Bayesian Analysis (2010) **5**, Number 4, pp. 677–682

Comment on Article by Vernon et al.

Pritam Ranjan*

The article is very well written and the methodology used here is quite practical. I would like to congratulate the authors for a thoughtful uncertainty analysis of the Galform model. The paper addresses an inverse problem where the objective is to find the set of all 17-dimensional input values of the computer simulator Galform that leads to a pre-specified output (in this case the measured observational data on the number of galaxies in the universe (Norberg et al. (2002))). A few of the major challenges in this problem are (a) the input space of the computer simulator is high dimensional, (b) the simulator runs are expensive, and (c) the simulator outputs are functional. The authors use the idea of history matching to address some of these problems (Craig et al. (1997) present a nice review on history matching). However, the paper also raises a few interesting questions and should provide opportunities for further research.

Most of my comments are not necessarily specific to this case study and apply generally to the history matching problems for expensive computer simulators with functional outputs. Section 1 discusses my concerns on the simplification of the simulator output data structure while solving the history matching problem. Though the idea of history matching is in the literatures for the last three decades, it is not commonly used in the area of computer experiments (Section 2 suggests a strategy for making it more popular). Finally, some modeling and design aspects of the problem that can use further attention are discussed in Section 3.

1 Simplification of the data structure

The Galform model takes a 17-dimensional input to run (Table 1 of the paper describes the input space), and produces several outputs related to various physical characteristics of the simulated galaxies. Two types of simulator outputs (the (log) number of galaxies per unit volume with respect to the luminosity functions bj and K) were used in the case study. Both of these outputs are functional in nature, and the history matching problem requires matching these luminosity curves with the observed data (Norberg et al. (2002)). In the spirit of Craig et al. (1997), the authors simplify the history matching problem by choosing a set of seven scalar values (three bj and four K luminosity values) from the two functional output curves, and use the Gaussian process modeling approach to construct seven different scalar valued emulators for the (log) number of galaxies per unit volume at these seven luminosity values. That is, the original history matching problem with the simulator producing multiple functional outputs is simplified to finding common solutions of seven inverse problems for simulators with scalar responses and 17-dimensional inputs.

Such a simplification is very appealing and can be useful in practice, but it is not

*Department of Mathematics and Statistics, Acadia University, Wolfville, NS, Canada, mailto:pritam.ranjan@acadiau.ca

 DOI:10.1214/10-BA524B