Computers & Industrial Engineering 57 (2009) 996–1007

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Metamodels for variable importance decomposition with applications to probabilistic engineering design

Hemalatha Sathyanarayanamurthy [a,1], Ratna Babu Chinnam [b,*]

[a] UniBoring Company, Inc., 2555 Clark Street, Detroit, MI 48209, USA
[b] Industrial and Manufacturing Engineering Department, Wayne State University, 4815 Fourth Street, Detroit, MI 48202, USA

ARTICLE INFO

Article history:
Received 8 March 2009
Received in revised form 4 April 2009
Accepted 6 April 2009
Available online 12 April 2009

Keywords:
Probabilistic engineering design
Metamodels
Sensitivity analysis
FAST
Sobol
Kriging

ABSTRACT

It is routine in probabilistic engineering design to conduct modeling studies to determine the influence of an input variable (or a combination) on the output variable(s). The output or the response can then be fine-tuned by changing the design parameters based on this information. However, simply fine-tuning the output to the desired or target value is not adequate. Robust design principles suggest that we not only study the mean response for a given input vector but also the variance in the output attributed to noise and other unaccounted factors. Given our desire to reduce variability in any process, it is also important to understand which of the input factors affect the variability in the output the most. Given the significant computational overhead associated with most Computer Aided Engineering models, it is becoming popular to conduct such analysis through surrogate models built using a variety of metamodeling techniques. In this regard, existing literature on metamodeling and sensitivity analysis techniques provides useful insights into the various scenarios that they suit the best. However, there has been a limitation of studies that simultaneously consider the combination of metamodeling and sensitivity analysis and the environments in which they operate the best. This paper aims at contributing to reduce this limitation by basing the study on multiple metrics and using two test problems. Two test functions have been used to build metamodels, using three popular metamodeling techniques: Kriging, Radial-Basis Function (RBF) networks, and Support Vector Machines (SVMs). The metamodels are then used for sensitivity analysis, using two popular sensitivity analysis methods, Fourier Amplitude Sensitivity Test (FAST) and Sobol, to determine the influence of variance in the input variables on the variance of the output variables. The advantages and disadvantages of the different metamodeling techniques, in combination with the sensitivity analysis methods, in determining the extent to which the variabilities in the input affect the variabilities in the output are analyzed.

1. Introduction

Probabilistic engineering design offers tools for computing the uncertainty associated with the input and the output parameters/design variables of complex engineering simulation models. Probabilistic engineering designs include *reliability-based design* (Carter, 1997; Grandhi & Wang, 1998; Melchers, 1999; Wu & Wang, 1996) and *robust design* (Chen, Allen, Mistree, & Tsui 1996; Du & Chen, 2002a, 2002b; Parkinson, Sorensen, & Pourhassan 1993; Phadke, 1989; Taguchi, 1993). While reliability-based designs emphasize high reliability of a design by ensuring the probabilistic constraint satisfaction at desired levels, robust designs focus on making the design inert to the variations of system input through optimizing mean performance of the system and minimizing its variance simultaneously.

Given that probabilistic engineering design often involves simulation of complex engineering models with typically 10–20 variables, the computational costs for design optimization through complex computer-aided engineering models/simulations (for example using finite element models) are enormous, even with the latest improvements in computational algorithms and computing hardware. This makes it highly impractical to invest time, money and other resources in such computations of complex engineering problems. One alternative to reduce costs for design optimization is to use *surrogate models*, also known as *metamodels* in place of the actual simulation model to evaluate designs for optimality. A variety of metamodeling techniques exist; Polynomial Regression, Kriging (Sacks, Welch, Mitchell, & Wynn, 1989), Multivariate Adaptive Regression Splines (MARS), Radial-Basis Functions (RBF), Multi-layer Perceptron Networks (MLP), and Support Vector Machines (SVM) (Vapnik, 1995, 1998) to name a few. A comprehensive comparison of some of these metamodeling techniques

* Corresponding author. Tel.: +1 313 577 4846; fax: +1 313 578 5902.
E-mail addresses: sathyah@uniboring.com (H. Sathyanarayanamurthy), r_chinnam@wayne.edu (R.B. Chinnam).
[1] Tel.: +1 313 297 1005; fax: +1 313 849 0129.

0360-8352/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cie.2009.04.003

can be found in Jin, Chen, and Simpson (2000), Haykin (1999), Simpson, Allen, and Mistree (1998) and Simpson, Mauery, Korte, and Mistree (1998).

One of the important tasks of probabilistic engineering design is sensitivity analysis (Sudjianto, Du, & Chen, 2005), to determine how the variability associated with the system input affects the system output (Saltelli, Chan, & Scott, 2000). Modelers conduct sensitivity analysis for a number of reasons including the needs to determine which input parameters contribute the most to output variability (and possibly, require additional research to strengthen the knowledge base) (Hoffman & Hammonds, 1994), which parameters are insignificant, critical parameter interactions, optimal regions within the parameters space for use in a subsequent studies and so on.

Consider the example of an engine block and head joint sealing assembly studied in Awad, Sudjianto, and Singh (2004), where the objective was to optimize the head gasket design factors, to minimize the "gap lift" of the assembly as well as its sensitivity to manufacturing variation. To best simulate the engine assembly process and its operation, a finite element model shown in Fig. 1 was used. Given that the finite element model was computationally intensive, a surrogate model was employed and Sobol technique was used for the purposes of sensitivity analysis to achieve the desired objective.

A comprehensive review of the different sensitivity analysis methods, including their advantages and disadvantages, can be found in Helton, Johnson, Sallaberry, and Storlie (2006), Frey and Patil (2002). Some of the popular sensitivity analysis methods include SOBOL (Sobol, 1990a, 1990b) and FAST (Cukier, Fortuin, Schuler, Petschek, & Schaibly, 1973). A comparison of these methods can be found in Saltelli, Ratto, Tarantola, and Campolongo (2006), Saltelli, Tarontola, and Chan (1999), Saltelli and Bolado (1998), Chan, Saltelli, and Tarantola (1997), Saltelli, Andres, and Homma (1993).

Although existing studies on metamodeling and sensitivity analysis techniques provide useful insights into the various scenarios that they suit the best, there has been a limitation of studies that simultaneously consider the combination of metamodeling and sensitivity analysis and the environment in which they operate the best. These limitations were brought to our attention by the probabilistic engineering design group within the Product Development Division of Ford Motor Company. In an attempt to address these limitations, this paper aims at studying the combination of metamodeling and sensitivity analysis methods, based on multiple metrics using two test problems. The test functions have been used to build metamodels, using Kriging, radial-basis function (RBF) networks, and Support Vector Machines (SVMs). Sensitivity analysis is then conducted on these metamodels using FAST and Sobol techniques to determine the influence of the variabilities in the input on the variability of the output. The advantages and disadvantages of the different metamodeling techniques in combination with the sensitivity analysis methods, in determining the extent to which the variabilities in the input affect the variabilities in the output, are analyzed.

The paper is organized as follows: Section 2 provides an overview of the three metamodeling techniques used to develop surrogate models as well as the two sensitivity analysis methods. Section 3 explains the two test problems and data sampling schemes employed for the study. In Section 4, the performance of the different metamodeling and sensitivity analysis methods on the test problems is discussed. Finally, some recommendations and concluding remarks are offered in Section 5.

2. Overview of metamodels and sensitivity analysis techniques

This section first presents a concise overview of the three most popular metamodeling techniques under investigation, Kriging, RBFs, and Support Vector Machines. Then, it briefly describes the two sensitivity analysis techniques, FAST and Sobol.

2.1. Metamodels

The efficacy of any design optimization study largely depends on the accuracy with which a metamodel can capture the general (global) tendency of the design behavior. The accuracy for optimization under uncertainty also relies on the accuracy of the metamodel in capturing the performance variations, which could be caused by small perturbations of design parameters.

2.1.1. Kriging

Originally developed for spatial statistics and geostatistics, Kriging is an interpolative approximation method based on an exponentially weighted sum of the sample data. A graphical representation of this method, in particular, the correlation between the unknown data point and the known sample is illustrated in Fig. 2. Kriging models are very flexible due to the wide range of correlation functions that can be chosen for building the model. Depending on the type of correlation function used, Kriging model can either "honor the data", providing an exact interpolation of the data, or "smooth the data" in the presence of numerical noise.

A Kriging model postulates a combination of a polynomial model and departures of the form (Simpson, Allen, et al., 1998; Simpson, Mauery, et al., 1998):

$$y(x) = f(x) + z(x) \tag{1}$$

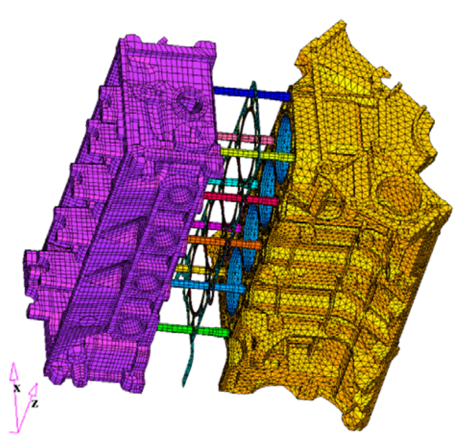

Fig. 1. Finite element model of head and block joint sealing assembly (source: Awad et al., 2004).

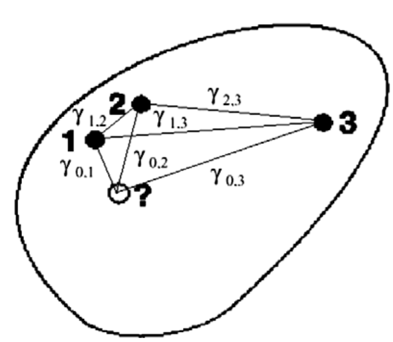

Fig. 2. Graphical representation of correlation between the unknown data point and the known sample (source: www.statios.com).