22nd International Conference
ENGINEERING MECHANICS 2016
Svratka, Czech Republic, 9 – 12 May 2016

PRESS WEIGHTED AVERAGE SURROGATE: TRIAL TESTS IN 2D

E. Myšáková[*], M. Lepš[**]

Abstract: *Surrogate modeling (Meta-modeling) is an often used tool for analysis of behavior of complex systems which are usually described by computationally demanding models. Surrogate models provide an approximation of the original model's response in a fraction of time and therefore are suitable when multiple evaluations are needed. Many types of meta-models exist and each suits another type of problem. On the other hand it is not always possible to select the right meta-model in advance. Therefore parallel construction of several meta-models and their subsequent comparison and combining can be utilized with advantage. A typical method called PRESS weighted average surrogate which uses the prediction sum of squares obtained by cross-validation for computation of the weights for linear combination of individual surrogates is discussed in this contribution and illustrated on several 2-dimensional benchmark examples using a group of different meta-models.*

Keywords: Meta-modeling, Ensemble of Surrogates, Cross-validation, Root Mean Square Error, Prediction Sum of Squares.

1. Introduction

Surrogate modeling (Meta-modeling) constitutes a tool usable for analyses of complex systems which are described by computationally demanding models. Such models cannot be used when multiple evaluations are necessary, for example in Monte Carlo based reliability assessment of the system. In such case the meta-model represents a convenient substitution: it provides an approximation of the original model's response in a fraction of time.

A surrogate model is constructed based on training data which consist of i) training points spread over the design domain as uniformly as possible and ii) responses of the original model in training points. The positions of the points are selected via Design of Experiments (DoE) (Montgomery, 2012). Many types of surrogate models exist and each of them consists of surrogates differing in particular settings which results in a wide group of available meta-models. Therefore it is a logical step to use at least several of them at parallel.

This contribution is focused on an approach for combination of multiple meta-models. The methodology follows the procedure described in (Goel et al., 2007; Viana et al., 2009). The resulting surrogate is created as a linear combination of predictions of the individual meta-models and the weights of individual meta-models are computed based on their quality estimated by cross-validation. The methodology is tested on several illustrative 2-dimensional examples to get a good overview of the procedure. Ten different meta-models are used within the combination.

2. Methodology

The methodology described in this contribution deals with a set of meta-models constructed on the same training data which are selected via LHS design (Iman & Conover, 1980). The number of training points, i.e. the size of the DoE, is 11, 25 and 100. Each constructed meta-model is then tested using four randomly generated testing data sets (10, 100, 1000 and 10000 testing points). Such testing is possible

[*] Ing. Eva Myšáková: Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29, Prague, CZ, eva.mysakova@fsv.cvut.cz

[**] doc. Matěj Lepš, PhD.: Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29, Prague, CZ, leps@cml.fsv.cvut.cz

only because the original model is not computationally demanding; in practical use the large testing data sets cannot be used. The second way of testing the meta-models is cross-validation where the training data are used also for testing. In particular, the leave-one-out cross-validation is utilized. Cross-validation represents a type of testing usable with a real demanding original model because it requires no more original model's evaluations.

2.1. Root Mean Square Error

Having $y(\mathbf{x})$ the actual simulation at the point $\mathbf{x}$, $\hat{y}(\mathbf{x})$ the surrogate's prediction and $e(\mathbf{x}) = y(\mathbf{x}) - \hat{y}(\mathbf{x})$ the error associated with this prediction, the actual root mean square error over the domain with volume V is given by:

$$RMSE_{actual} = \sqrt{\frac{1}{V}\int_V e^2(\mathbf{x})d\mathbf{x}}\ . \tag{1}$$

When using Monte Carlo integration with p_{test} testing points the formula transforms into:

$$RMSE = \sqrt{\frac{1}{p_{test}}\sum_{i=1}^{p_{test}} e_i^2}\ , \tag{2}$$

where $e_i = y_i - \hat{y}_i$ is the error at the i-th testing point.

2.2. Predicted Residual Sum of Squares

When the testing set of points cannot be used a common way for evaluation of the surrogates' accuracy is a cross-validation where training points are used also for testing. Having p training points the leave-one-out cross-validation is performed by constructing p surrogates each of them with one of the training points excluded. Each surrogate is then used for prediction of the excluded point's response. The prediction sum of squares firstly proposed in (Allen, 1974) is then computed using the vector of cross-validation errors (PRESS vector) $\tilde{\mathbf{e}}$. An estimation of the *RMSE* using PRESS vector is given by:

$$PRESS_{RMS} = \sqrt{\frac{1}{p}\tilde{\mathbf{e}}^T\tilde{\mathbf{e}}}\ . \tag{3}$$

2.3. PRESS Weighted Average Surrogate

When constructing n meta-models in parallel the possible and intuitive way of their utilization is their weighted averaging. Naturally, the weights are derived from the individual surrogate's prediction quality. The average surrogate's prediction $\hat{y}_{WAS}(\mathbf{x})$ is then given by a linear combination of individual meta-model's predictions $\hat{\mathbf{y}}(\mathbf{x})$ using weights $\mathbf{w}(\mathbf{x})$:

$$\hat{y}_{WAS}(\mathbf{x}) = \sum_{i=1}^{n} w_i(\mathbf{x})\hat{y}_i(\mathbf{x}), \quad \sum_{i=1}^{n} w_i(\mathbf{x}) = 1\ . \tag{4}$$

The heuristic computation of the weights proposed in (Goel et al., 2007) is based on the PRESS estimation of the root mean square error, $PRESS_{RMS}$. The resulting average surrogate is called PRESS Weighted Average Surrogate (PWS). The weights computation is given by:

$$w_i = \frac{w_i^*}{\sum_{j=1}^{n} w_j^*}, \quad w_i^* = (E_i + \alpha E_{avg})^\beta\ , \tag{5}$$

$$E_{avg} = \frac{1}{n}\sum_{i=1}^{n} E_i, \quad \beta < 0, \alpha < 1, \tag{6}$$

where E_i is the $PRESS_{RMS}$ of the i-th surrogate model and recommended values of parameters are $\alpha = 0.05$ and $\beta = -1$.

3. Used surrogates and examples

The methodology was tested using several types of surrogate models with various inner settings which resulted in ten different meta-models in total. Their short descriptions are listed in Table 1.

Tab. 1: Used surrogate models.

Meta-model No.	Description
1	Kriging (zero order polynomial regression function, cubic correlation function)
2	Kriging (first order polynomial regression function, spherical correlation function)
3	newrbe (Radial Basis Neural Network available in MATLAB)
4	RBFN (Radial Basis Function Network with zero order polynomial regression function)
5	RBFN (Radial Basis Function Network with first order polynomial regression function)
6	RSM (Response Surface Methodology with first order polynomial regression function)
7	RBFN (Radial Basis Function Network without regression part)
8	PCE (Polynomial Chaos Expansion using Hermite polynomials of 3^{rd} degree)
9	PCE (Polynomial Chaos Expansion using Legendre polynomials of 3^{rd} degree)
10	RSM (Response Surface Methodology with second order polynomial regression function)

Seven benchmarks from (Andre et al., 2001) were used for testing. All of them have 2 input parameters which result in 2-dimensional design domain as shown in Figure 1.

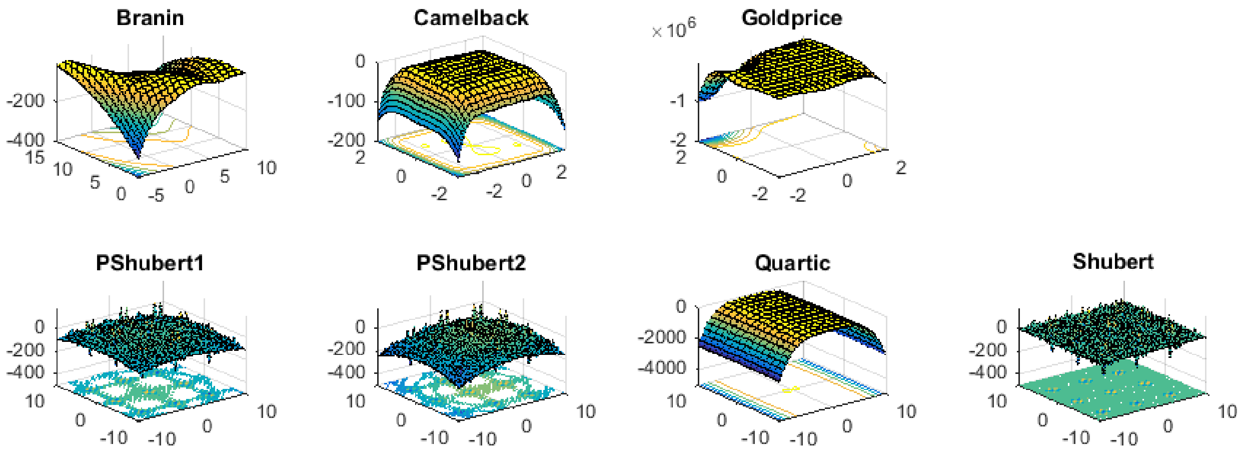

Fig. 1: Visualization of benchmark examples. Horizontal axes represent input parameters; vertical axis represents model's response.

4. Results and conclusions

The surrogate models were constructed based on 3 different optimized designs of experiments (with 11, 25 and 100 points). For each of them 100 random testing sets with 10^4, 10^3, 10^2 and 10 points were generated and used for computation of the *RMSE*. It was shown that the ranking of the meta-models remains almost the same no matter the number of training points or the testing points as depicted in Figure 2. Naturally, the spread of the values is significantly wider with decreasing number of testing points but the comparison of the surrogates' approximation quality is adequate even with the smallest testing set. Figure 3 shows the $PRESS_{RMS}$ obtained by the cross-validation of the training points which correspond to the weights of individual surrogates in the linear combination in the PWS. The results for the training set with 11 points were excluded because they would devalue the resulting graphs. The cross-validation led to very high errors in case of meta-models no. 8 and 9 (PCE). It is surprising because the *RMSE* for these meta-models come out well and ten training points used during the leave-one-out cross-validation is enough in case of the 3^{rd} degree of polynomials.

The stated simple testing on the PWS indicates possibility of usage of several surrogates at parallel and their follow-up combination. The results suggest that the weights for the linear combination could be computed using a small testing set rather than by the cross-validation based computation.

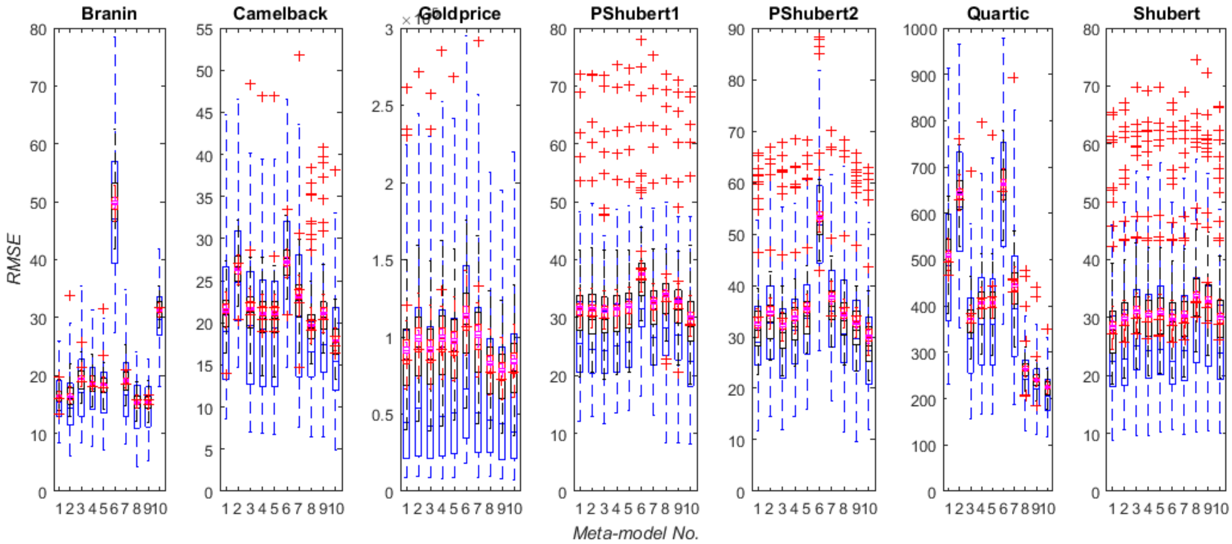

Fig. 2: Comparison of RMSE obtained by testing sets of different sizes (blue - 10^1, black - 10^2, red - 10^3, magenta - 10^4 points). Boxplots correspond to meta-models trained on the set with 11 points.

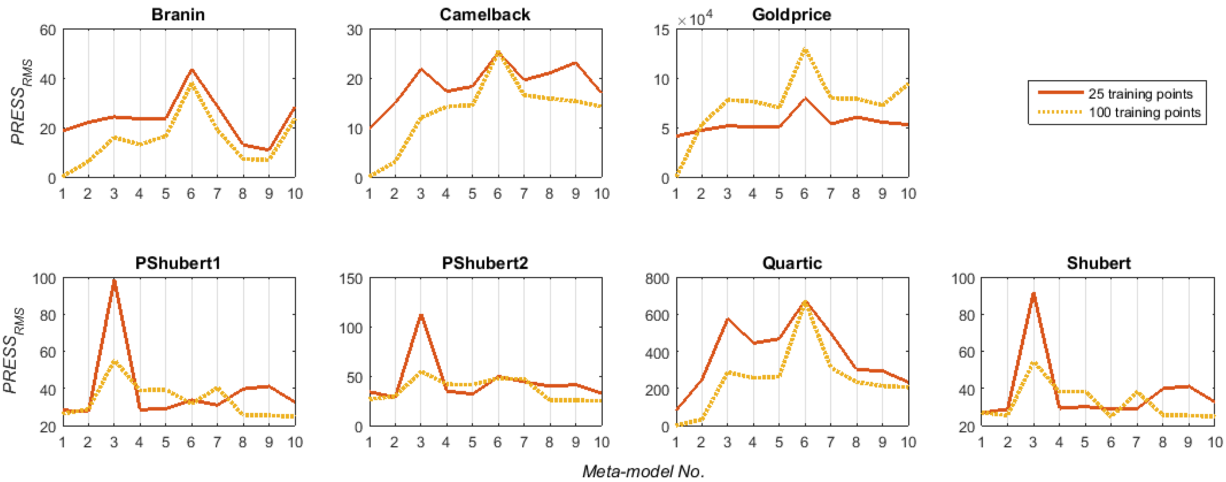

Fig. 3: Estimation of RMSE from PRESS vector obtained by leave-one-out cross-validation.

Acknowledgement

Financial support from the Ministry of Education, Youth and Sports of the Czech Republic, Project No. 8F15004, the Czech Science Foundation, Project No. 15-07299S, and the Grant Agency of the Czech Technical University in Prague, Project No. SGS16/037/OHK1/1T/11, is gratefully acknowledged.

References

Allen, D. M. (1974) The relationship between variable selection and data augmentation and a method for prediction. Technometrics, 16, pp.125-127.

Andre, J., Siarry, P. & Dognon, T. (2001) An improvement of the standard genetic algorithm fighting premature convergence in continuous optimization. Advances in Engineering Software, 32, 1, pp.49-60.

Goel, T., Haftka, R.Shyy, W. & Queipo, N. (2007) Ensemble of surrogates. Structural and Multidisciplinary Optimization, 33, 3, pp.199-216.

Iman, R. L. & Conover, W. J. (1980) Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. Communications in Statistics: Theory and Methods, A9, pp.1749-1842.

Montgomery, D. (2012) Design and Analysis of Experiments, 8th Edition. John Wiley & Sons, Incorporated.

Viana, F., Haftka, R. & Steffen, V. J. (2009) Multiple surrogates: How cross-validation errors can help us to obtain the best predictor. Structural and Multidisciplinary Optimization, 39, 4, pp.439-457.