Struct Multidisc Optim (2010) 42:811–821
DOI 10.1007/s00158-010-0541-7

RESEARCH PAPER

Estimating training data boundaries in surrogate-based modeling

Luis E. Pineda · Benjamin J. Fregly · Raphael T. Haftka · Nestor V. Queipo

Received: 3 September 2009 / Revised: 6 May 2010 / Accepted: 30 June 2010 / Published online: 27 July 2010

Abstract Using surrogate models outside training data boundaries can be risky and subject to significant errors. This paper presents a computationally efficient approach to estimate the boundaries of training data inputs in surrogate modeling using the Mahalanobis distance (MD). This distance can then be used as a threshold for deciding whether or not a particular prediction site is within the boundaries of the training data inputs, and has the potential of a likelihood/probabilistic interpretation. The approach is evaluated using two and four dimensional analytical restricted input spaces and a complex biomechanical six dimensional problem. The proposed approach: i) gives good approximations for the boundaries of the restricted input spaces, ii) exhibits reasonable error rates when classifying prediction sites as inside or outside known restricted input spaces and iii) reflects expected error trends for increasing values of the MDs similar to those obtained using a computationally expensive convex hull approach.

Keywords Surrogate modeling · Training data boundaries · Mahalanobis distance

Part of this work was presented at the 8th World Congress on Structural and Multidisciplinary Optimization, June 1–5, 2009, Lisbon, Portugal.

L. E. Pineda · N. V. Queipo (✉)
Applied Computing Institute, University of Zulia, Maracaibo, Venezuela
e-mail: nqueipo@ica.luz.ve

L. E. Pineda
e-mail: lpineda@ica.luz.ve

B. J. Fregly · R. T. Haftka
Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA

B. J. Fregly
e-mail: fregly@ufl.edu

R. T. Haftka
e-mail: haftka@ufl.edu

Nomenclature

BER	Balanced error rate
C	Covariance matrix
KS	Kolmogorov–Smirnov
LHS	Latin hypercube sampling
m	Number of training data
MD	Mahalanobis distance
n	Number of input variables
p	Probability of a prediction site being within the training data boundaries
R^p	Set of real numbers of dimension p
S	Surrogate model
T	Training data
x	Input variables
y	Response variables
α	Statistical significance level
χ^2_p	Chi-square distribution—p degrees of freedom
Δ	Difference
ε	Relative error
μ	Mean

Subindices

b	boundary estimation
$b20$	median of top 20% largest Mahalanobis distances
bl	largest Mahalanobis distance
T	training data

1 Introduction

Surrogate models, increasingly popular in the analysis and optimization of complex engineering systems (e.g., Queipo et al. 2005; Forrester and Keane 2009), are often constructed using training data in restricted input spaces; for example, as a result of correlation between input variables, physical limitations, or lack of information (e.g., Jacques et al. 2006; Lin et al. 2008). When using these models in the context of analysis and optimization, the modeler must decide whether or not it is reasonable to use the surrogate model at a given prediction site. A key consideration is whether using the prediction site would constitute extrapolation beyond the region defined by the training data inputs.

While there is no agreed upon definition for surrogate model extrapolation (i.e., the prediction site is outside the training data boundary) as opposed to interpolation, an appealing concept (see, for example, Missoum et al. 2007) is to consider extrapolation when prediction sites are outside the convex hull of the training data inputs. However, constructing a convex hull for a finite set of points—a computational geometry problem—can be complex and computationally expensive for high dimensional problems (Barber et al. 1996; Mount 2002). More specifically, the number of memory references can make impractical one of the best convex hull computation algorithms, i.e., Quickhull (Barber et al. 1996), even for mid-size problems.[1] On the other hand, the minimum volume ellipsoid covering the training data inputs is a scale-invariant alternative to the convex hull approach but can be sensitive to outliers and requires the solution of a mixed-integer semidefinite programming problem (Shioda and Tunçel 2007; Sun and Freund 2004).

The proposed approach estimates the training data boundaries using the Mahalanobis distance (MD), which can be used as a threshold for deciding whether or not a particular prediction site is consistent with the training data. The Mahalanobis distance is a scale-invariant metric widely used in cluster analysis and other classification techniques (Mardia et al. 1979) and can prove to be useful for estimating the training data boundaries since it: i) accounts for the covariance structure of the training data inputs, ii) can be easily extended to high dimensional problems as it essentially relies on matrix operations which can be computed quite efficiently, and iii) assuming a normal distribution of the training data inputs, it has a likelihood/probability interpretation. The effectiveness of the proposed approach is evaluated using different restricted input spaces and training sample sizes for two analytical test cases and a six dimensional biomechanical problem (i.e., contact force in a knee replacement).

The rest of the paper is outlined as follows: Section 2 describes the problem of interest, Section 3 explains the solution methodology, Section 4 discusses test cases including performance measures and extrapolation error trends, results and discussion is the subject of Section 5, and Section 6 presents the conclusions.

2 Problem of interest

Given a surrogate model S constructed using training data $T = ((\mathbf{x}_j, y_j) : 1 < j < m)$ for a function $y = f(\mathbf{x})$ defined in R^p, the goal is to find a distance to the training data boundaries to help decide whether a given prediction site $\mathbf{x}_k$ would constitute an extrapolation of S. The distance should be easy to compute and applicable to general surrogate modeling scenarios (S, T, f).

3 Solution methodology

The proposed methodology includes the following three steps: 1) estimate the training data boundaries using the Mahalanobis distance, 2) assess the consistency of prediction sites with the training data, and 3) give a probability interpretation of the MDs of prediction sites (if applicable).

3.1 Estimate the training data boundaries

In this context, Mahalanobis distance is a scale-invariant proximity measure between a given point $\mathbf{x}_l$ and the center of the training data inputs that accounts for their covariance structure. This distance measure is defined as:

$$MD(\mathbf{x}_l) = \sqrt{(\mathbf{x}_l - \mu_T)^T \, C^{-1} \, (\mathbf{x}_l - \mu_T)} \tag{1}$$

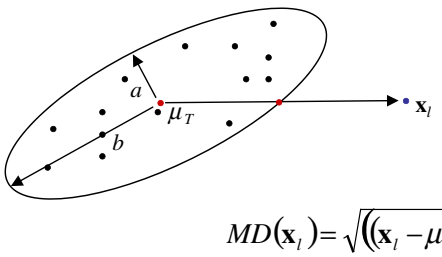

Fig. 1 Two dimensional Mahalanobis distance interpretation for non-diagonal covariance matrices, where the directions and length of the principal axes (a and b) of the ellipse are given by the eigenvectors and the square root of the eigenvalues of the covariance matrix C, respectively

[1] For example, the Matlab implementation of the Quickhull algorithm (convhulln) was unable to compute the convex hull associated with two hundred (200) training data in a ten (10) dimensional hyper-spherical restricted input space due to lack of memory, when using a computer with a 2.5 GHz Pentium IV processor and 2GB/5GB of RAM/virtual memory.

Table 1 Restricted input spaces characterization

Restricted input space	Number of training data	Covariance matrix of training data inputs (most dense sample)	Number of test data
Circle	15,30,45	$\begin{bmatrix} 3.72 & 0.62 \\ 0.62 & 4.56 \end{bmatrix}$	250
Ellipse	15,30,45	$\begin{bmatrix} 5.24 & -2.27 \\ -2.27 & 2.67 \end{bmatrix}$	250
Hyper-sphere	40,80,160	$\begin{bmatrix} 4.26 & -0.09 & 0.16 & 0.10 \\ -0.09 & 3.93 & -0.07 & 0.64 \\ 0.16 & -0.07 & 5.07 & -0.18 \\ 0.10 & -0.64 & -0.18 & 3.64 \end{bmatrix}$	1,000
Hyper-ellipsoid	40,80,160	$\begin{bmatrix} 4.45 & -2.92 & 0.31 & -0.06 \\ -2.92 & 2.62 & -0.29 & 0.03 \\ 0.31 & -0.29 & 4.22 & -2.59 \\ -0.06 & 0.03 & -2.59 & 6.05 \end{bmatrix}$	1,000

Analytical test cases

where μ_T and C are the mean and covariance matrix of the training data inputs, respectively. The mean and covariance matrix are calculated using standard formulas:

$$\mu_{Ti} = \frac{1}{m}\sum_{k=1}^{m} x_{k_i}, \quad i = 1, 2, ..., n \tag{2}$$

$$C_{ij} = \frac{1}{m-1}\sum_{k=1}^{m}\left(x_{k_i} - \mu_{Ti}\right)\left(x_{k_j} - \mu_{Tj}\right), \quad i, j = 1, 2, ..., n \tag{3}$$

where m represents the number of training data and n denotes the number of input variables, respectively. Note that the Mahalanobis and Euclidean distances coincide when the covariance is represented by the identity matrix.

Mahalanobis distance can be interpreted as follows. If the training data inputs have associated a diagonal covariance matrix, the MD along the x axis represents a normalized Euclidean distance given by $(x - \mu_T)/\sigma(x)$, where μ_T and $\sigma(x)$ are the mean and standard deviation of the training data inputs along the x direction, respectively. In the more general case of non diagonal covariance matrices, any MD can be shown to be equal to the ratio of: a) the distance between the point whose MD is sought (henceforth called prediction site) and μ_T, and b) the width of the ellipsoid/hyper-ellipsoid (measured from μ_T) in the direction of the prediction site (see Fig. 1). The principal axes and the widths of the ellipsoid/hyper-ellipsoid along these axes are given by the eigenvectors of the covariance matrix C and the squared root of the corresponding eigenvalues (i.e., the eigenvalues represent the variance along principal axes/directions), respectively. A principal component analysis of the data would reveal this fact since it essentially linearly transforms the data such that the covariance matrix C (for the new set of variables) becomes diagonal.

The training data boundaries are estimated using two alternative MDs denoted as MD_b: i) the largest MD (MD_{bl}), and ii) the median of the top 20% largest MDs (MD_{b20}),

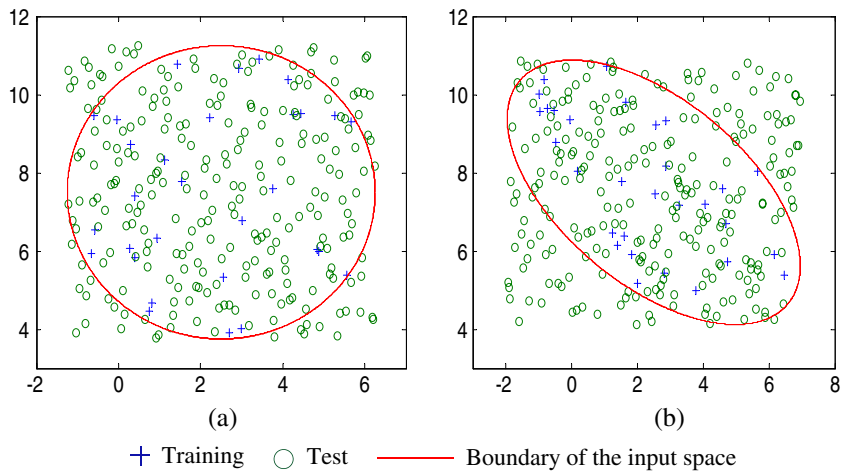

Fig. 2 Circular (**a**) and elliptical (**b**) two dimensional restricted input spaces, training (*plus sign*) and test data (*circle*)—30 training data

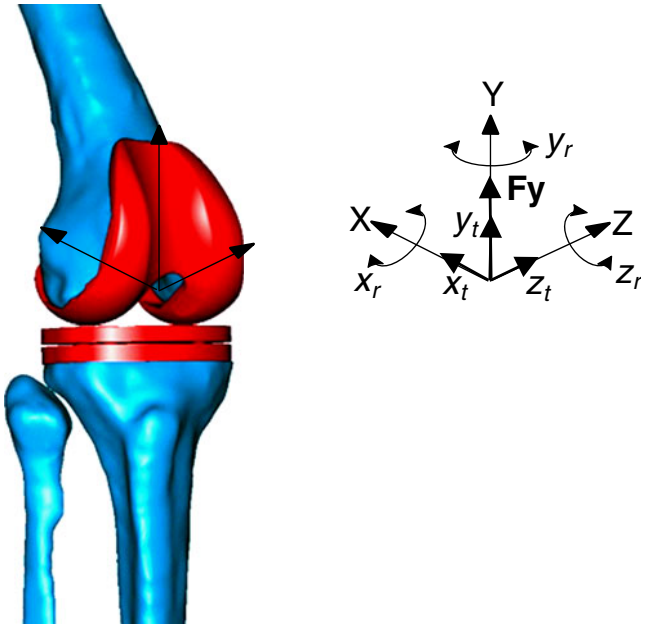

Fig. 3 Superior–inferior contact force and corresponding input variables for the biomechanical test case; the knee replacement is shown in *red*. F_y—superior–inferior contact force, *xt*—tibial anterior–posterior translation, *yt*—femoral superior–inferior translation, *zt*—tibial medial–lateral translation, *xr*—femoral varus–valgus rotation, *yr*—tibial internal–external rotation, and *zr*—femoral flexion–extension

which is less sensitive to outliers in the training data. The 20% percent threshold was empirically selected to provide a robust estimation of the training data boundaries; any other value though (e.g., top 10% MDs) could be used if it meets the robustness objective.

3.2 Assessing the consistency of prediction sites with the training data

To assess this consistency, the MD corresponding to the prediction site is computed and compared to the estimated MD for the training data boundaries (i.e., either MD_{b20} or MD_{bl}). If the MD for the prediction site is larger than the selected MD_b, then the use of the surrogate model is likely to represent extrapolation and the prediction could be at a significant error. The designer must decide—depending on the size of the mismatch—whether or not to use the prediction site in the context of surrogate-based analysis or optimization.

Table 2 Training data covariance (*italics*)/correlation matrix

	x_t	y_t	z_t	x_r	y_r	z_r
x_t	*7.281*	0.097	−0.087	−0.457	0.341	−0.026
y_t	*0.103*	*0.155*	0.260	−0.037	−0.044	−0.068
z_t	*−0.019*	*0.008*	*0.006*	0.124	−0.134	−0.414
x_r	*−0.465*	*−0.005*	*0.004*	*0.142*	−0.924	0.210
y_r	*1.099*	*−0.021*	*−0.013*	*−0.417*	*1.431*	−0.166
z_r	*−0.000*	*−0.000*	*0.000*	*0.002*	*−0.001*	*0.000*

Biomechanical test case

3.3 Probability interpretation of MDs associated with prediction sites (if applicable)

If the training data inputs are assumed to be normally distributed, each iso-probability contour (ellipsoid/hyper-ellipsoid) has a particular MD associated with it (McLachlan 1992). Hence, the likelihood/probability p that a prediction site is within the training data boundaries can be computed using its MD. This probability interpretation allows the user to account for uncertainty when classifying points as inside or outside training data boundaries. For example, if one assumes that the covariance matrix has the same variance in all directions (spherical/hyper-spherical), then, for points at one standard deviation from the mean of the training data input (MD = 1), the probability p can be obtained using a standard normal distribution and calculated as 1–0.68; for $MD = 2$, p is approximately 1–0.95; and for $MD = 3$, p is approximately 1–0.997.

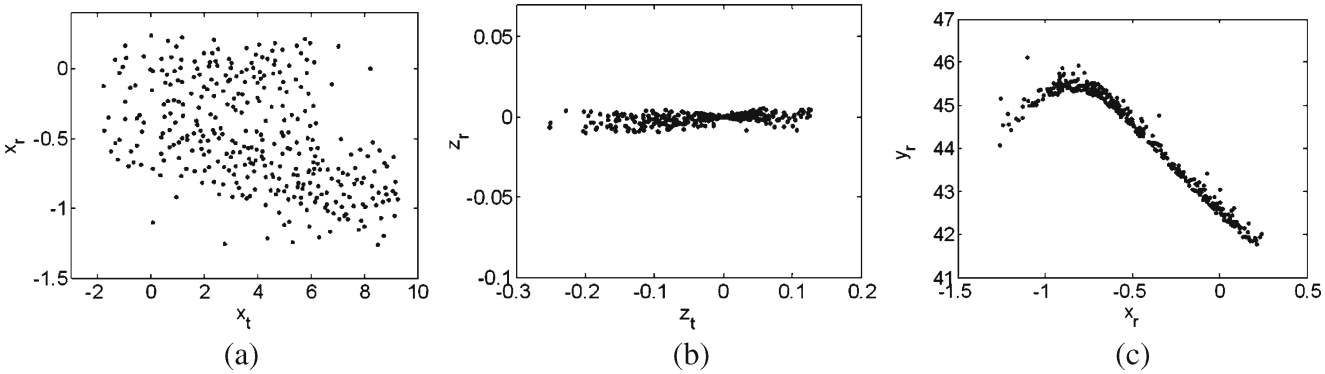

Fig. 4 Scatter plots for training data of sample pairs of input variables: **a** x_r & x_t, **b** z_r & z_t, **c** y_r & x_r. Biomechanical test case

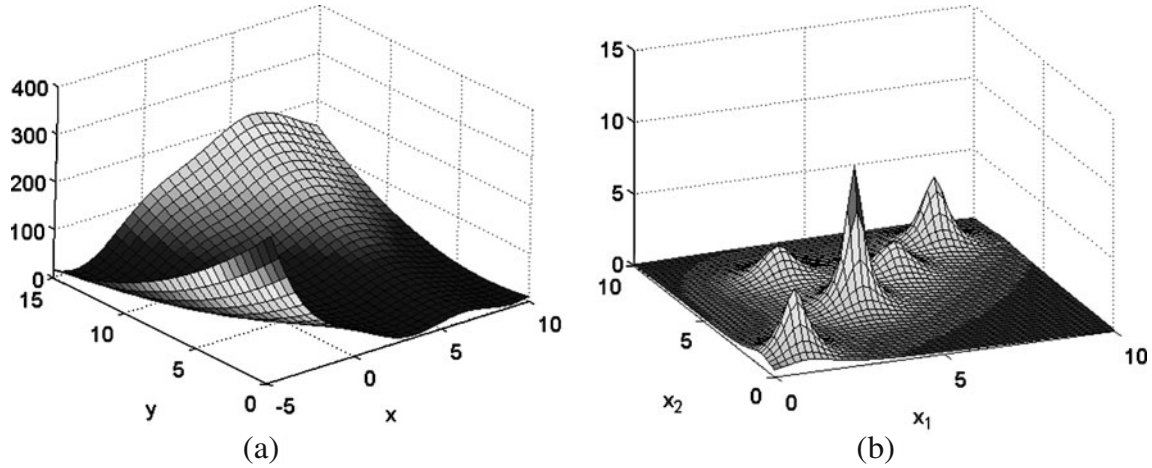

Fig. 5 Branin-Hoo test function (**a**) and a two dimensional illustration of the Shekel 4D test function (**b**)

For normally distributed training data inputs, the squared Mahalanobis distances MD^2 are Chi-square distributed with n degrees of freedom $\left(\chi_n^2\right)$, where n is the number of input variables for the surrogate model of interest (null hypothesis). A Kolmogorov–Smirnov (Stephens 1974) test evaluates if such null hypothesis cannot be rejected. This test compares the cumulative distribution of a χ_n^2 and the corresponding distribution of the MD^2 of the training data inputs and assesses whether their differences are statistically significant ($\alpha = 0.05$).

4 Test cases

This section presents performance measures and alternative restricted input spaces for evaluating the proposed approach. Two analytical test functions (Branin-Hoo and Shekel 4D) and a biomechanical test case are also included for scenarios when extrapolation error trends are of interest.

4.1 Performance measures

Performance measures are intended to evaluate the proposed approach using the following criteria: i) level of approximation of the boundaries of the restricted input spaces, ii) ability to classify prediction sites as inside or outside known restricted input spaces, and iii) capability to reflect expected error trends. Specifically, these measures are:

- For the analytical test cases, the difference between the estimated MD_b and the median of the MDs of the points on the theoretical boundaries of the given restricted input spaces. *Rationale: the estimated MD_b should give a good approximation to a robust measure of the boundaries of the given restricted input spaces.*
- For points outside the theoretical boundaries, the correlation between ΔMD_b and the Euclidean distance from the prediction site to the nearest vertex in the training data convex hull; ΔMD_b represents the difference between the MD at the prediction site and the MD_b. The convex hulls are obtained using the Quickhull algorithm (Barber et al. 1996) as implemented in the Matlab function `convhulln`. *Rationale: A statistically significant correlation should be observed between the two distance measures under consideration.*
- For the analytical test cases, balanced error rates (BER) when establishing if a given prediction site is inside or outside the boundaries of a known restricted input space. A prediction site is classified as inside the restricted input space if its MD is less than the MD_b

Table 3 MD_bs for the analytical and biomechanical test cases

Test case	MD_{b20}(sample size)			MD_{bl}(sample size)		
Circular input space (2D)	1.68 (15)	1.74 (30)	1.75 (45)	2.20 (15)	2.07 (30)	2.05 (45)
Elliptical input space (2D)	1.93 (15)	1.86 (30)	1.91 (45)	2.06 (15)	2.17 (30)	2.11 (45)
Hyper-spherical input space (4D)	2.37 (40)	2.31 (80)	2.38 (160)	2.81 (40)	2.76 (80)	2.62 (160)
Hyper-ellipsoidal input space (4D)	2.49 (40)	2.44 (80)	2.43 (160)	2.99 (40)	2.73 (80)	2.86 (160)
Biomechanical problem (6D)	3.33 (339)				6.48 (339)	

MD_{b20} and MD_{bl} are the median of the top 20% and the maximum of the MDs in the training data, respectively

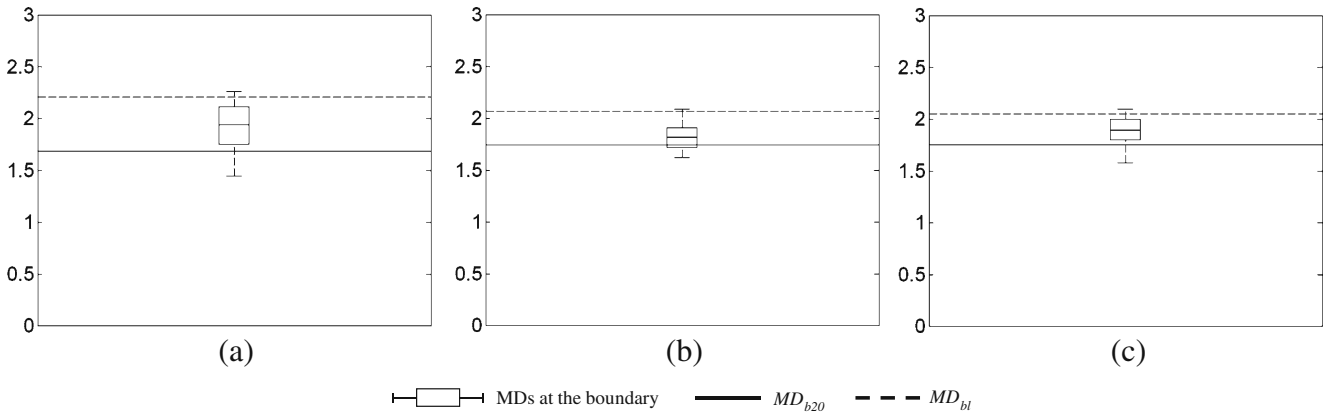

Fig. 6 Boxplots of the Mahalanobis distance for a sample of points in the boundary of the circular input space and estimated MD_b. Number of training data: **a** 15, **b** 30 and **c** 45. Two dimensional test case

(training data) and as outside otherwise. The BER measure (4) is defined as the average of the classifying relative error for each class (i.e., inside or outside). *Rationale: The proposed approach should exhibit a reasonable balanced error rate*.

$$BER = 0.5\left(\frac{\#\text{ of incorrectly classif. as outside}}{\#\text{ of prediction sites inside}} + \frac{\#\text{ of incorrectly classif. as inside}}{\#\text{ of prediction sites outside}}\right) \quad (4)$$

- For points outside the theoretical boundaries, plots of the logarithm of the relative extrapolation error versus ΔMD_b. The errors were computed using a Kriging model with a quadratic trend and Gaussian correlation model (Cressie 1993) implemented in Matlab (Lophaven et al. 2002). *Rationale: The error should increase with higher values of* ΔMD_b.
- The plots discussed above are compared with those obtained using the Euclidean distance to the nearest vertex in the training data convex hull. *Rationale: Similar trends should be observed between the results obtained using the proposed and convex hull approaches*.

4.2 Restricted input spaces

The proposed approach was evaluated using two, four and six dimensional test cases. For the two and four dimensional test cases, circular/hyper-spherical and elliptical/hyper-ellipsoidal restricted input spaces were considered. Table 1 shows a characterization of the shape of the restricted input spaces and training and test data; the hyper-sphere has a radius equal to five (5) and the hyper-ellipsoid satisfies the equation $\mathbf{x}^T\mathbf{A}\mathbf{x} = r$, with A = [1.5 1 0 0; 1 2 0 0; 0 0 1 0.25; 0 0 0.25 1] and $r = 5$.

The *training data inputs* were generated using Latin hypercube sampling (McKay et al. 1979) filtering out points that lied outside the selected restricted input spaces. Furthermore, to achieve a pre-specified number of training data, points were randomly eliminated. On the other hand, the *test data* was also generated using Latin hypercube sampling within box-like boundaries, i.e., the minimum rectangular cuboid covering the given restricted input space, but without any filtering. Figure 2 illustrates training (blue crosses) and test data inputs (green circles) for the two dimensional restricted input spaces.

A six dimensional restricted input space associated with a biomechanical problem (Bei and Fregly 2004) was also considered. The input variables (Fig. 3) are: tibial anterior–posterior translation *xt*, femoral superior–inferior translation *yt*, tibial medial–lateral translation *zt*, femoral varus–valgus rotation *xr*, tibial internal–external rotation *yr*, and femoral flexion–extension *zr*. These variables are of interest in the design of artificial knees for minimizing wear, one of the most significant problems in biomechanical engineering. The data was generated by performing Hammersley

Table 4 Normalized differences between the MD_bs and the median of the MDs on the theoretical boundaries

Restricted input space	MD_{b20}			MD_{bl}		
Circular	−0.129	−0.039	−0.074	0.141	0.139	0.082
Elliptical	−0.056	−0.110	−0.089	0.009	0.040	0.012
Hyper-spherical	−0.033	−0.062	−0.020	0.071	0.053	0.046
Hyper-ellipsoidal	−0.011	−0.023	−0.042	0.133	0.064	0.019

Two dimensional test cases

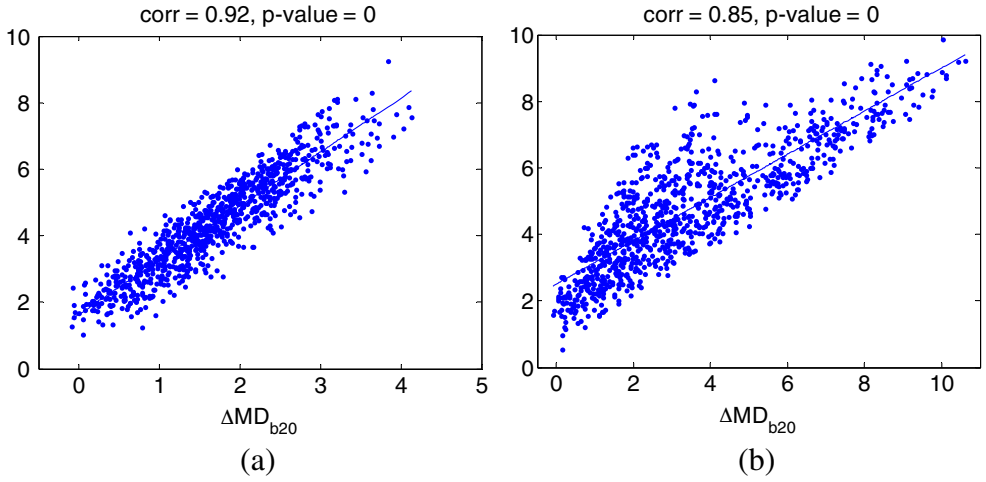

Fig. 7 Euclidean distance to the nearest vertex in the convex hull vs. ΔMD_{b20} corresponding to 1,000 random test data: **a** hyper-spherical input space, **b** hyper-ellipsoidal input space. Four dimensional test case—80 training data

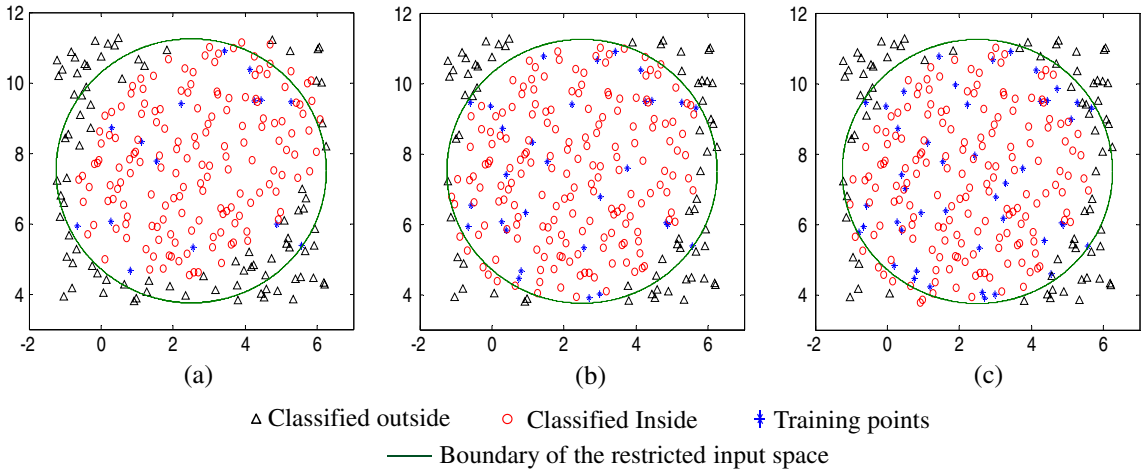

Fig. 8 Classification of test data with a restricted circular input space. Number of training data: **a** 15, **b** 30 and **c** 45. Two dimensional test case

Table 5 Classification results of the proposed approach for a hyper-ellipsoidal restricted input space

Reality/decision	Size of the training data					
	40 points		80 points		160 points	
	Inside	Outside	Inside	Outside	Inside	Outside
Inside	108	36	113	31	119	25
Outside	16	840	4	852	1	855

Four dimensional test case

Table 6 Classification balanced error rates

Hyper-spherical			Hyper-ellipsoidal		
40 points	80 points	160 points	40 points	80 points	160 points
0.133	0.113	0.051	0.134	0.110	0.087

Four dimensional test case

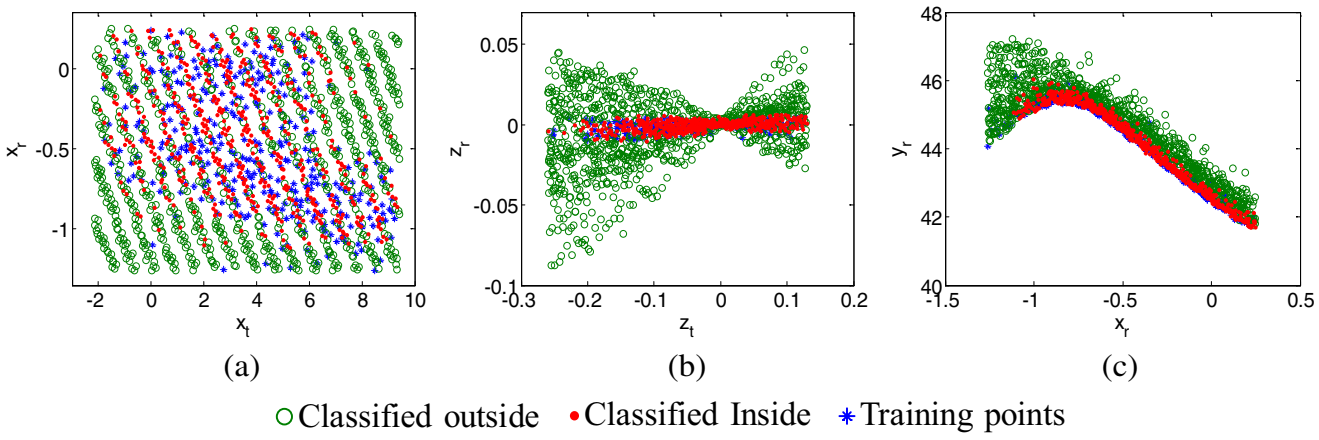

Fig. 9 Scatter plots of sample pairs of input variables for training data and classified prediction sites (339 training data—1,414 prediction sites): **a** x_r & x_t, **b** z_r & z_t, **c** y_r &x_r. Biomechanical test case

quasirandom sampling (Hammersley 1960) resulting in 339 training data, after filtering out physically unrealistic points using a reasonable design space (RDS) approach (Lin et al. 2008), and 1,414 test data. Training data scatter plots for three sample pairs of input variables are shown in Fig. 4; the selected sample pairs were those exhibiting the highest variability. Table 2 shows covariance values (lower triangular matrix/italics) and correlation values (off-diagonal upper triangular matrix) of the input variables.

4.3 Extrapolation error trends—Branin-Hoo, Shekel 4D test functions, contact force F_y in a knee replacement

Equations (5) and (6) show the mathematical expressions of the analytical test functions, with two dimensional illustrations depicted in Fig. 5.

$$\begin{aligned} Branin - Hoo(x,\ y) = &\left(\frac{5.1x^2}{4\pi^2} + \frac{5x}{\pi} - 6\right)^2 \\ &+ 10\left(1 - \frac{1}{8\pi}\right)\cos(x) + 10 \\ -5 < x < 10 \\ 0 < y < 15 \end{aligned} \tag{5}$$

$$\begin{aligned} Shekel4D(x_1,\ x_2,\ x_3\ x_4) = &\sum_{i=1}^{5} \frac{1}{c_i + \sum_{j=1}^{4}\left(x_j - a_{ji}\right)^2} \\ a = &\begin{bmatrix} 4 & 1 & 8 & 6 & 3 \\ 4 & 1 & 8 & 6 & 7 \\ 4 & 1 & 8 & 6 & 3 \\ 4 & 1 & 8 & 6 & 7 \end{bmatrix} \quad c = [0.1\ \ 0.2\ \ 0.2\ \ 0.4\ \ 0.4] \\ &0 < x_j < 10, \quad j = 1, 2, ..., 4 \end{aligned} \tag{6}$$

For the biomechanical problem, the output of interest is the superior–inferior contact force F_y in a knee replacement (Fig. 3), calculated through static analyses using an elastic-foundation (i.e., bed of springs) contact model of a knee replacement design. The model employed an element grid of 63 × 35 on the medial and lateral sides of the tibial insert (i.e., polyethylene piece on the top of the shin bone).

5 Results and discussion

Table 3 shows the estimated Mahalanobis distances at the boundary (MD_bs): largest MD (MD_{bl}) or median of top 20% largest MD (MD_{b20}) among those in the training data. These measures were relatively insensitive to the selected sample sizes. The differences between MD_{bl} and MD_{b20}—a robust estimator based on the median—can be significantly different, in particular, in the presence of outliers, as shown by the results corresponding to the biomechanical test case.

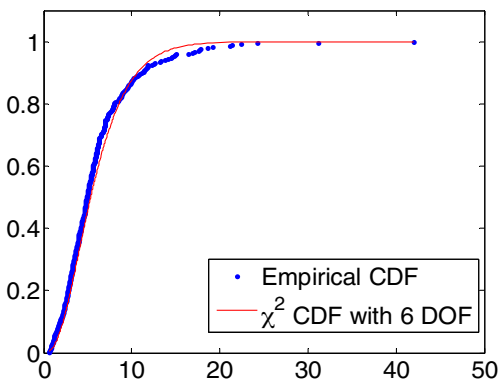

Fig. 10 Comparison of the training data and χ^2_6 CDF's. Biomechanical test case

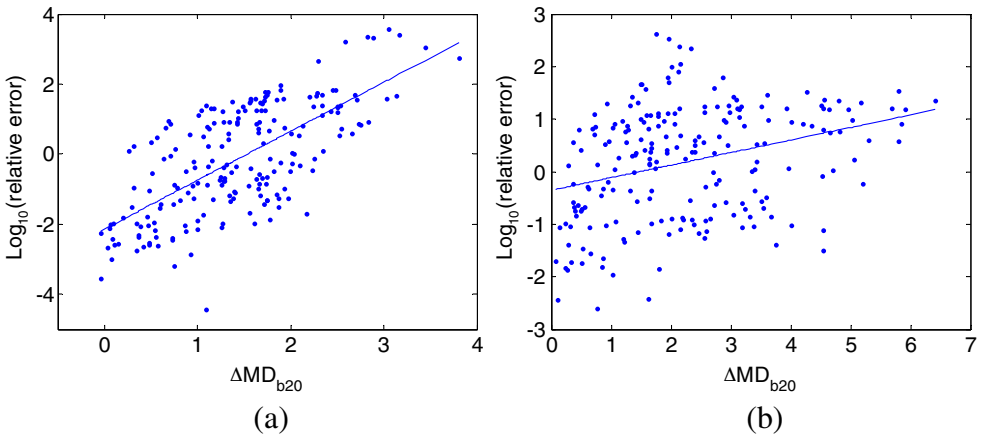

Fig. 11 Logarithm of the relative error vs. ΔMD_{b20} corresponding to 250 random test data: **a** circular input space, **b** elliptical input space. Branin-Hoo test case—30 training data

Figure 6 depicts the box-plots of the Mahalanobis distances of points on the circular theoretical boundaries and the estimated MD_bs, while Table 4 shows the normalized differences between the MD_bs and the median of the MDs on the theoretical boundaries for all analytical test cases. Note that, in most cases, the estimated MD_{b20} outperformed MD_{bl} approximating the median of the MDs of the theoretical boundaries. As a result, MD_{b20} was adopted throughout the rest of the study.

For points outside the theoretical boundary, the difference between the MD at the prediction site and the MD_b (ΔMD_b), and the Euclidean distances to the convex hull, are linearly related with correlation coefficients ranging from 0.81 to 0.94 for all analytical test cases and sample sizes under consideration. The statistical significance of the results is reflected by zero p values for the null hypothesis of uncorrelated inputs. The results show that ΔMD_b could represent a proxy of distances to the convex hull, and give an indication of how risky and prone to error is the surrogate prediction. As an example, Fig. 7 shows the relationship between the two cited measures corresponding to the four dimensional test case for hyper-spherical and hyper-ellipsoidal restricted input spaces and selected sample sizes.

Establishing whether or not it is reasonable to use a surrogate model at a given prediction site relies heavily on correctly classifying whether a prediction site is inside or outside the training data boundaries. In general, the classification results were good for all the restricted input spaces under consideration, with lower balanced error rates as the number of training data increases. For the circular and elliptical two dimensional restricted input spaces, the balanced error rates were in the intervals [9.7%, 17.1%] and [5.5%, 8.8%], respectively, with lower balanced error rates corresponding to the highest sample size. Figure 8 depicts the effectiveness of the proposed approach in the two dimensional test case with circular restricted input space.

On the other hand, for the hyper-spherical and hyper-ellipsoidal four dimensional restricted input spaces the balanced error rates were in the intervals [5.1%, 13.3%] and [8.7%, 13.4%], respectively. Tables 5 and 6 give details of the classifying performance for the hyper-ellipsoidal restricted input space and the balanced error rates for alternative sample sizes, respectively.

For the six dimensional biomechanical test case, the shape of the restricted input space is unknown, hence the classification capabilities of the proposed approach were only qualitatively evaluated. The evaluation criterion is the level of overlapping (as seen in scatter plots) between the training data and the prediction sites classified as inside the restricted input space. Five hundred and forty nine (549) and eight hundred and sixty five (865) prediction sites were classified as inside and outside of the restricted input space, respectively. Figure 9 shows scatters plots for three selected pairs of input variables; note that training data and prediction sites classified as inside are almost superimposed, reflecting the effectiveness of the proposed approach.

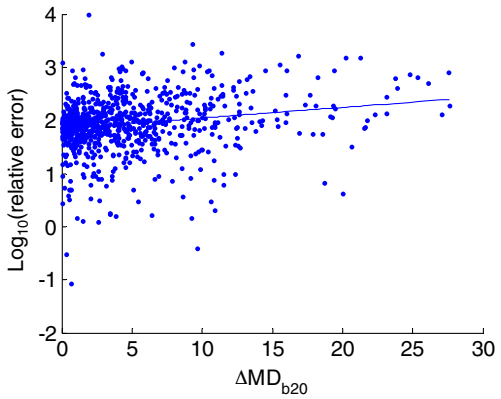

Fig. 12 Logarithm of the relative error vs. ΔMD_{b20}. Biomechanical test case

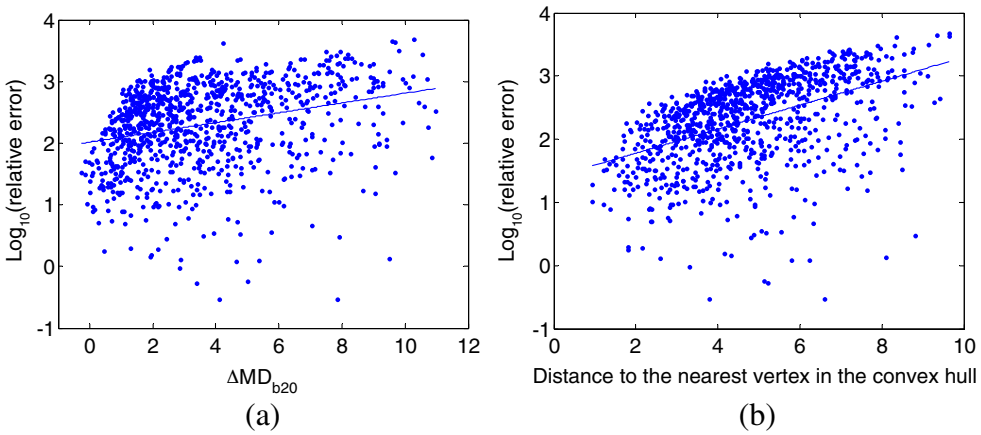

Fig. 13 Logarithm of the relative error vs. **a** ΔMD_{b20}, **b** Euclidean distance to the nearest vertex in the convex hull, corresponding to 1,000 random test data. Shekel 4D test case—hyper-ellipsoidal restricted input space—80 training data

In general, based on the results of the Kolmogorov–Smirnov tests, the training data inputs can be assumed normally distributed and hence allows for a probability interpretation. As an example associated with the biomechanical problem, Fig. 10 shows the cumulative distributions for the squared MDs of the training data and of a Chi-square distribution with six degrees of freedom $\left(\chi_6^2\right)$. The KS test resulted in a p-value of 0.30 well above the rejection threshold ($\alpha = 0.05$) for the null hypothesis. In this scenario, a prediction site with, for instance, $MD_{b20} = 3.34$, could be interpreted as having a probability $p = 0.08$ of being within the training data boundaries.

Regarding expected error trends, in all instances, as desired, the logarithm of the relative error increased with the ΔMD_b; this tendency was observed for all sample sizes. Figures 11 and 12 demonstrate such behavior for the Branin-Hoo, and biomechanical test data, respectively. Similar trends were obtained when using the Euclidean distance from the prediction site to the nearest vertex of the training data convex hull; for example, Fig. 13 shows the results corresponding to the Shekel 4D test function considering an hyper-ellipsoid restricted input space for a sample size of 80 training data.

6 Conclusions

Using a surrogate model to extrapolate outside the region defined by the training data can lead to large errors. Therefore, a strategy for assessing the consistency of prediction sites to the training data should always be employed. This paper presented a computationally efficient approach to estimate the training data boundaries in surrogate modeling with restricted input spaces using a Mahalanobis distance (MD). This distance can be selected as a threshold for deciding whether or not a particular prediction point is within the boundaries of the training data and has the potential for a likelihood/probabilistic interpretation.

Even with limited sample sizes, the proposed approach exhibited good approximations to the training data boundaries, and reasonable balanced error rates when classifying prediction sites as inside or outside known restricted input spaces, with error trends aligned with those obtained using the more computationally expensive convex hull approach.

Given the computational efficiency, performance, and potential for likelihood/probabilistic interpretation, the use of MD for estimating the boundaries of training data holds promise for becoming a popular method among practitioners when establishing whether or not it is reasonable to use a surrogate model at a given prediction site in the context of analysis and optimization efforts.

Acknowledgement This work was supported in part by the National Science Foundation CBET Division under Grant No. 0602996 to B. J. Fregly and R. T. Haftka.

References

Barber CB, Dobkin DP, Huhdanpaa HT (1996) The Quickhull algorithm for convex hulls. ACM Trans Math Softw 22(4):469–483

Bei Y, Fregly BJ (2004) Multibody dynamic simulation of knee contact mechanics. Med Eng Phys 26:777–789

Cressie NAC (1993) Statistics for spatial data. Wiley, New York

Forrester AIJ, Keane AJ (2009) Recent advances in surrogate based optimization. Prog Aerosp Sci 45:50–79

Hammersley JM (1960) Related problems. 3. Monte-Carlo methods for solving multivariable problems. Ann N Y Acad Sci 86(3):844–874

Jacques J, Lavergne C, Devictor N (2006) Sensitivity analysis in presence of model uncertainty and correlated inputs. Reliab Eng Syst Saf 91:1126–1134

Lin YC, Haftka RT, Queipo NV, Fregly BJ (2008) Dynamic simulation of knee motion using three dimensional surrogate contact modeling. In: Proceedings of the ASME 2008 summer bioengineering conference. SBC, Marco Island

Lophaven SN, Nielsen HB, Sondergaard J (2002) DACE—A Matlab kriging toolbox, version 2.0. Report IMM-TR-2002-12. Informatics and Mathematical Modeling. Technical University of Denmark

Mardia KV, Kent JT, Bibby JM (1979) Multivariate analysis. Academic Press

McKay MD, Beckman RJ, Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics (American Statistical Association) 21(2):239–245

McLachlan GJ (1992) Discriminant analysis and statistical pattern recognition. Wiley Interscience

Missoum S, Ramu P, Haftka RT (2007) A convex hull approach for the reliability-based design optimization of nonlinear transient dynamic problems. Comput Mehods Appl Mech Eng 196:2895–2906

Mount DM (2002) CMSC 754 Computational geometry. Lecture Notes, University of Maryland

Queipo NV, Haftka RT, Shyy W, Goel T, Vaidyanathan R, Kevin Tucker P (2005) Surrogate based analysis and optimization. Prog Aerosp Sci 41:1–28

Shioda R, Tunçel L (2007) Clustering via minimum volume ellipsoid. Comput Optim Appl 37:247–295

Stephens MA (1974) EDF statistics for goodness of fit and some comparisons. J Am Stat Assoc 69:730–737

Sun P, Freund MR (2004) Computation of minimum-volume covering ellipsoids. Oper Res 52(5):690–706