eta^\top \right\}. \quad (24)$$

The use of L'Hopital's rule on the appropriate terms in (24) yields

$$\lim_{|\alpha| \rightarrow 0} \mathbf{H}_j = \beta_j \phi''(0) \mathbf{I}.$$

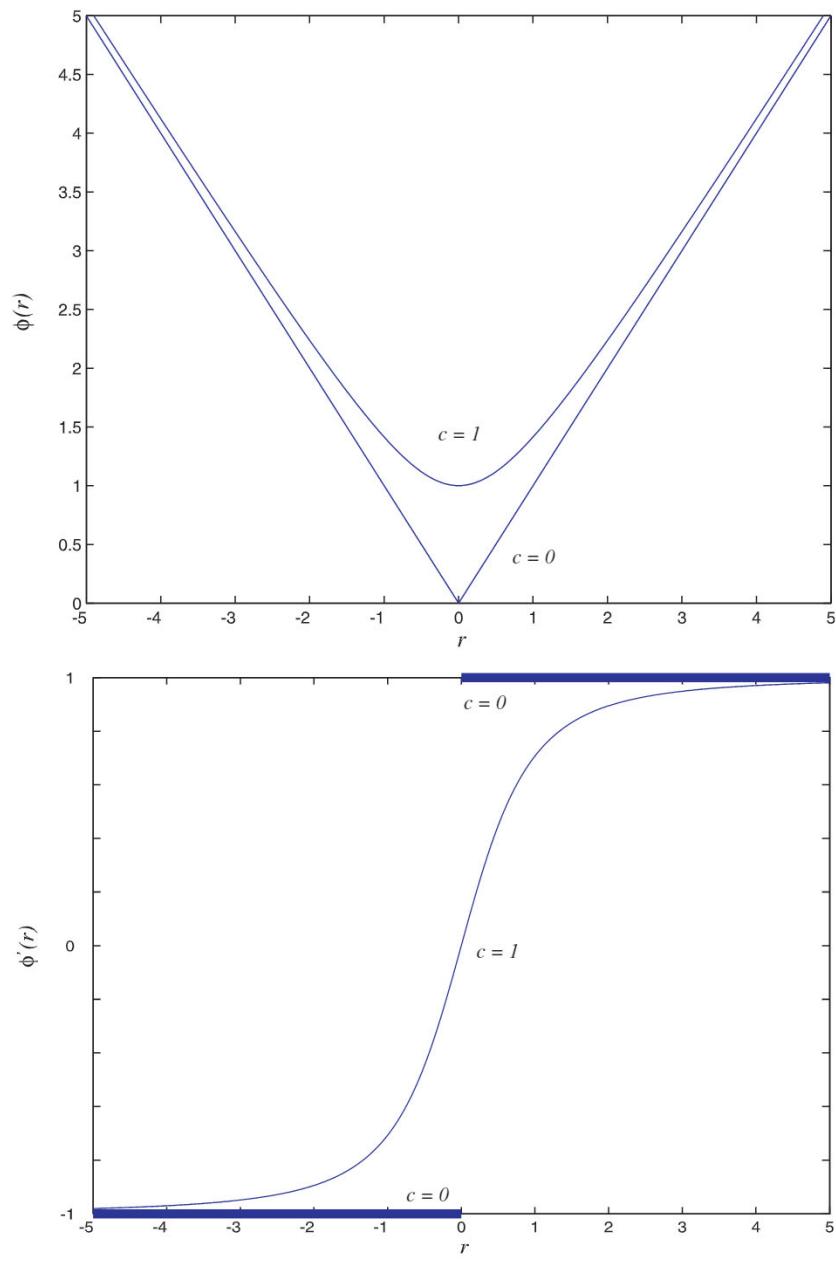


Fig. 5. Multiquadric function.

For a model to exist at a data point \mathbf{x}_j , $\phi'(0) = 0$ as before, and $\phi''(0)$ must exist. For the cubic function $\phi''(0)$ exists for all choices of c . For the thin plate spline function $\phi''(0)$ does not exist for any choice of c . For the Gaussian function $\phi''(0)$ exists for all choices of c . For the inverse multiquadric and multiquadric functions, $\phi''(0)$ exists for all $c \neq 0$. Fig. 5 illustrates how the choice of c influences differentiability properties for the multiquadric function.

4. Shaped charge oil well penetrator

The physical problem that motivated our study or response surface modelling is concerned with designing an explosively-formed oil well penetrator using simulation data provided by Lawrence Livermore National Laboratory. The system consists of a “shaped charge”, which is a high explosive charge with a lined cavity [1]. The liner of this cavity is typically a hollow metal cone or wedge, as illustrated in Fig. 6. The explosive charge is used to collapse the liner in such a manner that a high velocity jet is produced from a collapse point. The jet typically stretches to several times its original length before breaking into smaller pieces. The jet is very useful for its penetration properties, and this property is the principal focus of the design analysis. The jet impacts a wall surface, which is copper in our case, at a given distance from the original position of the shaped

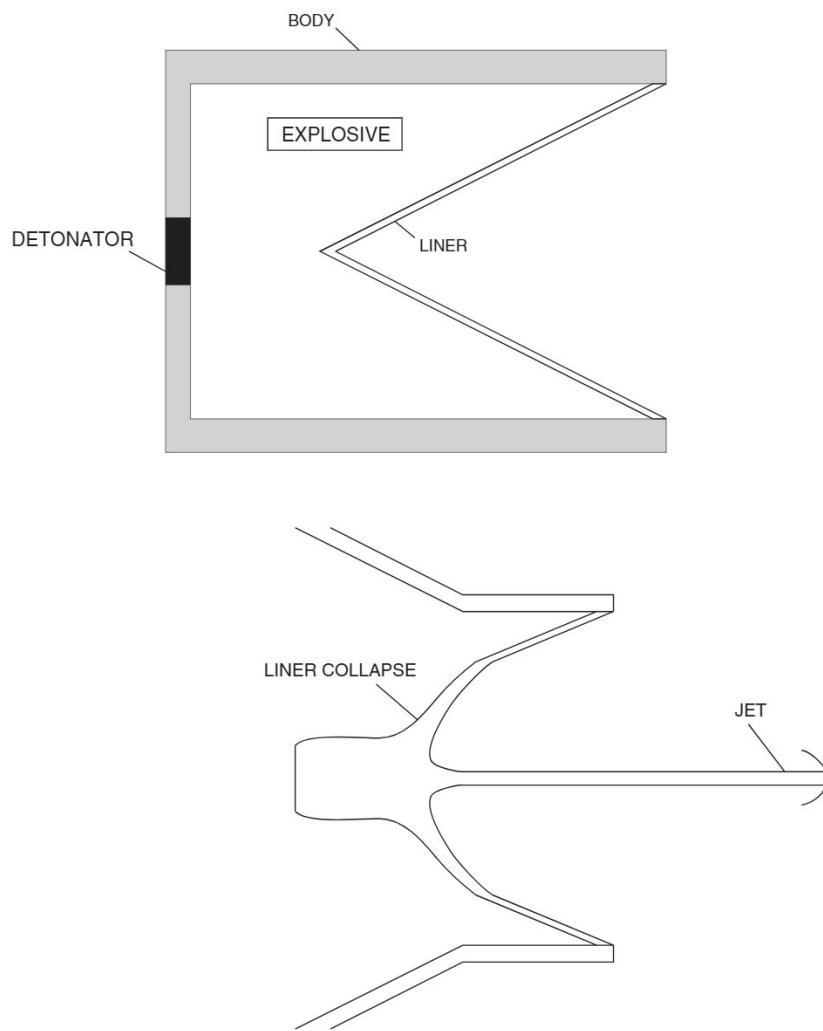


Fig. 6. Shaped charge explosive and jet formation.

charge, and creates a “hole” in the wall surface that is then studied. The resulting hole is the focus of the data analysis and its properties are used as performance measures to evaluate the effectiveness of the design. Fig. 6, adapted from [1], illustrates the initial shaped-charge configuration and shows the liner collapse and the jet formation. For the system that we consider there are seven design parameters, consisting of various geometric parameters of the shaped charge (radii and thicknesses at certain points along the liner, etc.).

As the jet impacts the wall surface the quality of the resulting hole is a measure of the success of the design. Several criteria have been investigated to judge the quality of the hole. Typically, the overall volume of the hole is used as a performance measure. The combination of design parameters that deliver the greatest volume might be deemed of greatest importance in some applications. However, alternate performance measures may be preferred. For example, the volume of a cylindrical shaft of maximum radius that would fit into the hole to 90% of the hole’s depth may be a performance measure. Preservation of a radius of a given magnitude may be vital for certain applications.

Regardless of the specific performance measure, the complexity of the system requires costly experimental or numerical analysis. The simulation approach used in this paper consists of finite element and finite difference programs to integrate the conservation equations (mass, momentum, energy, species concentrations, etc.) during explosive formation of the shaped-charge penetrator. From this, a continuum solution is obtained for the explosive detonation, shaped-charge liner collapse, and the jet formation [1].

4.1. Hydrocode simulation

A significant problem with this type of analysis is the amount of time needed to complete such a simulation. The data in our problem was generated by the use of two platforms; CALE and DYNA. CALE is an Arbitrary Lagrangian Eulerian hydrodynamics computer program written in the C programming language [9]. DYNA is an advanced finite element software program that is well suited for the analysis of nonlinear systems capable of design optimization, multi-physics coupling, and adaptive re-meshing. The data collected at Lawrence Livermore National Laboratory consisted of a figure of merit objective function evaluated at 480 data points, composed of seven parameters (variables). Approximately 48 h CPU simulation time was required to produce this data set.

4.2. Model development

The RSM Response Surface Model Fortran module was used to create eight response surface models using the radial basis function method. Once each model was created, a steepest ascent algorithm was used to locate the local maximum points. For each model the algorithm was run 128 times, beginning at each point of a 128 point grid ($x_i = -0.5$ or $x_i = 0.5$, $i = 1, \dots, 7$) so as to locate as many relative maxima as possible. From these models, we obtained candidate optimization points that were sent to LLNL where these optimal parameters were used for hydrocode simulation runs. These eight models and their objective function global maximum approximation values are shown in Table 2.

Originally, the optimization strategy was to use sequential quadratic programming [10,11]. This algorithm did not perform well on the response surface models for several reasons. For a function with multiple extrema

Table 2
RSM objective function maxima

#	Basis function	c	$\phi(r)$	f
1	Cubic	0	r^3	1562.0
2	Thin plate spline	1	$r^2 \ln r$	1548.9
3	Gaussian	0.8	$e^{(-0.8r^2)}$	1516.8
4	Gaussian	1	$e^{(-r^2)}$	1496.0
5	Inverse multiquadric	0.1	$(r^2 + .001)^{-1/2}$	1477.8
6	Inverse multiquadric	1	$(r^2 + 1)^{-1/2}$	1496.3
7	Multiquadric	0.1	$(r^2 + .01)^{1/2}$	1487.5
8	Multiquadric	1	$(r^2 + 1)^{1/2}$	1585.4

the quadratic approximation may not be accurate enough unless you are sufficiently close to an extreme point. We could not be sure that relative maxima would not be missed using this method. As an example, consider a surface which consists of a very flat valley floor with spikes generated from a rapidly decaying exponential function rising from the floor. A poor quadratic approximation may travel through the valley and miss some or all spikes if they are not in close proximity with them. The resulting list of global maximum candidates found by the RSM Fortran module was reduced by discarding all local maxima for which the model objective function was less than the maximum objective function value found in the data set. These global maximum candidates were then sent to LLNL, where the parameters of the hydrocode simulation were adjusted to match the RSM candidate values.

Four hydrocode simulation runs were carried out using the eight RBF models from Table 2 and three models developed by LLNL (LL1, LL2, LL3). The LL1 model was created using the parameters that produced the largest objective function value in the LLNL data set. The LL2 model and the LL3 model were developed independently by LLNL to re-evaluate and compare to the RSM models. These three models served as benchmarks to gauge the performance of the RSM models. Four simulation runs LLS1, LLS2, LLS3, and LLS4 were carried out to assess the performance of the eight RSM models and the three LLNL models. The objective function for LLS1 was based on the volume of a cylindrical rod whose length was 90% of the hole depth

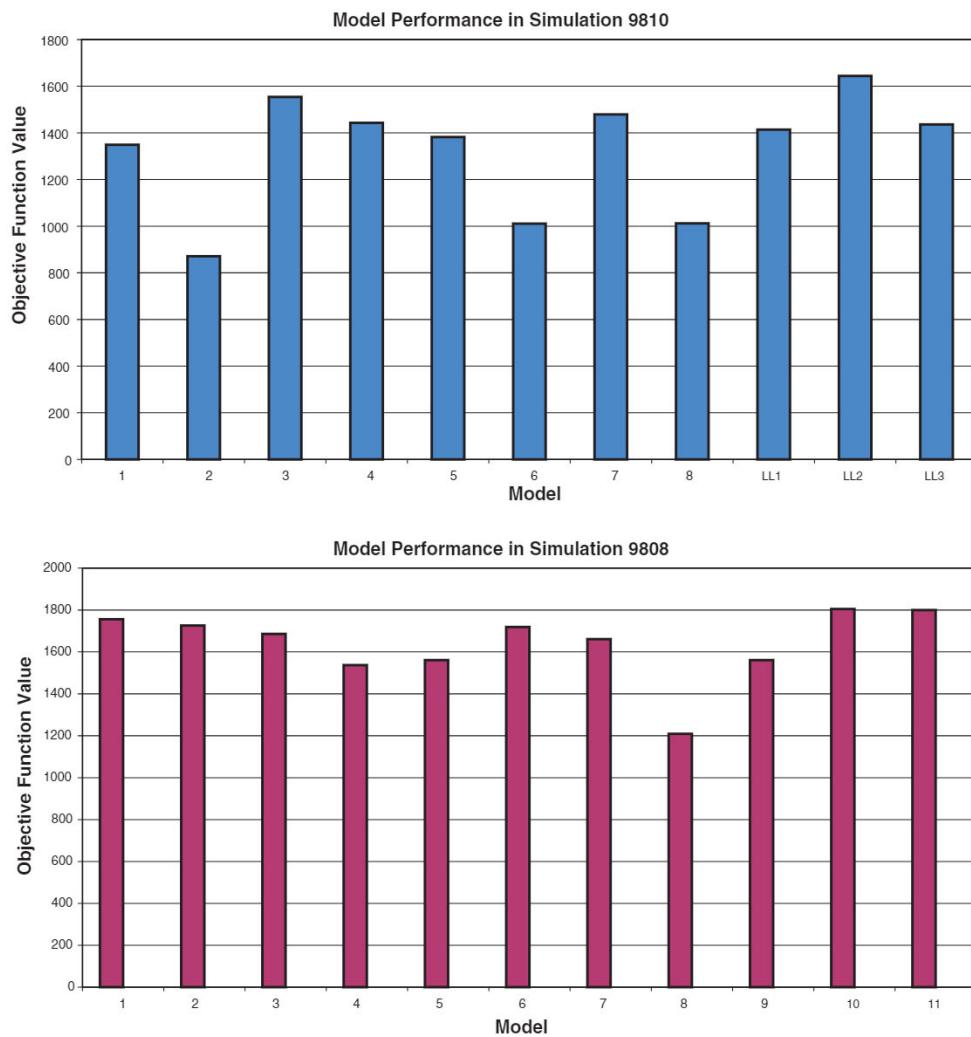


Fig. 7. Model performance in simulation.

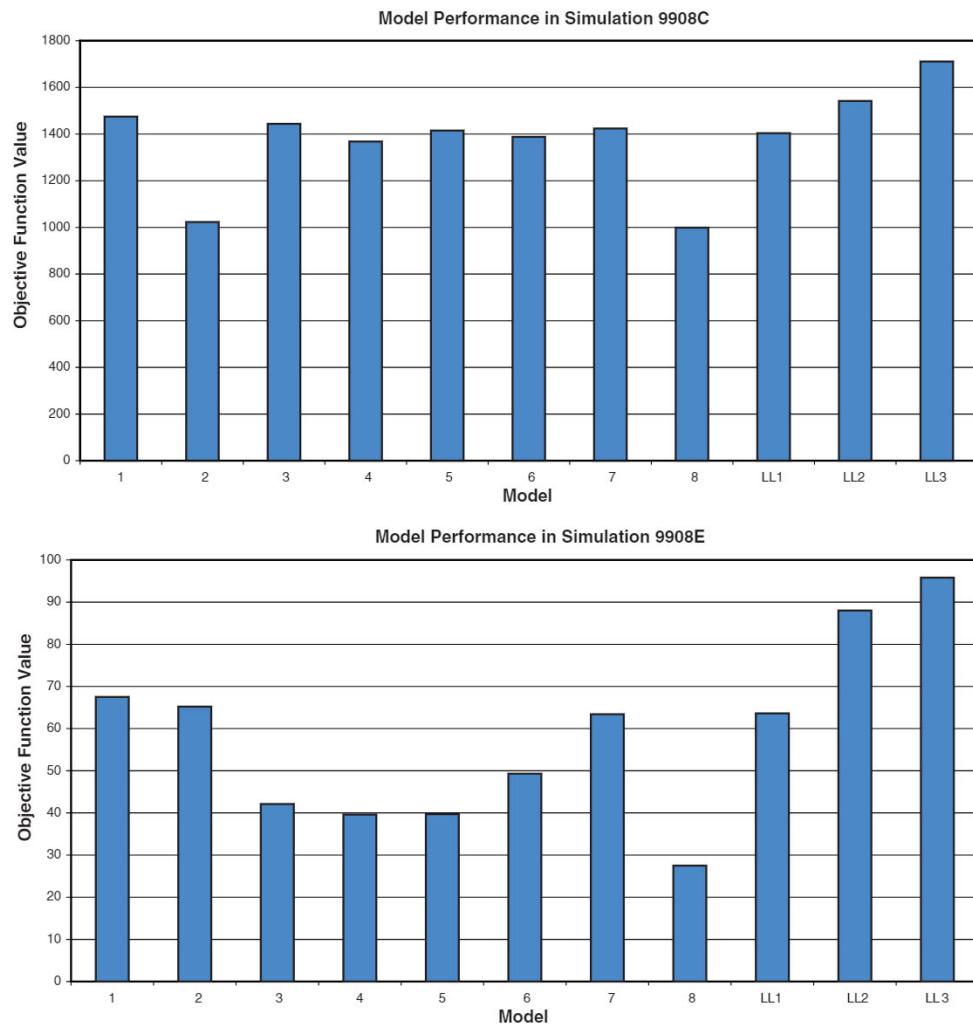


Fig. 8. Model performance in simulation.

and whose diameter was that of the minimum hole cross-sectional diameter. The objective function for LLS2 used volume based on the average hole cross-sectional area. The objective function for LLS3 used the volume of a cylinder whose depth corresponded to the 25–95% depth region of the hole and whose diameter was the minimum hole cross-sectional area. The objective function for the LLS4 simulation was the depth of a 4.7 cm diameter rod inserted into the hole as deeply as it would fit.

Of the RSM models, the Gaussian ($c = 0.8$, Model 3) and the multiquadric ($c = 0.1$, Model 7) performed best in the LLS1 simulation. The cubic ($c = 0.0$, Model 1), the other Gaussian ($c = 1$, Model 4), and the inverse multiquadric ($c = 0.1$, Model 5) also performed well. Overall, a 10% improvement in the objective function was realized with the use of the optimized parameter values obtained from these models. Figs. 7 and 8 graphically display these results.

5. Conclusion

Many of the most interesting scientific and engineering applications, such as those involving multiple time- and length-scales are very complex and difficult to analyze. The processing times for data collection and analysis can be quite lengthy; repeated use of such a simulation for optimization purposes is not practical.

Creation of a response surface model using radial basis functions allows for timely system analysis and optimization; simulation time may be reduced from days to minutes. Limited information collected from the complex system may be used to generate a model; this model may then be easily studied. Information gained in this process may be re-applied to the original system, as was done in this paper. Valuable information such as model function evaluations, gradient evaluations, and Hessian matrices can be computed easily. Further analysis, such as sensitivity to a particular parameter could also be accomplished in an efficient manner.

The value of response surface modelling using radial basis functions can be quite significant. A 10% improvement resulted when such methods were used to guide the optimization of a complex shaped-charge explosive oil well penetrator. Response surface models provide a powerful tool to simulate, analyze, and optimize complex systems in a cost-effective manner.

References

- [1] E.L. Baker, Modelling and Optimization of Shaped Charge Liner Collapse and Jet Formation, Ph.D. Dissertation, Washington State University, 1992.
- [2] R.H. Myers, D.H. Montgomery, *Response Surface Methodology*, Wiley, New York, 1995.
- [3] M.J.D. Powell, The theory of radial basis function approximation in 1990, in: W. Light (Ed.), *Numerical Analysis II, Wavelets, Subdivision, and Radial Basis Functions*, Oxford University Press, Oxford, 1992, pp. 105–210.
- [4] G.E.P. Box, N.R. Draper, *Empirical Model Building and Response Surfaces*, Wiley, New York, 1987.
- [5] T.L. Vincent, W.J. Grantham, *Optimality in Parametric Systems*, Wiley, New York, 1981.
- [6] D.S. Watkins, *Fundamentals of Matrix Computations*, Wiley, New York, 1991.
- [7] R.L. Hardy, Theory and applications of the multiquadric–biharmonic method, *Computers and Mathematics with Applications* 19 (1990) 163–208.
- [8] Press et al., *Numerical Recipes, the Art of Scientific Computing*, Cambridge University Press, New York, 1986.
- [9] R.E. Tipton, CALE Users Manual, Lawrence Livermore National Laboratory, Version 980201.
- [10] P.E. Gill, W. Murray, M.H. Wright, *Practical Optimization*, Academic Press, New York, 1981.
- [11] T.L. Vincent, W.J. Grantham, *Nonlinear and Optimal Control Systems*, Wiley, New York, 1997.