HYBRID AND UNCERTAINTY-BASED SURROGATE MODELING WITH APPLICATIONS TO WIND ENERGY

By

Jie Zhang

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: MECHANICAL ENGINEERING

The original of the complete thesis is on file
in the Rensselaer Polytechnic Institute Library

Examining Committee:

Prof. Achille Messac, Thesis Adviser
Prof. Luciano Castillo, Member
Prof. Prabhat Hajela, Member
Prof. Thomas C. Sharkey, Member

Rensselaer Polytechnic Institute
Troy, New York

July 2012
(For Graduation August 2012)

ABSTRACT

Surrogate models are extensively used in the analysis and the optimization of systems that involve computationally intensive simulation-based models. This dissertation develops a comprehensive methodology to accurately construct a surrogate and to quantify the uncertainty in the surrogate. A robust hybrid surrogate modeling method is first developed, which combines the favorable attributes of characteristically different surrogates using a weighted sum approach, in order to provide more accurate function estimation than provided by individual surrogates. This surrogate is called Adaptive Hybrid Functions (AHF). The AHF formulates a *reliable* trust region based on the crowding distance of the training points. A new *local measure of accuracy* is formulated using functions similar to Gaussian kernels, constructed within the trust region. The weights of the contributing surrogates are determined based on the *local measure of accuracy* estimated for each surrogate. This hybrid surrogate seeks to simultaneously capture the global trend of the function as well as the local deviations. The accuracy of the AHF is evaluated through applications in a variety of standard test problems and complex engineering problems. Surrogate-based infill optimization is also implemented using the AHF surrogate in conjunction with a recently developed adaptive sequential sampling method. Both local exploitation and global exploration aspects are considered for updating the surrogate during optimization.

In the second part of the methodology, we quantify the uncertainty in the surrogate introduced during the training process itself. In order to understand the fidelity of the proposed surrogate, the variation of errors over the design domain is characterized using a new technique, hereby called Domain Segmentation based on Uncertainty in the Surrogate. This technique segregates the design domain based on the level of cross-validation errors. The estimated errors are classified into physically meaningful classes based on the user's understanding of the system and/or the accuracy requirements for the concerned system analysis. In each class, the distribution of the cross-validation errors is estimated to represent the uncertainty in the surrogate. Support vector machine is implemented to determine the boundaries between error classes in the variable space, and to classify any new design (point) into a meaningful class. This method is uniquely helpful: (i) in improving the accuracy of surrogate modeling, and (ii) in enhancing the users' confidence in the application of the trained surrogates.

We successfully apply the surrogate modeling methodology to key aspects of wind

resource assessment and wind farm design. Using the AHF surrogate, we developed (i) a response surface-based cost model for wind farm design; and (ii) a response surface-based metric to assess the wind resource potential at a farm site. The response surface-based cost model provides interesting insights into cost variation with respect to critical engineering and economic parameters. This cost model addresses key interests of investors, project planning engineers, and policy makers in wind energy. The response surface-based wind resource potential metric can help decision makers in the assessment of wind farm sites and in the planning of the wind energy project. While the current resource potential metric (wind power density) only accounts for wind speed information, the proposed response surface-based metric yields more credible estimates by (i) exploiting the joint distribution of wind speed and direction; and (ii) estimating the (maximum) power generated by an optimized wind farm.