Clemson Uncertainty Quantification (UQ) Interest Group

Whitney Huang

Clemson

February 3, 2020

About Clemson UQ Group

- Initiated by Dr. Andrew Brown in 2019 Fall.
- Founding group members: Drs. Qiong Zhang, Andrew Brown, Whitney Huang
- We think it is an interesting and promising research area and we would like to invite you to join us!

What is Uncertainty Quantification (UQ)?

$\boldsymbol{x} \in \mathcal{X}$ $\quad$ $f : \mathcal{X} \mapsto \mathcal{Y}$ $\quad$ $y = f(\boldsymbol{x})$

Input → Model → Output

Computer Model

- deterministic
- computationally extensive

- deterministic

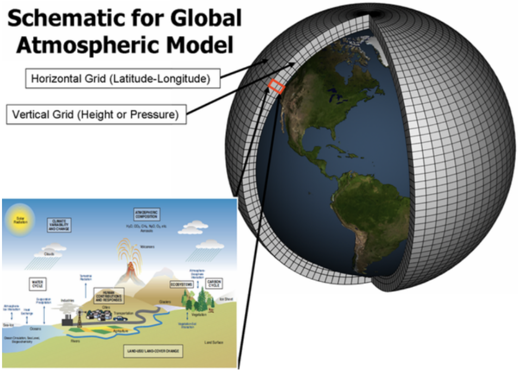
Statistical Science
1989, Vol. 4, No. 4, 409–435

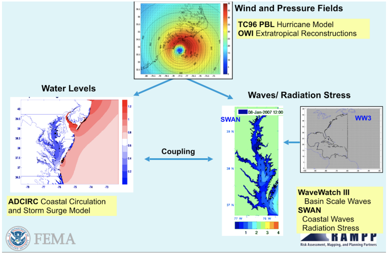
Design and Analysis of Computer Experiments

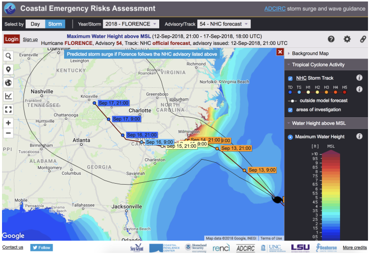
Jerome Sacks, William J. Welch, Toby J. Mitchell and Henry P. Wynn

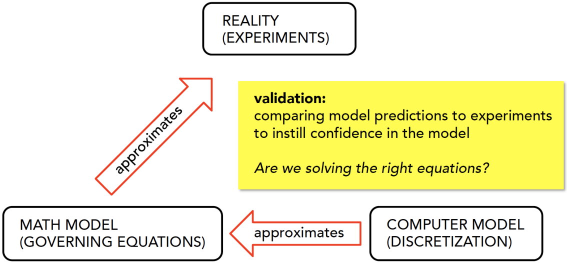
Abstract. Many scientific phenomena are now investigated by complex computer models or codes. A computer experiment is a number of runs of the code with various inputs. A feature of many computer experiments is that the output is deterministic—rerunning the code with the same inputs gives identical observations. Often, the codes are computationally expensive to run, and a common objective of an experiment is to fit a cheaper predictor of the output to the data. Our approach is to model the deterministic output as the realization of a stochastic process, thereby providing a statistical basis for designing experiments (choosing the inputs) for efficient prediction. With this model, estimates of uncertainty of predictions are also available. Recent work in this area is reviewed, a number of applications are discussed, and we demonstrate our methodology with an example.

Examples of Computer Models

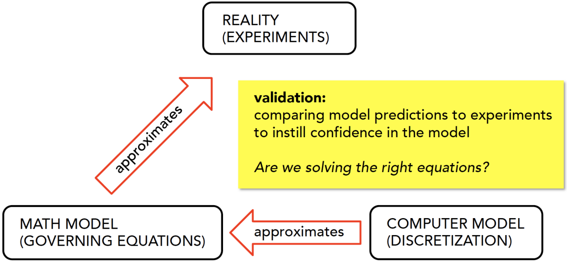








- *"Essentially, all models are wrong, but some are useful"*–George E.P. Box
- *"Experimental results are believed by everyone, except for the person who ran the experiment"*
- *"Computational results are believed by no one, except the person who wrote the code"*



- *"Essentially, all models are wrong, but some are useful"*–George E.P. Box
- *"Experimental results are believed by everyone, except for the person who ran the experiment"*
- *"Computational results are believed by no one, except the person who wrote the code"*

Computer Model Calibration

J. R. Statist. Soc. B (2001)
63, *Part* 3, *pp.* 425–464

Bayesian calibration of computer models

Marc C. Kennedy and Anthony O'Hagan

University of Sheffield, UK

[*Read before* The Royal Statistical Society *at a meeting organized by the* Research Section *on Wednesday, December 13th, 2000,* Professor P. J. Diggle *in the Chair*]

Summary. We consider prediction and uncertainty analysis for systems which are approximated using complex mathematical models. Such models, implemented as computer codes, are often generic in the sense that by a suitable choice of some of the model's input parameters the code can be used to predict the behaviour of the system in a variety of specific applications. However, in any specific application the values of necessary parameters may be unknown. In this case, physical observations of the system in the specific context are used to learn about the unknown parameters. The process of fitting the model to the observed data by adjusting the parameters is known as calibration. Calibration is typically effected by *ad hoc* fitting, and after calibration the model is used, with the fitted input values, to predict the future behaviour of the system. We present a Bayesian calibration technique which improves on this traditional approach in two respects. First, the predictions allow for all sources of uncertainty, including the remaining uncertainty over the fitted parameters. Second, they attempt to correct for any inadequacy of the model which is revealed by a discrepancy between the observed data and the model predictions from even the best-fitting parameter values. The method is illustrated by using data from a nuclear radiation release at Tomsk, and from a more complex simulated nuclear accident exercise.

MATH MODEL
(GOVERNING EQUATIONS)

$\xleftarrow{\text{approximates}}$

COMPUTER MODEL
(DISCRETIZATION)

verification:
ensuring the computer model is getting sufficiently close to the solution of the governing equations

Are we solving the equations right?

Statistical Science
2019, Vol. 34, No. 1, 1–22
https://doi.org/10.1214/18-STS660

Probabilistic Integration: A Role in Statistical Computation?[1]

François-Xavier Briol, Chris J. Oates, Mark Girolami, Michael A. Osborne and Dino Sejdinovic

Abstract. A research frontier has emerged in scientific computation, wherein discretisation error is regarded as a source of epistemic uncertainty that can be modelled. This raises several statistical challenges, including the design of statistical methods that enable the coherent propagation of probabilities through a (possibly deterministic) computational work-flow, in order to assess the impact of discretisation error on the computer output. This paper examines the case for probabilistic numerical methods in routine statistical computation. Our focus is on numerical integration, where a *probabilistic integrator* is equipped with a full distribution over its output that reflects the fact that the integrand has been discretised. Our main technical contribution is to establish, for the first time, rates of posterior contraction for one such method. Several substantial applications are provided for illustration and critical evaluation, including examples from statistical modelling, computer graphics and a computer model for an oil reservoir.

Journal of Global Optimization **13:** 455–492, 1998.

Efficient Global Optimization of Expensive Black-Box Functions

DONALD R. JONES[1], MATTHIAS SCHONLAU[2,⋆] and WILLIAM J. WELCH[3,⋆⋆]

[1]*Operations Research Department, General Motors R&D Operations, Warren, MI, USA;* [2]*National Institute of Statistical Sciences, Research Triangle Park, NC, USA;* [3]*Department of Statistics and Actuarial Science and The Institute for Improvement in Quality and Productivity, University of Waterloo, Waterloo, Ontario, Canada*

(Accepted in final form 30 June 1998)

Abstract. In many engineering optimization problems, the number of function evaluations is severely limited by time or cost. These problems pose a special challenge to the field of global optimization, since existing methods often require more function evaluations than can be comfortably afforded. One way to address this challenge is to fit response surfaces to data collected by evaluating the objective and constraint functions at a few points. These surfaces can then be used for visualization, tradeoff analysis, and optimization. In this paper, we introduce the reader to a response surface methodology that is especially good at modeling the nonlinear, multimodal functions that often occur in engineering. We then show how these approximating functions can be used to construct an efficient global optimization algorithm with a credible stopping rule. The key to using response surfaces for global optimization lies in balancing the need to exploit the approximating surface (by sampling where it is minimized) with the need to improve the approximation (by sampling where prediction error may be high). Striking this balance requires solving certain auxiliary problems which have previously been considered intractable, but we show how these computational obstacles can be overcome.

What is UQ?

One definition of "Capital UQ":

> "*The synergy between Statistics, Applied Mathematics, and domain sciences required to quantify uncertainties in inputs and the quantity of interest when models are too computationally complex to permit sole reliance on sampling-based methods*" – Ralph Smith, Distinguished University Professor, NCSU Math

- A Combined Physical-Statistical-Computational Approach to model input/output relationship
- Use statistical emulators to mimic (computationally extensive) simulators and to quantify its (epistemic) uncertainty

UQ Resources

- Statistical and Applied Mathematical Sciences Institute (SAMSI)
  - UQ Summer School (Link)
  - Model Uncertainty: Mathematical and Statistical (Link)
- SIAM UQ Activity Group (Link)
- Institute for Mathematics and its Applications (IMA)'s UQ workshop (Link)
- Isaac Newton Institute Uncertainty quantification for complex systems: theory and methodologies (Link)
- Robert Gramacy's new book Surrogates (Link)

Further Readings

- Santner, T. J., Williams, B. J., Notz, W.
  The Design and Analysis of Computer Experiments.
  Springer, 2003.
- Smith, R. C.
  Uncertainty quantification: theory, implementation, and applications.
  SIAM, 2014.
- Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P.
  Design and analysis of computer experiments
  Statistical science, 409–423, 1989
- Kennedy, M. C., & O'Hagan, A.
  Bayesian calibration of computer models (with Discussion)
  Journal of the Royal Statistical Society: Series B, 425–464, 2001