SciDAC Scientific Discovery through Advanced Computing

www.scidac.gov

Quantification of Uncertainty in Extreme Scale Computations

QUANTIFICATION OF UNCERTAINTY IN EXTREME SCALE COMPUTATIONS QUEST

SAND2012-3083C

www.quest-scidac.org

H. Najm[1], B. Debusschere[1], M. Eldred[1], C. Safta[1], J. Jakeman[1], D. Higdon[2], J. Gattiker[2], R. Ghanem[3], O. Knio[4], A. Alexanderian[5], Y. Marzouk[6], T. El Moselhy[6], J. Li[6], P. Conrad[6], T. Bui-Thanh[7], O. Ghattas[7], J. Martin[7], R. Moser[7], E. Prudencio[7], G. Stadler[7]

[1]Sandia National Laboratories, [2]Los Alamos National Laboratory, [3]University of Southern California, [4]Duke University, [5]The Johns Hopkins University, [6]Massachusetts Institute of Technology, [7]The University of Texas at Austin

SciDAC QUEST Institute

- QUEST is a SciDAC institute focused on uncertainty quantification (UQ) in large-scale scientific computations.
- It is a collaboration among six institutions with a history of in-depth collaborations on the development, implementation, and use of UQ algorithms/software in challenging high-performance computing environments.
- Our members have developed and maintain a range of UQ software products that have been applied in extreme scale computational environments, with challenging scientific application codes, including climate, geophysics, & combustion.

Key Objectives

- Delivering expertise, advice, and state of the art UQ tools to SciDAC application partnership projects utilizing extreme scale computations on advanced computational architectures.
- Shepherding forward our repertoire of UQ theory, algorithms, and software, and enhancing their robustness/effectiveness for relevant benchmark problems in extreme-scale computational settings

Member Institutions

- Sandia National Laboratories
- Los Alamos National Laboratory
- University of Southern California
- Duke University
- Massachusetts Institute of Technology
- University of Texas at Austin

Vision

Our vision encompasses all aspects of UQ in computational modeling, namely:

- Well-founded setup of the UQ problem
- Characterization of the input space given available data/information
- Local and global sensitivity analysis
- Adaptive dimensionality and order reduction
- Forward and inverse propagation of uncertainty
- Handling of application code failures, missing data, and fault tolerance
- Model comparison, validation, selection, and averaging

The nature of the UQ problem requires the seamless combination of data, models, and information across this whole landscape in a manner that provides a self-consistent quantification of requisite uncertainties in predictions from computational models. Accordingly, our UQ methods and tools span an interdisciplinary space across applied math, information theory, and statistics.

Motivation for UQ in Computational Modeling

- Quantification of uncertainty in computational predictions/inferences.
- Model validation and comparison, hypothesis testing, design optimization, and decision support.
- Information on global sensitivity of the computational model over the range of uncertainty in its parameters.
- Extraction of enhanced understanding about a physical system/model from computations.
- Guiding experimental studies in optimal directions that maximize information gain by targeting parameters with maximal contributions to uncertainty in predictions.

UQ Software Products

Key UQ software tools that are included in the QUEST universe are:

- DAKOTA (SNL) provides a variety of non-intrusive algorithms for design optimization, model calibration, uncertainty quantification, global sensitivity analysis, solution verification, and parameter studies. (dakota.sandia.gov)
- UQTk (SNL) is a lightweight C++ library, offering intrusive/non-intrusive Polynomial Chaos UQ tools. (www.sandia.gov/UQToolkit)
- QUESO (UT) is an MPI/C++ library that provides statistical algorithms for Bayesian inference, model calibration, model validation, and decision making under uncertainty.
- GPMSA (LANL) focuses on Bayesian inference, using a Gaussian process (GP) response surface, trained from an ensemble of forward model runs, to minimize the number of forward model calls required in the inference.

Software Development

- Pursuing enhanced software interoperability among DAKOTA, QUESO, GPMSA
- Pursuing enhanced functionality:
  - DAKOTA: Generalized sparse grids, adaptive h-refinement; Compressive sensing, L1 constraints; random fields.
  - QUESO: incorporating GP functionalities, improving user support
  - GPMSA: improving user support and tutorial material
  - UQTk: adding non-intrusive capabilities, random fields; matlab version.

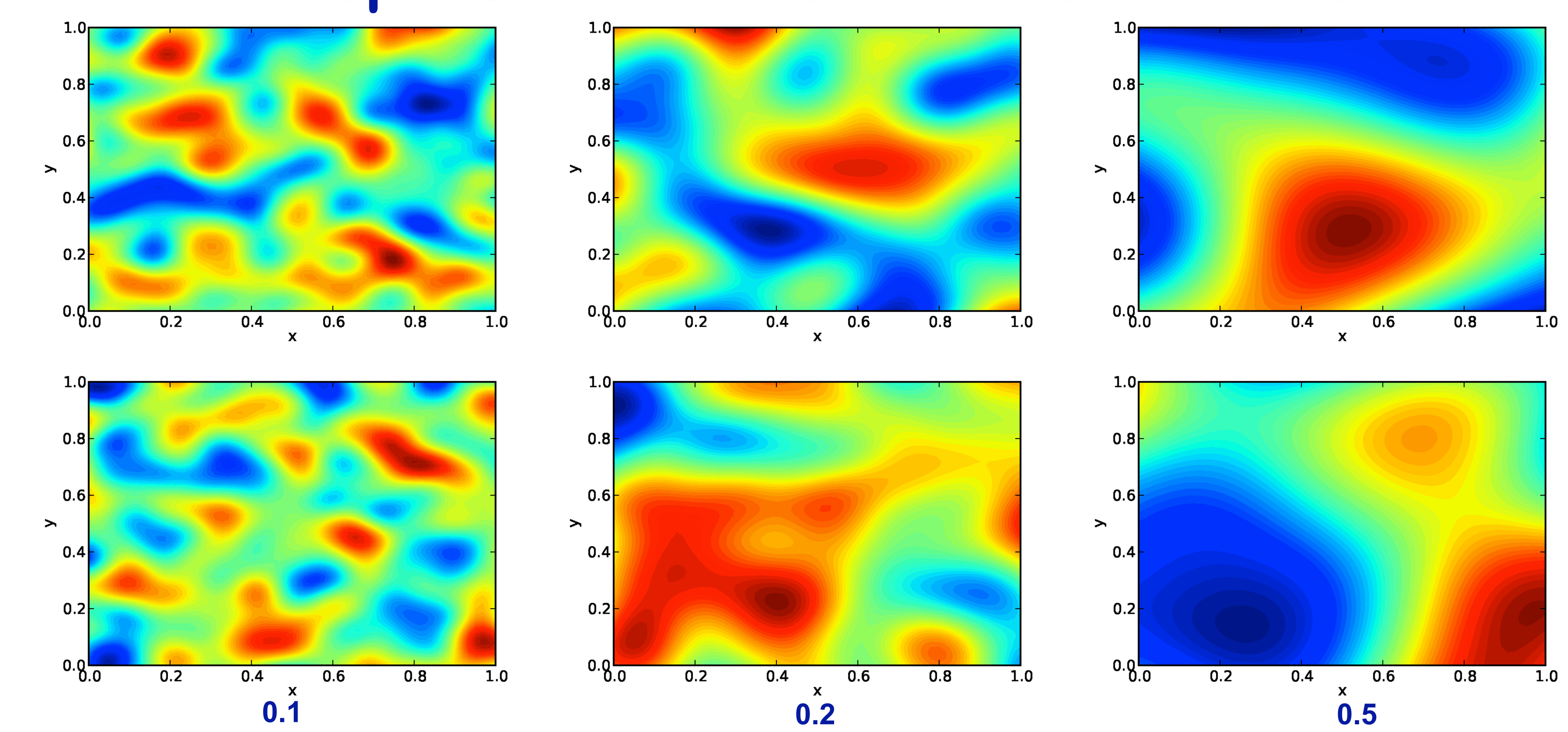
Algorithmic Development

- Implementation of Bayesian additive regression trees (BART) for response surface/surrogate construction
- Development of stochastic preconditioning methods for sparse representation of stochastic dynamical system predictions
- Development of methods for adaptive iterative data-driven surrogate construction targeting the support of the posterior
- Parallelization and efficient implementation of optimal map methods for Bayesian inference without MCMC
- Development of adaptive generalized sparse quadrature methods, with application to uncertain ocean dynamics in the Gulf of Mexico
- Development of reduced representations of high-dimensional uncertain model outputs, adapted to specific quantities of interest
- Development of methods for the solution of extreme-scale statistical inverse problems, with application in seismic inversion, ice sheet dynamics, mantle convection and porous media flows.

Impact

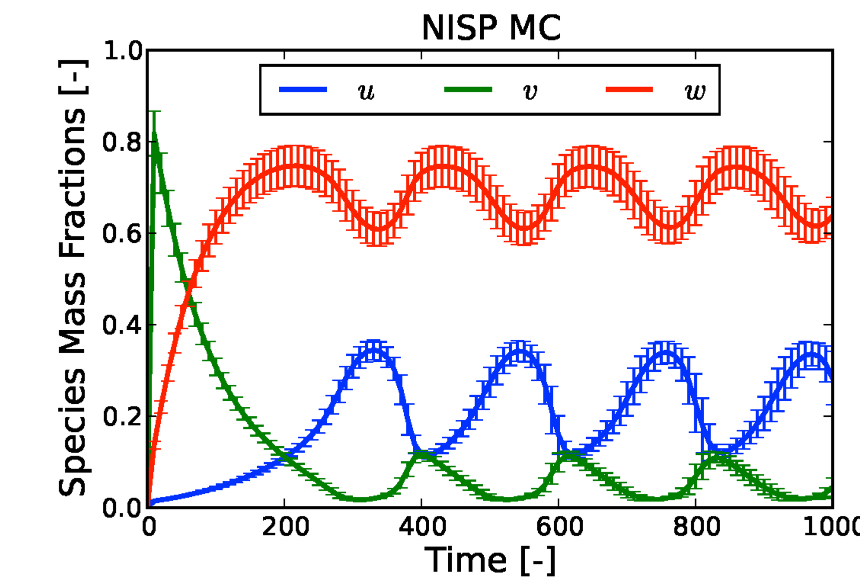
- Raise awareness of uncertainty in computational models and the need for UQ in scientific computations
- Lead to closer integration between experimental activities and extreme-scale scientific computing
- Closer integration and synergies between applied math and statistics in computational science
- Advance the state of the art in UQ theory, algorithms, and software, targeting extreme scale computations
- Enable UQ in a number of science application partnership projects

Representation of Random Fields



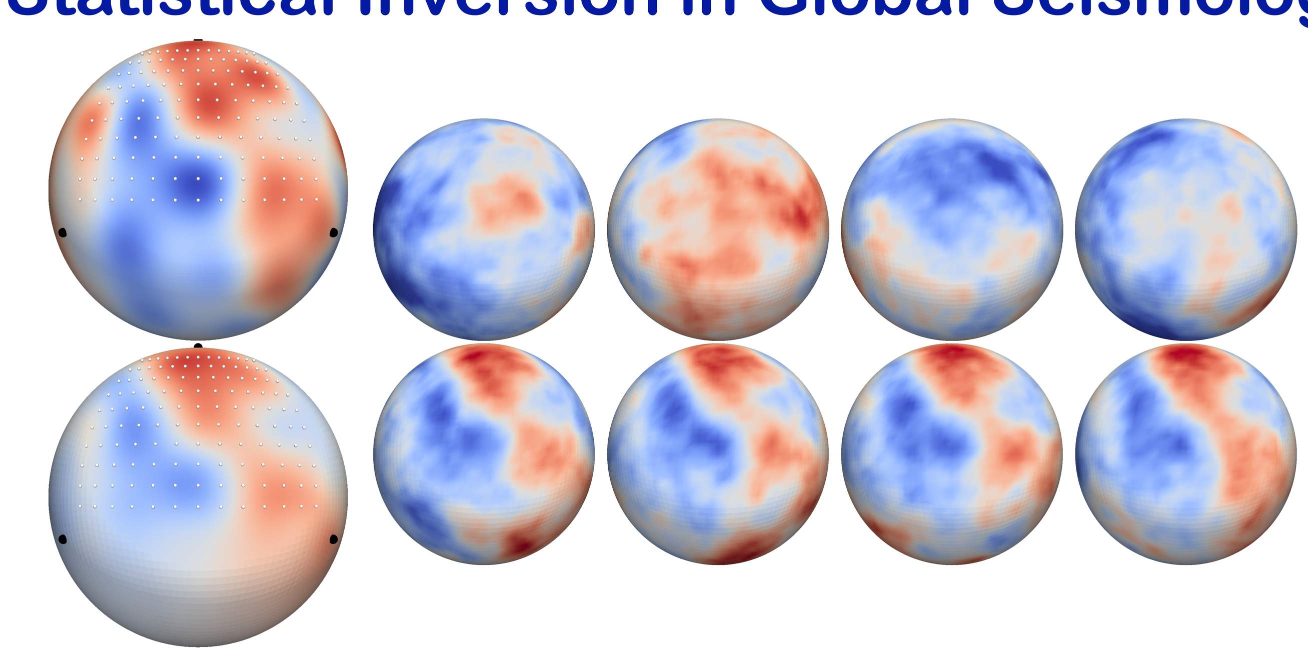
The Karhunen-Loeve expansion is used to generate realizations from a Gaussian random field over a range of correlation lengths. Two random realizations are shown for fields with correlation lengths of 0.1, 0.2, and 0.5, as indicated, highlighting the different range of length-scales in each case.

UQ in Chemical Systems



Non-intrusive Polynomial Chaos UQ application in a surface-reaction ODE model, with uncertainty in one chemical rate parameter.

Statistical Inversion in Global Seismology



Application of a low-rank based algorithm to the solution of a large-scale statistical inverse problem in global seismology. The left column shows a comparison of a "truth" earth wave speed field (top) with the MAP estimate of the posterior PDF (bottom). Locations of three earthquake sources are indicated by black dots, and locations of 130 seismometers in a quarter of the Northern hemisphere are indicated by white dots. The MAP estimate captures the structure of the truth wave speed reasonably well in the vicinity of observations. Eight images in the 2nd to 5th columns depict samples from the prior (top) and posterior (bottom) distributions. The difference between the top and bottom rows reflects the information gained from the observations. *Support for this work was also provided by AFOSR (Computational Math), DOE (ASCR Applied Math), and NSF (CMG and CDI).*

This work was supported by the Scientific Discovery through Advanced Computing (SciDAC) program funded by the US Department of Energy Office of Advanced Scientific Computing Research. Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

Sandia National Laboratories · Los Alamos NATIONAL LABORATORY EST. 1943 · USC University of Southern California · Duke UNIVERSITY · Massachusetts Institute of Technology · THE UNIVERSITY OF TEXAS AT AUSTIN