Guest Editors' Introduction to the Special Issue on "Computer Models and Spatial Statistics for Environmental Science"

Brian J. REICH and Murali HARAN

The December 2011 issue of the Journal of Agricultural, Biological, and Environmental Statistics is on the topic "Computer models and spatial statistics for environmental science." This is a topic of great interest as the study of complex environmental phenomena increasingly relies on deterministic computer models. These models, for example regional climate models or rainfall-runoff simulators, are mathematical models that describe the evolution in time of a physical process. Usually, they consist of complex differential or partial differential equations that are not solvable in closed form. Hence, these are typically solved using numerical techniques, yielding deterministic predictions of a process. In this special issue, researchers tackle several important statistical problems that arise in the analysis of computer model output, for example calibrating model output with observed data, comparing and combing output from several computer models and physical observations, and building statistical emulators for computer models to predict the outcome of the models for new sets of input conditions.

An important contribution of statisticians in the analysis of deterministic models is to quantify uncertainty in inferences and predictions in rigorous fashion. Uncertainty quantification is of great interest, especially as information from complex computer models and messy observational data is used for decision making. There are several types of uncertainty, including (1) parametric uncertainty in the model's inputs or tuning parameters and (2) structural uncertainty in the mathematical equations that define the model. In "First-Order Emulator Inference for Parameters in Nonlinear Mechanistic Models", Mevin B. Hooten, William B. Leeds, Jerome Fiechter, and Christopher K. Wikle provide a computationally-efficient method for quantifying parametric uncertainty. They approximate the complicated computer model with a more tractable statistical model, and use

Brian J. Reich (✉) is Assistant Professor, Department of Statistics, North Carolina State University, 2311 Stinson Drive 4264 SAS Hall, Box 8203, Raleigh, NC 27695, USA (E-mail: *brian_reich@ncsu.edu*). Murali Haran is Associate Professor, Department of Statistics, Penn State, 326 Thomas Building, University Park, PA 16802, USA (E-mail: *mharan@stat.psu.edu*)

Journal of Agricultural, Biological, and Environmental Statistics, Volume 16, Number 4, Pages 451–452
DOI: 10.1007/s13253-011-0071-9

the statistical model to estimate model parameters. Structural uncertainty is addressed by "Quantifying simulator discrepancy in discrete-time dynamical simulators" by Richard D. Wilkinson, Michail Vrettas, Dan Cornford, and Jeremy E. Oakley and "Verifying reification with application to a rainfall-runoff computer simulator" by Leanna House. Both papers address structural uncertainty in rainfall-runoff simulators, but with very different approaches. Wilkinson et al. exploit the dynamic structure of the simulator and model discrepancy between consecutive time steps, whereas House utilizes expert knowledge about the effects of various model components.

Computer models often produce huge spatial or spatiotemporal fields that are too large to be analyzed using standard spatial statistical methods. Several papers address this crucial problem using reduced-rank spatial processes and Bayesian analysis via MCMC. "Climate Projections Using Bayesian Model Averaging and Space-Time Dependence" by K. Sham Bhat, Murali Haran, Adam Terando, and Klaus Keller and "Modeling Space-Time Dynamics of Aerosols Using Satellite Data and Atmospheric Transport Model Output" by Catherine A. Calder, Candace Berrett , Tao Shi, Ningchuan Xiao, and Darla K. Munroe use kernel-mixing methods to combine information from climate models and multiple air pollution models respectively. An alternative reduced-rank approach is predictive process modeling, which is utilized in "Comparing and blending regional climate model predictions for the American Southwest" by Esther Salazar, Bruno Sansó, Andrew O. Finley, Dorit Hammerling, Ingelin Steinsland, Xia Wang and Paul Delamater and "Improving crop model inference through Bayesian melding with spatially-varying parameters" by Andrew O. Finley, Sudipto Banerjee, and Bruno Basso to analyze climate and agricultural model output, respectively.

In addition to methods for melding huge spatial datasets, the special issue features new methods for analyzing non-standard spatial data, such as non-Gaussian and multivariate data. "Calibration of numerical model output using nonparametric spatial density functions" by Jingwen Zhou, Montserrat Fuentes, and Jerry Davis proposes a method for adjusting for model bias using observational data. The method aims to adjust not only the mean of the model output, but all quantiles simultaneously, which adjusts for model discrepancies in the variance and tail behavior. Computer model output is often high dimensional, and thus multivariate methods are useful to borrow strength across output dimensions. In "Multivariate Spatial Analysis of Climate Change Projections" by Tamara A. Greasby and Stephan R. Sain, the authors analyze changes in climate across the four seasons using a multivariate spatial model to pool information across both season and spatial location to identify differences between climate model projections.

In addition to some of the issues and methods outlined above, the papers in this special issue address a number of other interesting modeling and computational challenges that arise when using computer models in conjunction with physical observations. This special issue should therefore be of great interest not only to scientists working on related problems in environmental science but also, more generally, to the growing number of researchers developing statistical approaches for inference involving complex computer models.

[Published Online October 2011.]