8th World Congress on Structural and Multidisciplinary Optimization
June 1 - 5, 2009, Lisbon, Portugal

Safety Margins for Conservative Surrogates

Felipe A. C. Viana[1], Victor Picheny[2], Raphael T. Haftka[3]

[1] University of Florida, Gainesville, USA, fchegury@ufl.edu
[2] University of Florida, Gainesville, USA, vpicheny@ufl.edu
[3] University of Florida, Gainesville, USA, haftka@ufl.edu

1. Abstract
Using surrogate models for learning or optimization creates a risk associated to the fitting error that must be accounted for. Conservative surrogates are metamodels designed to safely estimate the actual response of the system. In this work we use safety margins to generate conservative surrogates. Given a desired level of conservativeness (percentage of safe predictions), we propose the use of cross-validation for estimating the required safety margin. We also explore how multiple surrogates and cross-validation can be used to minimize the loss of accuracy inherent in conservative surrogates. The approach was tested on two algebraic examples for ten basic surrogates including different instances of kriging, polynomial response surface, radial basis neural networks and support vector regression surrogates. For these examples we found that cross-validation (i) is effective for selecting the safety margin; and (ii) allows us to select a surrogate with the best compromise between conservativeness and loss of accuracy.

2. **Keywords:** Conservative surrogates, Cross-validation, Multiple surrogates, Safety margin.

3. Introduction
The use of surrogates for facilitating optimization and statistical analysis of computationally expensive simulations has become commonplace [1]-[4]. Sophisticated surrogates such as kriging [5], [6]; radial basis neural networks [7], [8]; and support vector regression [9], [10]; increasingly share place with the traditional polynomial response surface [11], [12]. Usually, surrogate models are fit to be unbiased, so that predictions are equally likely to be below and above the actual value of the response (i.e., the error expectation is zero). However, in many applications, we want to obtain approximations that are expected to be as close as possible but on the safe side of the actual response [13]-[16]. For example, in structural analysis, stress or strain values must not be underestimated in order to avoid failure. In this paper, when estimates are higher than the true response we call them conservative. Hence, conservative surrogates tend to overestimate the actual response; and as a consequence, there is a trade-off between accuracy and conservativeness.

One of the most widely used methods for conservative estimation is to bias the prediction response by additive or multiplicative constants (termed safety margin and safety factors, respectively). The choice of the constant is often based on previous knowledge of the problem. However, for surrogate-based analysis, there is no established practice for choosing the safety margin. In their previous work [17], the authors explored and compared different alternatives to produce conservative predictions with surrogates. They found that safety margins and estimators based on the surrogate error distribution were comparable in performance, but the safety margin approach lacked a basis for selecting the magnitude of the margin.

In this paper we propose the use of cross-validation for estimating the safety margin. For unbiased surrogates, there has been research pointing to the utility of cross-validation [18]-[20] for selecting the most accurate surrogate from several fitted models. Thus we also explore the use of cross-validation for selecting a surrogate with desired combination of safety and accuracy; since different surrogates may have different performance in terms of the tradeoff between conservativeness and accuracy.

The rest of the paper is organized as follows. Section 4 reviews the basis of conservative prediction using safety margins. Section 5 introduces the proposed approach for designing the safety margins based on cross-validation. Sections 6 and 7 define the numerical experiments and presents results and discussion. Finally, the paper is closed by recapitulating salient points and concluding remarks.

4. Background

4.1. Conservative Surrogates

We denote by y the response of a numerical simulator or function that is to be studied:

$$\begin{array}{llll} y: & D \subset \mathbb{R}^d & \rightarrow & \mathbb{R} \\ & \mathbf{x} & \mapsto & y(\mathbf{x}) \end{array}, \tag{1}$$

where $\mathbf{x} = \{x_1, \ldots, x_d\}^T$ is a d-dimensional vector of input variables.

Since the response $y(\mathbf{x})$ is expensive to evaluate, we approximate it by a cheaper model $\hat{y}(\mathbf{x})$ (surrogate model), based on (i) assumptions on the nature of $y(\mathbf{x})$; and (ii) on the observed values of $y(\mathbf{x})$ at a set of points, called the design of experiment (DOE).

A conservative surrogate, $\hat{y}_C(\mathbf{x})$, obtained by adding a safety margin, s, to an unbiased surrogate model, $\hat{y}(\mathbf{x})$, is an empirical estimator of the type:

$$\hat{y}_C(\mathbf{x}) = \hat{y}(\mathbf{x}) + s \ . \tag{2}$$

In this paper, when we check the accuracy of a conservative surrogate, we use the root mean square error (e_{RMS}).

$$e_{RMS} = \left(\int_D \left(\hat{y}_C(x) - y(x) \right)^2 dx \right)^{\frac{1}{2}} \tag{3}$$

The e_{RMS} is computed by Monte-Carlo integration at a large number of p_{test} test points[*]:

$$e_{RMS} \approx \sqrt{\frac{1}{p_{test}} \sum_{i=1}^{p_{test}} e_{Ci}^2} \ , \tag{4}$$

$$e_{Ci} = \hat{y}_{Ci} - y_i = \hat{y}_i - y_i + s \ , \tag{5}$$

where $\hat{y}_{Ci}$ and y_i are values of the conservative prediction and actual simulation at the i-th test point, respectively.

4.2. Cross-Validation

Cross-validation is a process of estimating errors by constructing the surrogate without some of the points and calculating the errors at these left out points. The process is repeated with different sets of left-out points in order to get statistically significant estimates of errors. The process proceeds by dividing the set of p data points into k subsets. The surrogate is fit to all subsets except one, and error is checked in the subset that was left out. This process is repeated for all subsets to produce a vector of cross-validation errors, $\mathbf{e}_{XV}$ (also known as the PRESS vector, where PRESS stands for prediction sum of squares). Figure 1 illustrates the cross-validation errors for a kriging surrogate.

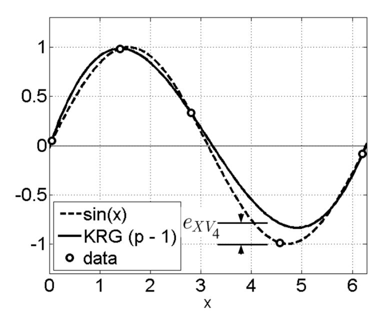


Figure 1: Cross-validation error at the fourth point of the DOE, e_{XV_4}, illustrated by fitting a kriging model (KRG) to $p = 5$ data points of the function sin(x).

---

[*] We computed the actual errors using a Latin Hypercube design [24] of 10,000 points created by the MATLAB function lhsdesign [25] set with the "maxmin" option with 10 iterations.

For comparing surrogates based on the data only at the p points of the design of experiment (DOE), we use cross-validation errors, $\mathbf{e}_{XV}$. The e_{RMS} is estimated from $\mathbf{e}_{XV}$:

$$PRESS_{RMS} = \sqrt{\frac{1}{p}\mathbf{e}_{XV}{}^T\mathbf{e}_{XV}} \ . \tag{6}$$

Since $PRESS_{RMS}$ is an estimator of e_{RMS}, one possible way of using multiple surrogates is to select the model with best (i.e., smallest) $PRESS_{RMS}$ value. We call such model the BestPRESS surrogate. See Appendix A for more details about surrogate selection based on cross-validation errors.

When we add a safety margin to a predictor, we do not need to repeat the costly process of cross-validation to assess the new *PRESS* vector, because the vector of cross-validation errors associated with $\hat{y}_C\left(\mathbf{x}\right)$, $\mathbf{e}_{XVC}$, is simply:

$$\mathbf{e}_{XVC} = \mathbf{e}_{XV} + s \ . \tag{7}$$

And with that, the computation of PRESS for the conservative surrogate does not require any extra computation beyond the cross-validation errors of the unbiased surrogate.

4.3. Percent Conservative Errors and Loss in Accuracy

There are different measures of the conservativeness of an approximation including the average error or the maximum non-conservative error. Here we use the percentage of conservative errors (i.e. positive):

$$\%c = 100\int_D I\left[\hat{y}_C\left(x\right) - y\left(x\right)\right]dx \tag{8}$$

where, $I\left(e\right)$ is the indicator function, which equals 1 if $e > 0$ and 0 otherwise. $\%c$ can be estimated by Monte-Carlo integration:

$$\%c \approx \frac{100}{p_{test}}\sum_{i=1}^{p_{test}} I\left(e_i\right) , \tag{9}$$

Ideally, $\%c = 50\%$, when $s = 0$.

As stated before, conservative estimators tend to overestimate the true values. As a consequence the accuracy of the surrogate is degraded. Figure 2 illustrates the effect of using a safety margin. It is easy to see that the reduction in non-conservative errors entails increasing the root mean square error (e_{RMS}).

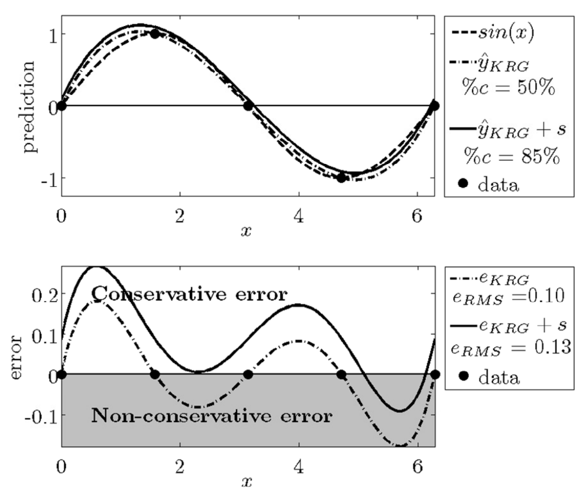


Figure 2: Effects of adding a safety margin to a kriging model (KRG) for the $\sin\left(x\right)$ function fitted with five data points. Conservativeness comes at the price of loosing accuracy.

We also define the relative loss in accuracy:

$$l_a = \frac{e_{RMS}}{e_{RMS}\big|_{ref}} - 1 , \tag{10}$$

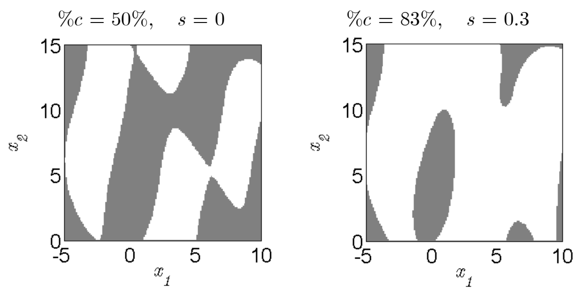
where e_{RMS} is taken at a given target conservativeness; and $e_{RMS}\big|_{ref}$ is the e_{RMS} value of reference. In the first part of the study, we use as reference the value of e_{RMS} when the target conservativeness is 50%. In the second part, when we compare the set of surrogates, we use as reference the e_{RMS} of the most accurate surrogate when no safety margin is added.

5. Selecting Safety Margin using Cross-Validation

In terms of the cumulative distribution function (CDF) of the errors, $F\left(e\left(\mathbf{x}\right)\right)$, the safety margin for a given conservativeness, $\%c$, is given as:

$$s = F^{-1}\left(\frac{\%c}{100}\right). \tag{11}$$

According to (9), the value of the safety margin is such that it makes a desired percentage of the errors to become conservative. Figure 3 illustrates this concept. Figure 3-(a) shows how adding a safety margin makes the surrogate being more conservative. Figure 3-(b) demonstrates how safety margin and errors are related.



(a) Two-dimensional visualization.

$$\mathbf{e} = \hat{\mathbf{y}} - \mathbf{y} = \begin{bmatrix} -0.5 & -0.3 & -0.15 & 0.05 & 0.2 & 0.4 \end{bmatrix}^T$$
$$\mathbf{e}_c = \mathbf{e} + s$$

$$s = 0.5 \Rightarrow \mathbf{e}_c = \begin{bmatrix} 0 & 0.2 & 0.35 & 0.55 & 0.7 & 0.9 \end{bmatrix}^T \Rightarrow \%c = 100\%$$
$$s = 0.3 \Rightarrow \mathbf{e}_c = \begin{bmatrix} -0.2 & 0 & 0.15 & 0.35 & 0.5 & 0.7 \end{bmatrix}^T \Rightarrow \%c = 83\%$$
$$s = 0.15 \Rightarrow \mathbf{e}_c = \begin{bmatrix} -0.35 & -0.15 & 0 & 0.2 & 0.35 & 0.55 \end{bmatrix}^T \Rightarrow \%c = 67\%$$

(b) Conservativeness and errors.

Figure 3: Illustration of relationship between safety margin and error. Safety margin makes the predictor to achieve a desired conservativeness. Figure 3-(a): in white, area where error is conservative (positive) for a two-dimensional example. Percent conservativeness calculated by integration. Figure 3-(b): how percent conservativeness is obtained from a vector of errors.

Besides estimating the e_{RMS} using (6), in this work, we also propose using cross-validation to estimate the required safety margin, s. This is done by using in (9) the vector of cross-validation errors $\mathbf{e}_{XVC}$. Then in (11), we use the experimental CDF of the cross-validation errors.

We also use cross-validation to predict the loss in accuracy l_a using $PRESS_{RMS}$ in (10), instead of e_{RMS}:

$$l_{aXV} = \frac{PRESS_{RMS}}{PRESS_{RMS}\big|_{ref}} - 1 \tag{12}$$

When considering a single surrogate, $PRESS_{RMS}\big|_{ref}$ is the value when the target conservativeness is 50%. When considering multiple surrogates, the $PRESS_{RMS}\big|_{ref}$ value is for the surrogate with smallest $PRESS_{RMS}$ of the set (BestPRESS surrogate).

6. Numerical Experiments

6.1. Basic Surrogates

Table 1 details the six different basic surrogates used during the investigation. The DACE toolbox of Lophaven et al. [27], SURROGATES toolbox of Viana [28], the native neural networks MATLAB toolbox [25], and the code developed by Gunn [29] were used to execute the kriging, polynomial response surface, radial basis neural network, and support vector regression algorithms, respectively. The SURROGATES toolbox was also used for easy manipulation of the surrogates. In this work, we use multiple instances of different surrogates (in the same fashion of [20] and [30]). This is possible because some techniques such as kriging and support vector regression allow different instances by changing parameters such as basis, correlation and loss functions.

Table 1: Information about the set of 10 surrogates.

Surrogates	Details
KRG0	**Kriging models:** KRG0, KRG1, and KRG2 indicate zero, first, and second order polynomial trend model, respectively. In all cases, a Gaussian correlation and $\theta_{0i} = \left(p^{-1/n_v}\right)$, and $10^{-3} \le \theta_i \le 2 \times \theta_{0i}$, $i = 1,2,\ldots,n_v$ were used as initial values and bound, respectively, for the correlation parameters estimation. We chose 3 different kriging surrogates by varying the trend.
KRG1	
KRG2	
RBNN	**Radial basis neural network:** $Goal = \left(0.05\bar{y}\right)^2$ and $Spread = 1/3$.
GRBF-Full	**Support vector regression:** GRBF and Poly indicate the kernel function (Gaussian and second order polynomial respectively). All use loss function as ε-insensitive and quadratic, respectively. Full and Short refer to different values for the regularization parameter, C, and for the insensitivity, ε. Full adopts $C = \infty$ and $\varepsilon = 1 \times 10^{-4}$, while Short uses the selection of values according to Cherkassky and Ma [26]: $\varepsilon = \sigma_y / \sqrt{k}$; and for both $C = \max\left(\left\|\bar{y} + 3\sigma_y\right\|, \left\|\bar{y} - 3\sigma_y\right\|\right)$, where $\bar{y}$ and σ_y are the mean value and the standard deviation of the function values at the design data, respectively. We chose 4 different SVR surrogates by varying the kernel function and the SVR parameters (*C* and ε).
GRBF-Short	
Poly-Full	
Poly-Short	
PRS2	**Polynomial response surface:** Full models of degree 2 and 3.
PRS3	

6.2. Test Problems

As test problems, we employed two following widely used analytical benchmark problems [31]:

Branin-Hoo (2 variables):

$$y(\mathbf{x}) = \left(x_2 - \frac{5.1x_1^2}{4\pi^2} + \frac{5x_1}{\pi} - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10 \ , \quad -5 \le x_1 \le 10, \ \ 0 \le x_2 \le 15 \ . \tag{13}$$

Hartman6 (6 variables):

$$y(\mathbf{x}) = -\sum_{i=1}^{4} a_i \exp\left(-\sum_{j=1}^{6} b_{ij}\left(x_j - d_{ij}\right)^2\right) , \quad 0 \le x_j \le 1 \ , \ j = 1,\ 2,\ \ldots,\ 6 \ , \tag{14}$$

$$\mathbf{B} = \begin{bmatrix} 10.0 & 3.0 & 17.0 & 3.5 & 1.7 & 8.0 \\ 0.05 & 10.0 & 17.0 & 0.1 & 8.0 & 14.0 \\ 3.0 & 3.5 & 1.7 & 10.0 & 17.0 & 8.0 \\ 17.0 & 8.0 & 0.05 & 10.0 & 0.1 & 14.0 \end{bmatrix}, \ \mathbf{D} = \begin{bmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{bmatrix}$$

To investigate the effect of the sampling density, we fitted the Branin-Hoo function using both 17 and 34 points and the Hartaman6 function with 56 and 110 points. The quality of fit, and thus the performance, depends on the design of experiment (DOE). Hence, for all test problems, a set of 1000 different Latin Hypercube designs were used as a way of averaging out the DOE dependence of the results. We used the MATLAB function *lhsdesign*, set with the *"maximin"* option with 1000 iterations to generate the DOEs for fitting.

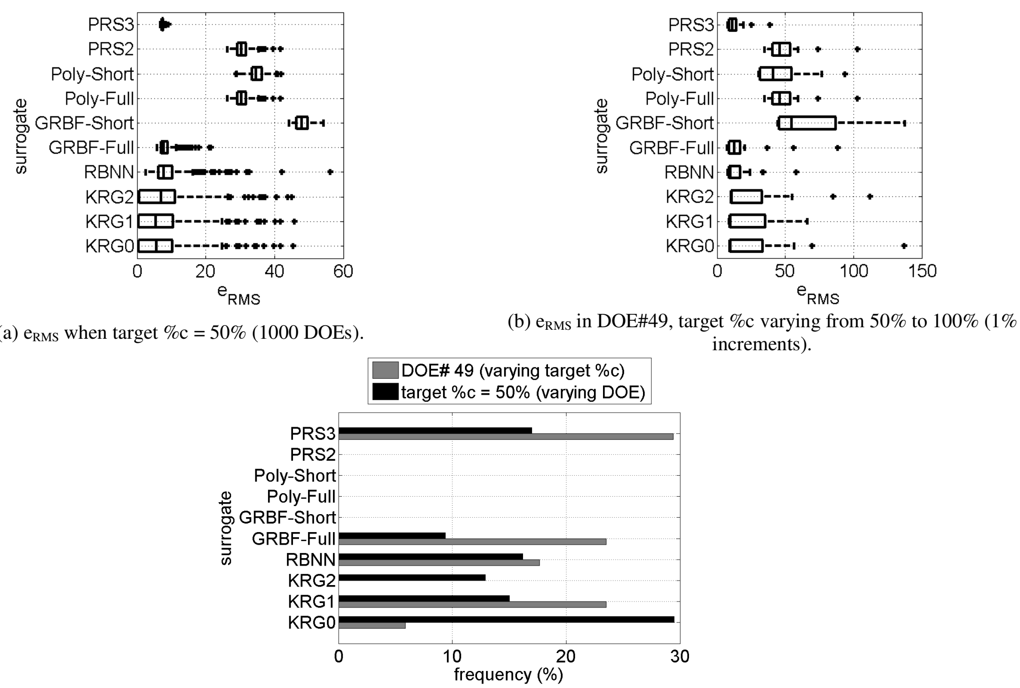
7. Results and Discussion

As stated before, the literature (e.g., [20]) has established that in comparing multiple surrogates, the most accurate one depends both on the problem and on the data set (meaning density and distribution). Figure 4-(a) illustrates this idea showing how the e_{RMS} changes for each of the unbiased surrogates over the 1000 DOEs. In a given DOE, it is natural to look at the variation of the e_{RMS} with respect to the target conservativeness. Figure 4-(b) shows how the e_{RMS} varies in an arbitrarily chosen DOE. The fact that the e_{RMS} values assume a large range either by looking at the unbiased surrogate or at different conservativeness makes us to believe that the selection of the most accurate surrogate depends both on the DOE and on the conservativeness. Figure 4-(c) shows in black the frequency of different unbiased surrogates being most accurate out of 1000 DOEs for the Branin-Hoo function fitted with 34 points. We can see that PRS3, GRBF-Full, and KRG1 are competitive. We then check if in the case of conservative predictors, the best surrogate also changes with the target conservativeness. The gray bars in Figure 4 show the frequency in which the surrogates are most accurate for an arbitrary DOE when the target conservativeness varies from 50% to 100%. For this DOE, PRS3 has the best overall performance, but it is most accurate only for 15 of the 51 levels of conservativeness. We conclude that using multiple surrogates becomes important to avoid further losses by selecting a poorly fitted model.

Next, we check whether cross-validation offers a practical approach for the design of the safety margin. Figure 5 illustrates the performance of three surrogates in estimating the conservativeness level. One is the BestPRESS surrogate (model with smallest $PRESS_{RMS}$ value), the other two are the surrogates that, out of the 1000 DOEs, frequently appear as the most and less accurate unbiased surrogate, respectively.

We can see that:

- For small number of points, we can incur large errors in the selection of safety margin. However, increasing the number of points allows better accuracy in selection of safety margin. Note that sparsity does not seem to be important, as the Hartman function with 56 points does well even though there is less than one point per orthant.
- Selecting the most accurate unbiased surrogate may not lead to best accuracy of safety margin. Figure 5-(b) shows that KRG0 (which is the most accurate unbiased surrogate) poorly performs in the design of the safety margin. KRG0 frequently overestimates the actual conservativeness and its performance is very sensitive to the design of experiment (see large spread).

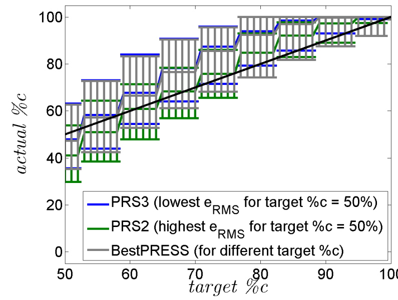


(a) e_{RMS} when target %c = 50% (1000 DOEs).

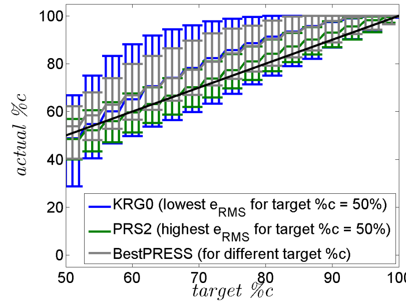
(b) e_{RMS} in DOE#49, target %c varying from 50% to 100% (1% increments).

(c) Frequency of most accurate surrogate

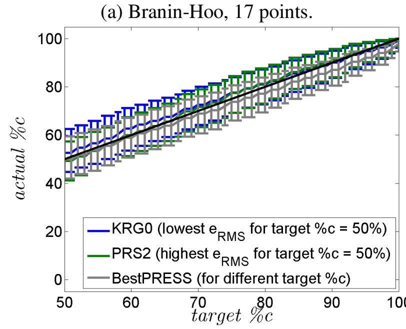
Figure 4: e_{RMS} analysis for Branin-Hoo, 34 points: 1000 DOEs and 51 target %c (from 50% to 100%). Most accurate surrogate changes not only with the design of experiment but also with the target %c.



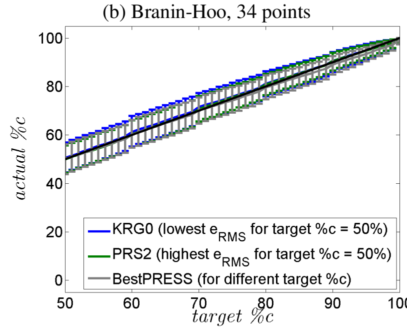
(a) Branin-Hoo, 17 points.



(b) Branin-Hoo, 34 points



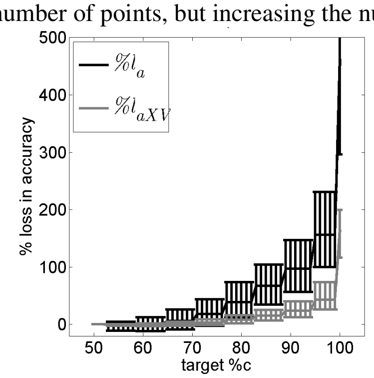
(c) Hartman6, 56 points.



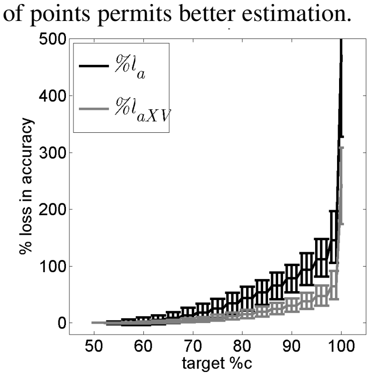
(d) Hartman6, 110 points.

Figure 5: Errorbars with the [10 50 90] percentiles (out of 1000 DOEs) of the actual conservativeness. Increasing the number of points (in spite of sparsity) allows better accuracy in selection of safety margin.

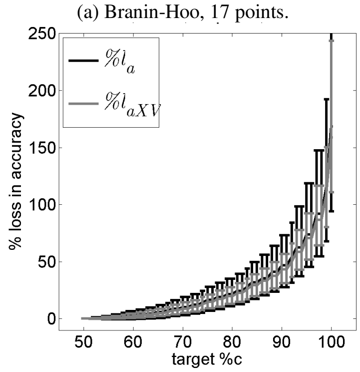
We also checked whether cross-validation allows estimation of loss in accuracy. Figure 6 compares the actual and the estimated loss in accuracy as a function of the target conservativeness for PRS2. Once again, the estimates are poor for small number of points, but increasing the number of points permits better estimation.



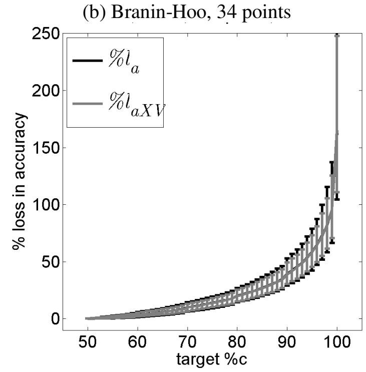
(a) Branin-Hoo, 17 points.



(b) Branin-Hoo, 34 points



(c) Hartman6, 56 points.



(d) Hartman6, 110 points.

Figure 6: Spreads with the [25 50 75] percentiles of the loss in accuracy for PRS2. More points allow good estimation of loss in accuracy.

Finally, we reinforce the benefits of multiple surrogates in the loss of accuracy. Figure 7 gives the example of the Hartman6 function fitted with 110 points. This figure compares four surrogates:

- KRG0 is the most accurate surrogate for low level of conservativeness.
- RBNN is the most accurate surrogate for high values of conservativeness.
- BestPRESS is the surrogate chosen at each conservativeness level using cross validation.
- BestRMSE is the best surrogate chosen at each conservativeness level using the actual value of e_{RMS}.

The figure shows that for the same target conservativeness, the loss in accuracy depends on the surrogate. It means that if we maintain the selection a single surrogate, we will sustain a 10% of loss of accuracy for low level of conservativeness or an 18% loss of accuracy for high levels. We further see that PRESS successfully selects the surrogate for minimum loss of accuracy, since BestPRESS performs almost as well as BestRMSE (which is not available practically).

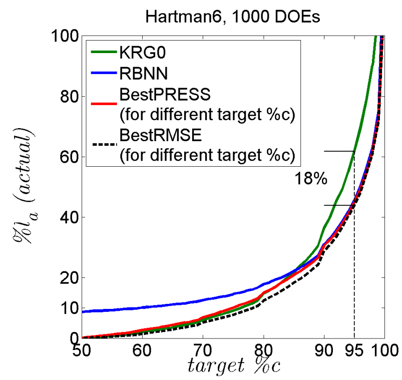


Figure 7: Median of the actual loss in accuracy ($\%l_a$) versus target conservativeness for the Hartman6 with 110 points. We can see that (i) most accurate surrogate changes with target $\%c$; and (ii) cross-validation successfully selects the best choice.

8. Conclusions

We proposed using cross-validation for designing conservative estimators and multiple surrogates for improved accuracy. The approach was tested on two algebraic examples for ten basic surrogates including different instances of kriging, polynomial response surface, radial basis neural networks and support vector regression surrogates. For these examples we found that:

- The best surrogate changes with sampling points (density and location) and with target conservativeness.
- Cross-validation appears to be useful for both estimation of safety margin and selection of surrogate. However, it may not be accurate enough when the number of data points is small.

9. Acknowledgements

This work has been supported by the NASA Constellation University Institute Program (CUIP). Authors gratefully acknowledge this grant.

10. Appendix

A. Cross-Validation for Surrogate Selection:

Because the quality of fit, and thus the performance, depends on the design of experiment (DOE), the BestPRESS surrogate may vary DOE to DOE. The main advantage of creating a diverse and large set of surrogates is reduced chance of accepting poorly fitted surrogates (for example, if the DOE is not good for some surrogates but better for others). Obviously, the success of using BestPRESS for selecting a surrogates depends on the diversity of the set of surrogates and on the fidelity of the $PRESS_{RMS}$ to estimates the true e_{RMS}. Table 2 summarizes surrogates that are considered in the present paper for selection using different criteria in a given design of experiment (DOE). See Table 3 for an illustrative example.

Table 2: Selection of surrogates according to different criteria (in a given DOE). BestRMSE is defined based on testing points; BestPRESS is obtained using cross-validation.

Criterion	Comments
e_{RMS}	Global measure of error in prediction: the smaller, the better.
PRESS	Estimator of the e_{RMS} : the smaller, the better
Surrogate	**Comments**
BestRMSE	Surrogate with the smallest e_{RMS} (most accurate surrogate in terms of prediction).
BestPRESS	Surrogate with smallest $PRESS_{RMS}$.

Table 3: Illustration of selection of surrogates according to different criteria for a hypothetic DOE. Bold face indicates which surrogate is selected. Different criteria may point to different surrogates, and the selection may vary from DOE to DOE. Note that BestRMSE is not practical since it is not available based on the data. Cross-validation succeeds since RBNN represents both BestRMSE and BestPRESS.

Criterion	KRG	PRS	RBNN	SVR
e_{RMS}	3.7	2.8	**2.4**	4.3
PRESS	5.5	4.1	**4.0**	6.1

11. References

[1] J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn, "Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4 (4), pp. 409-435, 1989.

[2] V.V. Toropov, "Simulation approach to structural optimization," *Structural and Multidisciplinary Optimization*, Vol. 1, No. 1, pp. 37-46, 1989.

[3] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P.K. Tucker, "Surrogate-based Analysis and Optimization," *Progress in Aerospace Sciences*, Vol. 41, pp. 1-28, 2005.

[4] T.W. Simpson, V.V. Toropov, V.O. Balabanov, and F.A.C. Viana, "Design and Analysis of Computer Experiments in Multidisciplinary Design Optimization: a Review of How Far We Have Come – or Not," In: *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Victoria, Canada, September 10-12, 2008.

[5] M.L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*, Springer, 1999.

[6] T.W. Simpson, T.M. Mauery, J.J. Korte and F. Mistree, "Kriging Models for Global Approximation in Simulation-Based Multidisciplinary Design Optimization," *AIAA Journal*, Vol. 39, No. 12, pp. 2233-2241, 2001.

[7] J. Park and I.W. Sandberg, "Universal Approximation Using Radial-Basis-Function Networks," *Neural Computation*, Vol. 3, No. 2, pp. 246-257, 1991.

[8] B. Cheng and D.M. Titterington, "Neural networks: a review from a statistical perspective," *Statistical Science*, Vol. 9, No. 1, pp. 2-54, 1994.

[9] A.J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Statistics and Computing*, Vol. 14, No. 3, pp. 199-222, 2004.

[10] B. Scholkopf and A.J. Smola, *Learning with Kernels*, The MIT Press, Cambridge, USA, 2002.

[11] G.E.P. Box, J.S. Hunter and W.G. Hunter, *Statistics for experimenters: an introduction to design, data analysis, and model building*, John Wiley & Sons, Inc, New York, USA, 1978.

[12] R.H. Myers and D.C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, Inc, New York, USA, 1995.

[13] J.H. Starnes, Jr. and R.T. Haftka, "Preliminary Design of Composite Wings for Buckling, Stress and Displacement Constraints," *Journal of Aircraft*, Vol. 16, pp. 564-570, 1979.

[14] V. Picheny, N-H. Kim, R.T. Haftka and J. Peters, "Conservative Estimation of Probability of Failure," In: *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Portsmouth, Virginia, USA, September 6-8, 2006, AIAA-2006-7038.

[15] E. Acar, A. Kale, and R.T. Haftka, "Comparing Effectiveness of Measures that Improve Aircraft Structural Safety," *Journal of Aerospace Engineering*, Vol. 20, No. 3, pp. 186-199, 2007.

[16] E. Acar, R.T. Haftka and T.F. Johnson, "Tradeoff of Uncertainty Reduction Mechanisms for Reducing Structural Weight," ASME Journal of Mechanical Design, Vol. 129, No. 3, pp. 266-274, 2007.

[17] V. Picheny, N-H. Kim, R.T. Haftka and N.V. Queipo, "Conservative Predictions Using Surrogate Modeling," In: *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Schaumburg, USA, April 7-10, 2008.

[18] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pp. 1137–1143, 1995.

[19] M. Meckesheimer, R.R. Barton, T.W. Simpson, and A. Booker, "Computationally Inexpensive Metamodel Assessment Strategies," *AIAA Journal*, Vol. 40 (10), pp. 2053-2060, 2002.

[20] F.A.C. Viana, R.T. Haftka, and V. Steffen Jr, "Multiple surrogates: how cross-validation errors can help us to obtain the best predictor," *Structural and Multidisciplinary Optimization*, Available online, 2009 (DOI: 10.1007/s00158-008-0338-0).

[21] T.W. Simpson, D.K.J. Lin, and W. Chen, "Sampling Strategies for Computer Experiments: Design and Analysis," *International Journal of Reliability and Applications*, Vol. 2, No. 3, pp.209-240, 2001.

[22] A.A. Giunta, S.F. Wojtkiewicz, M.S. Eldred, "Overview of modern design of experiments methods for computational simulations," *41st AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, AIAA-2003-0649, 6–9 January 2003.

[23] T. Goel, R.T. Haftka, W. Shyy, and L.T. Watson, "Pitfalls of using a single criterion for selecting experimental designs," *International Journal for Numerical Methods in Engineering*, Vol. 75, No. 2, pp. 127–155, 2008.

[24] M.D. Mckay, R.J. Beckman, and W.J. Conover, "A Comparison of Three Methods for Selecting Values of Input Variables from a Computer Code," *Technometrics*, Vol. 21, pp. 239-245, 1979.

[25] Mathworks contributors, *MATLAB® The language of technical computing*, Version 7.6 R2008a, The MathWorks Inc, 2008.

[26] V. Cherkassky, and Y. Ma, "Practical Selection of SVM Parameters and Noise Estimation for SVM Regression," *Neural Networks*, Vol. 17, No. 1, pp. 113-126, 2004.

[27] S.N. Lophaven, H.B. Nielsen, and J. Søndergaard, *DACE - A MATLAB Kriging Toolbox*, Technical Report IMM-TR-2002-12, Informatics and Mathematical Modelling, Technical University of Denmark, 2002.

[28] F.A.C. Viana, *SURROGATES ToolBox*, http://fchegury.googlepages.com, 2008.

[29] S.R. Gunn, *Support Vector Machines for Classification and Regression*, Technical Report, Image Speech and Intelligent Systems Research Group, University of Southampton, UK, 1997.

[30] E. Sanchez, S. Pintos and N.V. Queipo, "Toward an Optimal Ensemble of Kernel-based Approximations with Engineering Applications," *Structural and Multidisciplinary Optimization*, Vol. 36, No. 3, pp. 247-261, 2008.

[31] L.C.W. Dixon, G.P. Szegö, *Towards Global Optimization 2*, North Holland, Amsterdam, The Netherlands, 1978.