Uncertainty Quantification (UQ) and Sensitivity Analysis (SA) in GDSA

U.S. Nuclear Waste Technical Review Board
Fall 2021 Board Meeting
November 03-04, 2021
Virtual Meeting

Laura Swiler
Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.
Outline

- What are the objectives and strategy for developing uncertainty and sensitivity analysis tools for GDSA Framework?
- International collaboration
- What UQ and SA tools have been incorporated into GDSA Framework?
- Describe examples of how these UQ/SA tools have been applied to reference case simulations (crystalline reference case).
Objectives/Strategy for UQ/SA
Objectives and Strategy

- **Use well-established methods** for the conceptual and computational framework for UQ/SA in performance assessment
  - Allow for treatment of epistemic and aleatory uncertainty
  - Use approaches that address regulatory requirements
  - Use Latin Hypercube Sampling (LHS), correlation coefficients, scatterplots, and regression
  - Leverage existing algorithms implemented by the Dakota team and others

- **Keep abreast of new UQ/SA methods**
  - **Use variance-based sensitivity analysis**, which has become a standard approach.
  - **Use surrogate models** to explore the input parameter space of expensive simulations (in computational time and labor)
  - Develop methods that allow efficiency gains and extract information (multi-fidelity models)

- **Maintain leadership in UQ/SA** for geologic repository performance assessment
  - **Participate in an international working group** on sensitivity analysis
International collaboration
The Joint Sensitivity Analysis (JOSA) Group

- JOSA is an informal ad-hoc group:
  - dedicated to sensitivity analyses (SA) in the context of geologic disposal of radioactive waste. This includes **exchanging information on sensitivity analysis and methods**.
  - emerged from earlier bi-/trilateral activities (US, Germany, UK),
  - is being informally supported by OECD/NEA's Integration Group for the Safety Case

- Participants: GRS (Germany), Posiva & FORTUM (Finland), SCK-CEN (Belgium), Sandia (USA), TUC (Germany), and IBRAE (Russia).
We carried out comparative sensitivity analyses. Existing datasets were provided by “case owners” and analysed by various participants.

Outline
1. Introduction
2. Sensitivity Analysis Methods
3. Calculation Case Selection
4. GRS Clay Case
5. SNL Shale Case
6. Dessel Case
7. IBRAE Groundwater Case
8. Summary
<table>
<thead>
<tr>
<th>Sensitivity analysis approaches investigated by JOSA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graphical</strong></td>
</tr>
<tr>
<td>Scatterplots</td>
</tr>
<tr>
<td>Cumulative Sum of Normalized Reordered Output (CUSUNORO)</td>
</tr>
<tr>
<td><strong>Correlation &amp; Regression Analysis</strong></td>
</tr>
<tr>
<td>Pearson Correlation &amp; Partial Correlation</td>
</tr>
<tr>
<td>Spearman Rank Correlation &amp; Partial Rank Correlation</td>
</tr>
<tr>
<td>Regression Coefficients (Linear, Rank, Stepwise)</td>
</tr>
<tr>
<td><strong>Variance-based</strong></td>
</tr>
<tr>
<td>Sobol’ Indices</td>
</tr>
<tr>
<td>Fourier Amplitude Sensitivity Test (FAST), extended FAST (eFAST)</td>
</tr>
<tr>
<td>Effective Algorithm for Sensitivity Indices, Cosine Sensitivity (EASI, COSI)</td>
</tr>
<tr>
<td>Random Balance Designs</td>
</tr>
<tr>
<td><strong>Moment-independent</strong></td>
</tr>
<tr>
<td>Borgonovo's $\delta$</td>
</tr>
<tr>
<td>Pianosi and Wagener (PAWN)</td>
</tr>
</tbody>
</table>
Correlation coefficients and linear regression approaches continue to be used and are informative.

**The first order variance-based index estimates are now easily generated using a variety of approaches and are a main SA approach.**

- Results showed the same most important parameter but differed on lower ranked parameters
- Surrogate model type may play a role in accuracy of SA indices

Data transformations may be employed for variables which vary over orders of magnitude.

Graphical methods such as CUSUNORO also provide additional visualization which can show influences over the range of a variable.

**Bottom line: the international group is a valuable way for us to collaborate with and learn from the international community. We plan to continue this effort with a set of additional case studies in 2022.**
UQ/SA capabilities in GDSA
Computational Workflow

Next Gen Workflow

- Input Parameters
  - Parameter database

Uncertainty Sampling and Sensitivity Analysis

Computational Support
- Processing
  - Vorocrust
  - Python
- Visualization
  - ParaView
  - Python

Multi-Physics Simulation and Integration

- Source Term and EBS Evolution Model
  - Inventory
  - Decay, ingrowth
  - WF degradation
  - WP degradation
  - Radionuclide release
  - Thermal, mechanical
  - Gas generation

- Flow and Transport Model
  - Advection, diffusion, dispersion
  - Discrete fracture networks
  - Multiphase flow
  - Sorption, solubility, colloids
  - Isotope partitioning
  - Decay, ingrowth
  - Thermal effects
  - Chemical reactions

- Biosphere Model
  - Exposure pathways
  - Uptake/transfer
  - Dose calculations

Results
Dakota is a long-standing software framework (27 years) developed to perform parameter studies, optimization, etc. with computationally expensive codes.

Flexible interface to simulation codes: one interface; many methods

Continual advanced algorithm R&D to tackle computational challenges

Scalable parallelism on a variety of platforms

Publicly available: https://dakota.sandia.gov
UQ and SA methods in Dakota (methods in red used in GDSA)

Uncertainty Quantification
- **Sampling**
  - Monte Carlo
  - Latin Hypercube sampling
  - Quasi MC
  - Structured experimental designs
  - Parameter studies
- **Reliability Methods**
  - FORM/SORM
- **Stochastic expansions**
  - Polynomial chaos
- **Epistemic methods**
  - Nested Sampling, Interval bounds analysis, Dempster-Shafer
- **Multifidelity UQ**

Sensitivity Analysis
- **Correlation**
  - Pearson (on raw data)
  - Rank (Spearman)
  - Partial on both raw and rank data
- **Graphical methods**
  - Scatterplots
- **Variance-based sensitivity (Sobol’ indices)**
  - From sampling only
  - From surrogate analysis
- **Morris One-At-A-Time Methods**
- **Surrogates:**
  - Polynomial regression
  - Gaussian processes
  - Polynomial Chaos Expansion
  - Mars, NN, others
Epistemic/Aleatory Nested Sampling Capability

Epistemic: Lack of knowledge about the appropriate value to use for a quantity; reducible
Aleatory: Inherent variability, randomness, irreducible

Three areas of research focus:

- **Variance-based Decomposition (VBD):** Sobol’ Indices are sensitivity indices which summarize how response variability can be apportioned to individual input factors.
  - **Main effect** $S_i$ measures the effect of varying $x_i$ alone (averaging over other factors).
  - **Total effect** $T_i$ measures the effect of varying $x_i$ including its interactions with other variables.
  - The calculations require repeated sets of samples: this is very expensive. Surrogates are typically used to calculate these indices.

- **Polynomial Chaos expansion**
  - Uses an orthogonal polynomial approximation of the response
  - Analytically calculates statistics from the approximation instead of approximating the statistics with MC samples (makes it easy to obtain estimates for the Sobol’ indices!)

- **Multifidelity uncertainty quantification methods**
  - Exploit an *ensemble of models* with varying fidelities and cost to achieve *greater statistical accuracy* at *less computational cost*.
  - Ideas rooted in control variates and variance reduction
Multifidelity Results

Horizontal slices of the permeability tensor in the x-direction for meshes with cell sizes $d = 10$, 20, and 40 m, from left to right.

A vertical slice of the simplified crystalline domain, taken at $y = 500$ m.

Horizontal slices of the permeability tensor in the x-direction for meshes with cell sizes $d = 10$, 20, and 40 m, from left to right.
Application of UQ/SA tools to GDSA Reference Case
The model domain is approximately 3000 m in length, 2000 m in width, and 1260 m in height.

The repository is located at a depth of 585 m. Forty-two disposal drifts contain 40 12-PWR waste packages each (1680 total waste packages).

Drifts are backfilled with bentonite buffer and are surrounded by a 1.67-m thick DRZ.

The model domain contains 4.8 million cells.

Within the repository, grid cells are as small as 1.67-m on a side; elsewhere grid cells are 15-m on a side.
Crystalline Reference Case

- Used the dfnWorks software (from Los Alamos: https://dfnworks.lanl.gov/) to generate the discrete fracture networks
- These were meshed in Cubit and the simulation was run in PFLOTRAN (https://www.pflotran.org/)
Performed nested sampling, outer loop represented DFNs, inner loop represented epistemic parameters. 1000 PFLOTRAN runs.

### Measures of Spatial Heterogeneity

<table>
<thead>
<tr>
<th>DFN Graph Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>STT</td>
<td>The relative shortest travel time between repository and aquifer.</td>
</tr>
<tr>
<td>aveDegree</td>
<td>Average number of intersections per fracture. A measure of how connected the network is over the entire domain.</td>
</tr>
<tr>
<td>Intersections</td>
<td>Number of fractures intersecting the repository. A measure of number of potential flow pathways out of the repository region.</td>
</tr>
</tbody>
</table>

### Epistemic Variables

<table>
<thead>
<tr>
<th>Input</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rateUNF</td>
<td>Fractional dissolution rate of spent (used) nuclear fuel</td>
</tr>
<tr>
<td>kGlacial</td>
<td>Glacial till permeability</td>
</tr>
<tr>
<td>pBuffer</td>
<td>Buffer porosity</td>
</tr>
<tr>
<td>permDRZ</td>
<td>DRZ permeability</td>
</tr>
<tr>
<td>permBuffer</td>
<td>Buffer permeability</td>
</tr>
<tr>
<td>meanWPrate</td>
<td>Mean of the waste package corrosion rate</td>
</tr>
<tr>
<td>stdWPrate</td>
<td>Standard deviation of the waste package corrosion rate</td>
</tr>
<tr>
<td>IRF</td>
<td>Instant release fraction</td>
</tr>
</tbody>
</table>

### Quantities of Interest (QoIs)

<table>
<thead>
<tr>
<th>QoI</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak_Total I129_M</td>
<td>Maximum I-129 concentration in the aquifer [M]</td>
</tr>
<tr>
<td>Fractional Mass Flux from Repo_1Myr</td>
<td>The instantaneous fractional loss rate of tracer remaining in repository at one million years. It is an indicator of repository retention.</td>
</tr>
<tr>
<td>Rock Aq_Rock Eb_1Myr</td>
<td>This is the ratio of two water fluxes (upward vs. horizontal flow): the rock to the aquifer vs. the rock to the east boundary at 1 million years</td>
</tr>
<tr>
<td>Fraction of Spike in Repository_1Myr</td>
<td>The fraction of a tracer remaining in repository at 1 million years. It is an indicator of repository retention.</td>
</tr>
</tbody>
</table>
Crystalline Reference Case Results

Adding graph metrics to the SA significantly changes the results, showing the influence of DFNs. The fracture network and where fractures land has a larger effect on peak 129-I than source term and EBS uncertainties.
Crystalline Reference Case Results

Analysis for Max $^{129}\text{I}$ Concentration in Aquifer (M)

Capability to plot sensitivity indices as a function of time gives us additional insight and physical interpretation.
Next steps/additional tools and methods

- Continue investigation into advanced sensitivity analysis methods
- Additional work on multifidelity methods, especially with respect to models having different spatial representations of the discrete fracture network
- Investigation into efficient methods for estimating tail probabilities
- Methods to assess surrogate accuracy on the fly

In summary:
- We have focused on UQ/SA capability development. **We have a rich set of capabilities**, including established methods, variance-based indices, and surrogates.
- We have applied these capabilities to a variety of cases.
- The references cases have been very useful for demonstrating certain features we need to address, such as spatial heterogeneity from the discrete fracture networks.
- Sensitivity analysis is useful for helping understand the behavior and the importance of processes evolving over time within the models.

Swiler, Laura P. and Dirk-Alexander Becker, Dusty Brooks, Joan Govaerts, Lasse Koskinen, Pekka Kupiainen, Elmar Plischke, Klaus-Jürgen Röhlig, Elena Saveleva, Sabine M. Spiessl, Emily Stein, Valentina Svitelman. “Sensitivity Analysis Comparisons on Geologic Case Studies: An International Collaboration.” SAND2021-11053. Note: this is the report that is the result of the international SA working group comparisons on four case studies.

Swiler, L.P., E. Basurto, D.M. Brooks, A.C. Eckert, R. Leone, P.E. Mariner, T. Portone, M. L. Smith and E.R. Stein. “Uncertainty and Sensitivity Analysis Methods and Applications in the GDSA Framework (FY2021).” SAND2021-9903R. This document describes the most recent version of multifidelity UQ methods, the DFN analysis, and the crystalline reference case including the plots showing Sobol’ indices over time.


Additional References