

# ***Uncertainty and Sensitivity Analysis Methods and Applications in the GDSA Framework (FY2021)***

Fuel Cycle Research and Development

*Prepared for  
U.S. Department of Energy  
Spent Fuel and Waste Science and  
Technology Campaign*

*L.P. Swiler, E. Basurto, D.M. Brooks, A.C. Eckert,  
R. Leone, P.E. Mariner, T. Portone, M. L. Smith and  
E.R. Stein*

*Sandia National Laboratories  
August 13, 2021*

M3SF-  
21SN010304042  
SAND2021-9903R



#### **DISCLAIMER**

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.



**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.

## APPENDIX E

### NFCSC DOCUMENT COVER SHEET<sup>1</sup>

Name/Title Uncertainty and Sensitivity Analysis Methods and Applications in the GDSA Framework (FY2021)  
 Deliverable/Milestone/Revision No. M3SF-21SN010304042  
 Work Package Title and Number GDSA-Uncertainty and Sensitivity Analysis Methods-SNL SF-21SN01030404  
 Work Package WBS Number 1.08.01.03.04  
 Responsible Work Package Manager Laura Swiler *Laura Swiler*  
 (Name/Signature)  
 Date Submitted August 13, 2021

Quality Rigor Level for Deliverable/Milestone <sup>2</sup>	<input type="checkbox"/> QRL-1 <input type="checkbox"/> Nuclear Data	<input type="checkbox"/> QRL-2	<input checked="" type="checkbox"/> QRL-3	<input type="checkbox"/> QRL-4 Lab QA Program <sup>3</sup>
--	---	--------------------------------	---	---

This deliverable was prepared in accordance with Sandia National Laboratories  
 (Participant/National Laboratory Name)

QA program which meets the requirements of  
 DOE Order 414.1       NQA-1       Other

**This Deliverable was subjected to:**

Technical Review

**Technical Review (TR)**

**Review Documentation Provided**

- Signed TR Report or,  
 Signed TR Concurrence Sheet or,  
 Signature of TR Reviewer(s) below

**Name and Signature of Reviewers**

Gianluca Geraci *Gianluca Geraci* Chapter 2  
Tara LaForce *Tara LaForce* Chapter 3

Peer Review

**Peer Review (PR)**

**Review Documentation Provided**

- Signed PR Report or,  
 Signed PR Concurrence Sheet or,  
 Signature of PR Reviewer(s) below

Laura Price *Laura Price* Chapter 4  
Michael Nole *Michael Nole* Chapter 5

**NOTE 1:** Appendix E should be filled out and submitted with the deliverable. Or, if the PICS:NE system permits, completely enter all applicable information in the PICS:NE Deliverable Form. The requirement is to ensure that all applicable information is entered either in the PICS:NE system or by using the NFCSC Document Cover Sheet.

- In some cases there may be a milestone where an item is being fabricated, maintenance is being performed on a facility, or a document is being issued through a formal document control process where it specifically calls out a formal review of the document. In these cases, documentation (e.g., inspection report, maintenance request, work planning package documentation or the documented review of the issued document through the document control process) of the completion of the activity, along with the Document Cover Sheet, is sufficient to demonstrate achieving the milestone.

**NOTE 2:** If QRL 1, 2, or 3 is not assigned, then the QRL 4 box must be checked, and the work is understood to be performed using laboratory QA requirements. This includes any deliverable developed in conformance with the respective National Laboratory / Participant, DOE or NNSA-approved QA Program.

**NOTE 3:** If the lab has an NQA-1 program and the work to be conducted requires an NQA-1 program, then the QRL-1 box must be checked in the work Package and on the Appendix E cover sheet and the work must be performed in accordance with the Lab's NQA-1 program. The QRL-4 box should not be checked.

## **ACKNOWLEDGEMENTS**

This report has benefited from the insights and assistance of many. The authors thankfully acknowledge Michael Nole for his contributions to PFLOTRAN and supporting many advanced features for the crystalline reference case. Participants in the international uncertainty and sensitivity analysis collaboration, especially Klaus Röhlig (TU Clausthal) and Dirk Becker (GRS), provided valuable discussion and perspective. The authors gratefully acknowledge the significant contributions of S. David Sevougian (retired) in repository science and of Jon Helton in performance assessment methodology. The authors appreciate and thank the technical reviewers for comments and suggestions that improved the quality of the manuscript. Finally, the authors give a special thanks to Prasad Nair and Tim Gunter of the Department of Energy's Spent Fuel and Waste Science and Technology campaign for their on-going support and guidance.

## EXECUTIVE SUMMARY

The Spent Fuel and Waste Science and Technology (SFWST) Campaign of the U.S. Department of Energy (DOE) Office of Nuclear Energy (NE), Office of Fuel Cycle Technology (FCT) is conducting research and development (R&D) on geologic disposal of spent nuclear fuel (SNF) and high-level nuclear waste (HLW). Two high priorities for SFWST disposal R&D are design concept development and disposal system modeling. These priorities are directly addressed in the SFWST *Geologic Disposal Safety Assessment* (GDSA) control account, which is charged with developing a geologic repository system modeling and analysis capability, and the associated software, *GDSA Framework*, for evaluating disposal system performance for nuclear waste in geologic media. *GDSA Framework* is supported by SFWST Campaign and its predecessor the Used Fuel Disposition (UFD) campaign.

This report fulfills the GDSA Uncertainty and Sensitivity Analysis Methods work package (SF-21SN01030404) level 3 milestone, *Uncertainty and Sensitivity Analysis Methods and Applications in GDSA Framework (FY2021)* (M3SF-21SN010304042). It presents high level objectives and strategy for development of uncertainty and sensitivity analysis tools, demonstrates uncertainty quantification (UQ) and sensitivity analysis (SA) tools in *GDSA Framework* in FY21, and describes additional UQ/SA tools whose future implementation would enhance the UQ/SA capability of *GDSA Framework*. This work was closely coordinated with the other Sandia National Laboratory GDSA work packages: the GDSA Framework Development work package (SF-21SN01030405), the GDSA Repository Systems Analysis work package (SF-21SN01030406), and the GDSA PFLOTRAN Development work package (SF-21SN01030407). This report builds on developments reported in previous *GDSA Framework* milestones, particularly M3SF-20SN010304032.

## Contents

Acknowledgements.....	i
Executive Summary.....	ii
1. Introduction.....	1-1
1.1 Overview of this Report.....	1-2
1.2 GDSA Framework.....	1-4
1.2.1 PFLOTRAN.....	1-5
1.2.2 Dakota.....	1-5
1.2.3 Visualization Tools.....	1-7
1.2.4 Automated Analysis Workflow Development for GDSA.....	1-8
1.2.5 GDSA Software Strategy.....	1-8
1.3 References: Chapter 1.....	1-9
2. UQ Methodology: Multifidelity methods.....	2-1
2.1 Multifidelity Polynomial Chaos Expansion.....	2-1
2.2 Simplified crystalline problem.....	2-3
2.2.1 Problem Description.....	2-3
2.2.2 Assessment of Model Hierarchy.....	2-6
2.3 Results.....	2-9
2.3.1 No Spatial Uncertainty.....	2-10
2.3.2 With Spatial Uncertainty.....	2-12
2.4 Conclusions and Future Work.....	2-20
2.5 References: Chapter 2.....	2-20
3. Discrete fracture networks.....	3-1
3.1 Transmissivity Relationships.....	3-1
3.2 Graph Metrics characterizing the DFNs.....	3-3
3.3 Incorporating Graph Metrics into Polynomial Chaos Expansion.....	3-4
3.4 Analysis of Transmissivity Relationships.....	3-5
3.4.1 Approach.....	3-5
3.4.2 Results.....	3-7
3.4.3 Conclusions from correlated-constant vs. depth-dependent transmissivity study.....	3-14
References: Chapter 3.....	3-14
4. Crystalline reference case: Updates.....	4-1
4.1 Uncertainty Analysis (UA).....	4-1
4.1.1 Uncertainty Categorization.....	4-1
4.1.2 Uncertainty Implementation.....	4-3
4.1.3 Instant Release Fraction.....	4-5
4.2 Model Set-up.....	4-8
4.2.1 Model Domain.....	4-8
4.2.2 Discrete Fracture Networks.....	4-15
4.2.3 Waste Package Corrosion Model.....	4-16

4.2.4	Initial and Boundary Conditions .....	4-17
4.2.5	Observation Points .....	4-17
4.2.6	Timestep Size .....	4-19
4.2.7	Quantities of Interest (QoIs) .....	4-19
4.3	Results .....	4-22
4.3.1	Sensitivity Results for Scalar QoIs .....	4-22
4.3.2	Sensitivity Results for Time-Dependent QoIs .....	4-32
4.4	Summary .....	4-41
4.5	References: Chapter 4 .....	4-43
5.	Crystalline Reference Case UA Nested Workflow .....	5-1
5.1	Crystalline Reference Case UA Nested Workflow Development.....	5-2
5.1.1	Crystalline Reference Case UA Nested Workflow Structure Summary .....	5-2
5.1.2	Crystalline Reference Case UA Nested Workflow Updates .....	5-5
5.1.3	Updates to Improve Crystalline Reference Case UA Nested Workflow Reliability .....	5-7
5.2	Crystalline Reference Case UA Nested Workflow Application .....	5-7
5.3	Outcomes of Crystalline Reference Case UA Nested Workflow Use .....	5-9
5.4	References: Chapter 5 .....	5-10
6.	Conclusions .....	6-1

## Table of Figures

Figure 1-1. The GDSA Framework .....	1-5
Figure 1-2. Dakota interfacing to a computational model such as a repository simulator .....	1-7
Figure 2-1. A vertical slice of the simplified crystalline domain, taken at $y = 500$ m.....	2-4
Figure 2-2. 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. ....	2-5
Figure 2-3. Scatterplots for the median residence time for the repository for the 20 m vs 10 m mesh (left) and 40 m vs 20 m mesh (right) .....	2-7
Figure 2-4. Scatterplots of the $x$ location of the peak $^{129}\text{I}$ concentration as a function of $k_{\text{Glacial}}$ (left) and of the mean travel time at the location of the peak $^{129}\text{I}$ concentration as a function of $k_{\text{Glacial}}$ (right), for the 10 m, 20 m, and 40 m meshes. ....	2-8
Figure 2-5. Scatterplots for the median residence time for the repository for the 20 m vs 10 m mesh, with no spatial uncertainty (left) and including spatial uncertainty (right).....	2-9
Figure 2-6. Comparison of Sobol' indices from the reference PCE computed with 828 high-fidelity samples (blue) and for the MF PCE (orange). ....	2-11
Figure 2-7. Histograms of sample model (PFLOTRAN) evaluations compared to samples from a single-fidelity PCE constructed using all 828 samples from the highest-fidelity (10 m mesh) model and samples from a multifidelity PCE constructed using $\sim 21$ equivalent high-fidelity evaluations.....	2-12
Figure 2-8. A comparison of the reference max-3 <sup>rd</sup> order cross-validation PCE with a 3 <sup>rd</sup> order PCE without cross-validation, constructed using the same number of samples. In (a), sample profiles from each PCE is compared to those computed from PFLOTRAN (model). In (b), the coefficients of each PCE (note that the cross-validation PCE coefficients are truncated after 1 <sup>st</sup> order). In (c), the Sobol' indices for each PCE compared to those computed from a GP surrogate.....	2-14
Figure 2-9. Sobol' indices for peak $^{129}\text{I}$ concentration in the aquifer [M] and median residence time of the repository [yr], compared between the cross-validation PCE, the GP, and the quadratic polynomial surrogate used to compute them.....	2-15
Figure 2-10. A comparison of the main effects indices from three MF PCEs with different sample profiles compared to the reference PCE. ....	2-17
Figure 2-11. Main effects indices for two representative performance QoIs. Each subplot is a comparison of the Sobol' indices computed from the reference PCE (blue) with those computed from the PCE indicated in the title of the subplot (orange). ....	2-18
Figure 2-12. Main effects indices for performance QoIs where either the MF PCE or the 20 m PCE outperformed the other. Each subplot is a comparison of the Sobol'	

indices computed from the reference PCE (blue) with those computed from the PCE indicated in the title of the subplot (orange). .....	2-19
Figure 3-1. Normal density fits to the graph metrics for the crystalline reference case, along with histograms using a varying number of bins to show the goodness of fit over a range of bin sizes. ....	3-5
Figure 3-2. Interval plots for the geometric mean permeabilities of each depth zone. ....	3-8
Figure 3-3. Interval plots for the specified quantities of interest. ....	3-9
Figure 3-4. Boxplots for the geometric mean permeabilities of each depth zone. ....	3-10
Figure 3-5. Boxplots for the specified quantities of interest. ....	3-11
Figure 3-6. Scatterplot displaying the relationship between the total maximum <sup>129</sup> I concentration in the aquifer and the geometric mean permeability for depth zone one. ....	3-12
Figure 3-7. Scatterplots displaying the relationship between (A) the ratio of the mass flow from the aquifer to the east boundary to the mass flow rate from the rock to the east boundary (at 1 million years) and (B) the ratio of the mass flow rate from the rock to the aquifer to the mass flow rate from the rock to the east boundary (at 1 million years) with respect to the geometric mean permeability for depth zone one. ....	3-13
Figure 4-1. Sampling loops used in the Uncertainty Analysis of the Crystalline Reference Case presented in this Chapter. ....	4-4
Figure 4-2 Scatter plot of peak <sup>129</sup> I concentration (M) in the aquifer versus <i>IRF</i> . ....	4-7
Figure 4-3 Scatterplot of the peak <sup>129</sup> I concentration in the aquifer (M) versus <i>rateUNF</i> . ....	4-8
Figure 4-4 Sensitivity of peak I-129 concentration to various inputs. ....	4-8
Figure 4-5. Cut-away of DFN 1 realization mapped to porous medium grid, showing the full repository and the far half of the model domain. ....	4-9
Figure 4-6. FY20 model domain XY slice through the repository colored by material: dark blue, granite; light blue, granite at the repository level; red, buffer; maroon, WPs. Spacing between WPs is 10 m, 80 WPs/drift, 3360 total WPs. Inset zoom on the top left of the figure shows four WPs. ....	4-10
Figure 4-7. FY21 model domain XY slice through the repository colored by material: dark blue, granite; light blue, granite at the repository level; red, buffer; maroon, WPs. Spacing between WPs is 20 m, 40 WPs/drift, 1680 total WPs. Inset zoom on the top left of the figure shows four WPs. ....	4-11
Figure 4-8. Plot of heat generation per waste package over time for the previously used 12-PWR WPs in the FY20 case (dashed purple) and the recalculated 12-PWR WPs for the FY21 case (solid turquoise). ....	4-12
Figure 4-9. Plot comparing the temperature (°C) of the centermost WP of the repository between the FY20 (dashed purple) and FY21 case (solid turquoise). ....	4-13

Figure 4-10. FY21 PFLOTRAN integral flux planes and depth zones visualized within the model domain. .... 4-15

Figure 4-11. XY slice through the repository colored by material ID shows all twenty-seven newly added observation points represented by white dots ..... 4-18

Figure 4-12. Global sensitivity analysis results for peak  $^{129}\text{I}$ ..... 4-24

Figure 4-13. Scatterplots of peak  $^{129}\text{I}$  versus uncertain parameters and graph metrics. .... 4-24

Figure 4-14. Global sensitivity analysis results for the mean travel time at peak  $^{129}\text{I}$ ..... 4-27

Figure 4-15. Scatterplots of mean travel time at peak  $^{129}\text{I}$  versus uncertain parameters and graph metrics. .... 4-27

Figure 4-16. Global sensitivity analysis results for the median residence time of tracer in repository..... 4-28

Figure 4-17. Scatterplots of the median residence time of tracer in repository versus uncertain parameters and graph metrics..... 4-28

Figure 4-18. Global sensitivity analysis results for the fractional mass flux from the repository at 3000 years ..... 4-29

Figure 4-19. Scatterplots of the fractional mass flux from the repository at 3000 years versus uncertain parameters and graph metrics..... 4-29

Figure 4-20. Global sensitivity analysis results for the ratio between the aquifer to east boundary flux and the rock to east boundary flux at 1 Ma. .... 4-30

Figure 4-21. Scatterplots of the ratio between the aquifer to east boundary flux and the rock to east boundary flux at 1 Ma versus uncertain parameters and graph metrics. .... 4-30

Figure 4-22. Global sensitivity analysis results for the ratio between the rock to aquifer flux and the rock to east boundary flux at 1 Ma. .... 4-31

Figure 4-23. Scatterplots of the ratio between the rock to aquifer flux and the rock to east boundary flux at 1 Ma versus uncertain parameters and graph metrics..... 4-31

Figure 4-24. Global sensitivity analysis results over time for the maximum  $^{129}\text{I}$  concentration in the aquifer ..... 4-33

Figure 4-25. Scatterplots of the maximum  $^{129}\text{I}$  concentration versus *meanWPrate* and the number of intersections with the repository travel time at four time points. .... 4-34

Figure 4-26. Global sensitivity analysis results over time for the log (base 10) transformed maximum  $^{129}\text{I}$  concentration in the aquifer. .... 4-36

Figure 4-27. Global sensitivity analysis results over time for the fraction of tracer remaining in the repository. .... 4-37

Figure 4-28. Global sensitivity analysis results over time for the water flux from the rock to the aquifer..... 4-38

Figure 4-29. Global sensitivity analysis results over time for the rock to east boundary flux normalized by the flux at 1 Ma. .... 4-39

Figure 4-30. Global sensitivity analysis results over time for the aquifer to east boundary flux normalized by the flux at 1 Ma. .... 4-40

Figure 4-31. Global sensitivity analysis results over time for the ratio between the rock to aquifer flux and the aquifer to east boundary flux. .... 4-41

Figure 5-1. High-level view of the Crystalline Reference Case UA Nested Workflow structure. .... 5-3

Figure 5-2. Detailed view of the Crystalline Reference Case UA Nested Workflow structure showing the four workflows along with important files, actions, and codes that are employed for each. .... 5-5

Figure 5-3. Addition of graph metrics to the Aleatory/Spatial Sampling Workflow, outlined in red. .... 5-6

Figure 5-4. Screenshot of aleatory input parameters ..... 5-9

Figure 5-5. Screenshot of HPC user inputs..... 5-9

## Table of Tables

Table 2-1. Uncertain parameters in PFLOTRAN problem specification. ....	2-5
Table 2-2. Absolute and relative costs of models in hierarchy.....	2-6
Table 2-3. The number of terms and samples associated with total-order PCEs over varying order with 9 input parameters.....	2-16
Table 3-1. Crystalline reference case DFN parameterization.....	3-2
Table 3-2. Transmissivity relationship parameterizations considered in this study .....	3-3
Table 3-3. Graph Metrics used in crystalline reference case GSA. ....	3-3
Table 3-4. Depth zone definitions.....	3-6
Table 4-1. Uncertainty distributions propagated in crystalline reference case UA. ....	4-5
Table 4-2. I-129 IRF data used to bound IRF uncertainty for spent fuel in crystalline reference case .....	4-6
Table 4-3. Implemented values of soil compressibility for all material properties. ....	4-14
Table 4-4. FY2021 QoIs for crystalline reference case sensitivity analysis.....	4-21

## ACRONYMS AND DEFINITIONS

Abbreviation	Definition
CDF	cumulative distribution function
CRC	crystalline reference case
DOE	U.S. Department of Energy
DFN	discrete fracture network
DRZ	disturbed rock zone
ECPM	equivalent continuous porous medium
GDSA	Geologic Disposal Safety Assessment
Ma	mega annum (one million years)
MF	multifidelity
NGW	Next Generation Workflow
PA	Performance Assessment
PCE	Polynomial Chaos Expansion
SA	sensitivity analysis
SNL	Sandia National Laboratories
UQ	uncertainty quantification
WP	waste package

## 1. INTRODUCTION

This report presents high level objectives and strategy for development of uncertainty and sensitivity analysis tools in *Geologic Disposal Safety Assessment (GDSA) Framework*, a software toolkit for probabilistic post-closure performance assessment (PA) of systems for deep geologic disposal of nuclear waste. *GDSA Framework* is supported by the Spent Fuel and Waste Science and Technology (SFWST) Campaign of the U.S. Department of Energy (DOE) Office of Nuclear Energy (NE) and its predecessor the Used Fuel Disposition (UFD) campaign.

This report fulfills the GDSA Uncertainty and Sensitivity Analysis Methods work package (SF-21SN01030404) level 3 milestone – *Uncertainty and Sensitivity Analysis Methods and Applications in GDSA Framework (FY2021)* (M3SF-21SN010304042). It presents high level objectives and strategy for development of uncertainty and sensitivity analysis tools, demonstrates uncertainty quantification (UQ) and sensitivity analysis (SA) tools in *GDSA Framework* in FY21, and describes additional UQ/SA tools whose future implementation would enhance the UQ/SA capability of *GDSA Framework*.

This work was closely coordinated with the other Sandia National Laboratory GDSA work packages: the GDSA Framework Development work package (SF-21SN01030405), the GDSA Repository Systems Analysis work package (SF-21SN01030406), and the GDSA PFLOTRAN Development work package (SF-21SN01030407). This report builds on developments reported in previous *GDSA Framework* milestones, particularly M3SF-20SN010304032 [1] and M3SF-19SN010304032 [2].

The repository community must maintain leadership in UQ/SA methods. Computational codes are becoming increasingly complex and require high performance computers to run, resulting in costly sample evaluations. Geologic repository performance assessment in the U.S. involves a code base that includes coupled, multiphysics modeling at high resolution. Having relatively few simulation samples highlights the need to consider surrogate models to sample and explore the input parameter space more extensively. However, this must be done in a careful way so that surrogate accuracy can be tracked and understood in the context of UQ/SA results. Variance-based sensitivity indices are now a standard practice in the sensitivity analysis community but require many evaluations of the predictive model. Much research has focused on calculating variance-based sensitivity indices while keeping the computational cost reasonable. We note that other sensitivity analysis methods [2] may be better than variance-based methods at identifying patterns of behavior or trends. Another recent approach is to employ “multifidelity” UQ in which many low-fidelity simulation runs (e.g., coarser mesh, simpler physics) augment a small number of high-fidelity runs [1]. Keeping abreast of improvements to existing UQ/SA methods as well as employing new methods is critical to performing sensitivity and uncertainty analysis of new repository systems which will involve large parameter spaces and computationally expensive simulations.

## 1.1 Overview of this Report

This report provides documentation of the UQ/SA work performed in FY 2021. The outline of this report is as follows:

- Chapter 2 provides an overview of an exciting new uncertainty quantification method involving models at multiple levels of fidelity, hence the name “Multifidelity Uncertainty Quantification.” The main idea in multifidelity UQ is to extract information from a limited number of high-fidelity model evaluations and complement them with a much larger number of a set of lower fidelity evaluations. The final result is an estimator of the mean or statistic of the response (e.g. percentile) with a lower variance: a more accurate and reliable estimator can be obtained at lower computational cost. In 2020, we demonstrated the use of multifidelity UQ on a simplified 1-D test problem. This year, we provide a more realistic example using a modified crystalline reference case. We emphasize how multifidelity approaches can be used to efficiently estimate sensitivity indices.
- Chapter 3 discusses the discrete fracture networks (DFN) that were generated this year and the additional analyses performed on the DFNs using graph metrics. We also utilized the depth-dependent transmissivity in *dfnWorks*. These results are documented in Chapter 3.
- In Chapter 4, we extend the sensitivity analysis of the crystalline reference case performed in FY20. In FY20, we focused on extracting indicators of fracture connectivity as well as information about the flow between the repository and the aquifer. To facilitate this, PFLOTRAN added several new quantities of interest, including peak concentrations (of  $^{129}\text{I}$  and tracers) in the aquifer at each time step, along with the location of the peaks, mean travel time from the repository to various locations, median residence time of a tracer within the repository, and total water fluxes in various directions. These quantities of interest (QoIs) were used in the FY20 sensitivity analysis.

This year, we modified the repository in the crystalline reference case to make it cooler. That is, we used half the waste packages as in FY20 but the same repository footprint to obtain a maximum waste package temperature of 120-130 °C. We generated new DFNs and are extracting additional quantities of interest from the results, including fraction of tracers remaining and fractional fluxes of tracers at certain time points. We added new features to the PFLOTRAN model, including soil compressibility. We also treated the instant release fraction as an uncertain parameter. Finally, in the nested sampling approach, we generated different realizations for each epistemic inner loop. These differences and the new sensitivity results for the 2021 crystalline reference case are outlined in Chapter 4.

- Chapter 5 documents the *GDSA Workflow* for a UQ study. Uncertainty quantification workflows are not trivial to define and get running, even when using tools such as Dakota to generate nested studies involving sampling loops over both aleatory and epistemic samples. The analyst has to spend a significant amount of time writing scripts to interface the sample values to PFLOTRAN, extract the results, and put the entire workflow on a high-performance computing platform. Further, as the quantities of interest increase and many vectors of results are generated for each simulation, plotting and aggregating the results in a variety of ways (e.g. averaging over epistemic or aleatory slices as a function of time) becomes very involved. To address this, the *GDSA Workflow* was developed in FY20. This workflow couples Dakota, PFLOTRAN, and NGW (the Next Generation Workflow) to present the user with a unified

interface where the actual workflow can be dictated in an easy-to-use graphical format. This workflow also allows greater reproducibility and traceability of the actual files and scripts used for a particular study. In 2021, we significantly enhanced the *GDSA Workflow* and made it production ready, including the capability to launch PFLOTRAN jobs on high performance computing systems where the workflow is run on a local machine. We also added significant postprocessing capabilities to help automate sensitivity analysis. We used the GDSA Workflow to generate the 1000 PFLOTRAN runs (25 DFNs x 40 epistemic samples) for the crystalline reference case sensitivity analysis documented in Chapter 4.

- Chapter 6 provides a summary.

## 1.2 GDSA Framework

*GDSA Framework* (Figure 1-1) capabilities include multi-physics simulation of coupled processes affecting deep geologic repository performance, uncertainty and sensitivity analysis, pre- and post-processing, and visualization. For a given performance assessment, these tools will be linked to a version-controlled parameter database and an automated run-control system. The overall objectives of *GDSA Framework* development are to:

- create a framework that is flexible enough to take advantage of future advances in hardware, software, and simulation and analysis methods;
- leverage existing high-performance computing capabilities (e.g., meshing, simulation, analysis, and visualization);
- enable increasingly coupled, mechanistic multi-physics modeling;
- provide analysis methods for prioritization of SFWST Disposal Research R&D activities;
- provide transparent implementation of simulation and analysis methods;
- develop and distribute in an open-source environment so that software is freely available to stakeholders ([1-6]).

Objectives specific to the uncertainty and sensitivity analysis capability in *GDSA Framework* are to make available standard sampling-based methods of uncertainty propagation, sensitivity analysis, and uncertainty quantification typically used within U.S. nuclear waste disposal programs (e.g., DOE 2008 [7], DOE 2014 [8], RESS2000[14], RESS2014[15]); and to enable future adoption of new methods consistent with the current standard of practice in the UQ/SA community and appropriate for high-dimensional, highly coupled, nonlinear problems resulting from the implementation of mechanistic multi-physics simulations. Having a consistent, common framework which enables a user to perform a range of sensitivity analysis and UQ approaches for a particular problem or set of simulations allows for reproducibility, comparative analyses, use of verified algorithms, and documentation of best practices. These are important goals for performance assessments now and in the future.

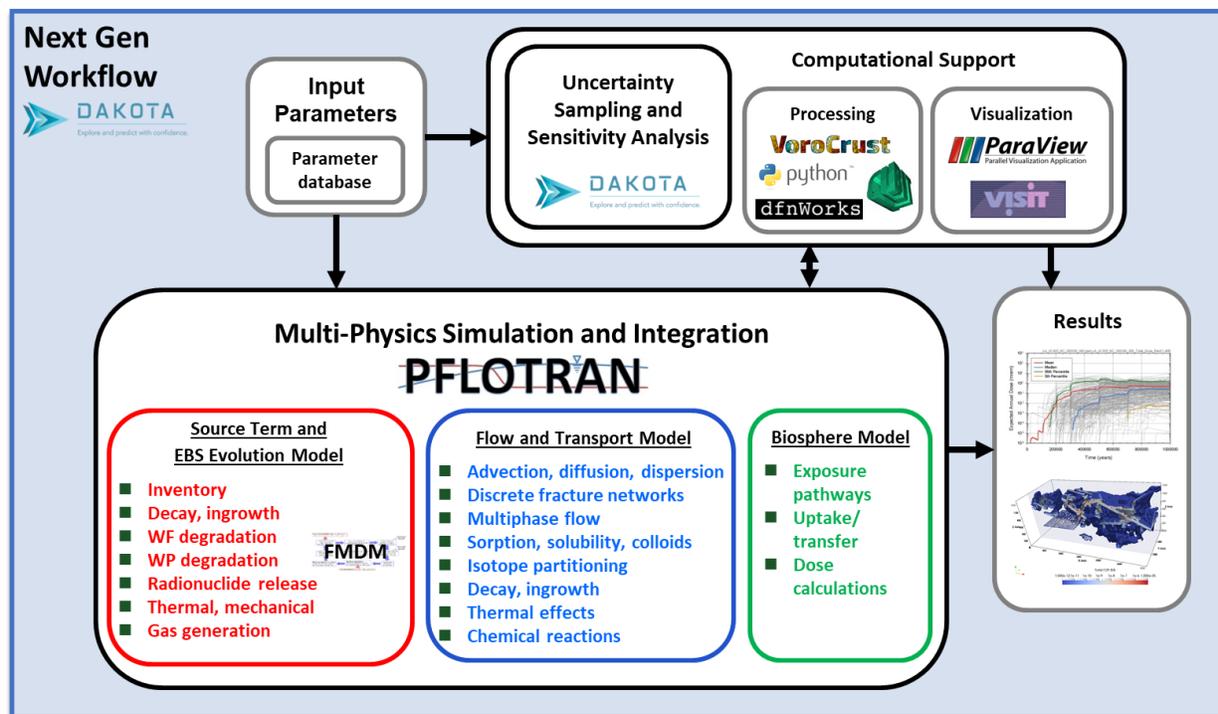


Figure 1-1. The GDSA Framework

The following sections highlight the key components of *GDSA Framework*. More information about each can be found by following the links at <https://pa.sandia.gov>.

### 1.2.1 PFLOTRAN

PFLOTRAN is an open source, state-of-the-art, massively parallel subsurface flow and reactive transport simulator ([9-11]). Written in object-oriented Fortran 2003, PFLOTRAN is a porous medium continuum code for modeling multicomponent, multiphase flow and transport, heat conduction and convection, biogeochemical reactions, geomechanics, and isotope decay and ingrowth. The code is developed under a GNU LGPL license allowing for third parties to interface proprietary software with the code. The availability and continuing development of PFLOTRAN are due to an ongoing collaborative effort of several DOE laboratories led by Sandia. PFLOTRAN development for *GDSA Framework* is described by Mariner et al. ([3-5]) and Sevougian et al. 2018 [12]. PFLOTRAN installation instructions are available at <https://www.pflotran.org/>.

### 1.2.2 Dakota

Dakota is an open-source toolkit of algorithms that contains both state-of-the-art research and robust, usable software for optimization and uncertainty quantification (UQ). It is available at: <https://dakota.sandia.gov> [13]. The algorithms allow a user to explore a computational simulation to answer questions such as:

- what is the best design?
- how safe is this design?
- what are the most important parameters?
- what effects do uncertainties have on my system?

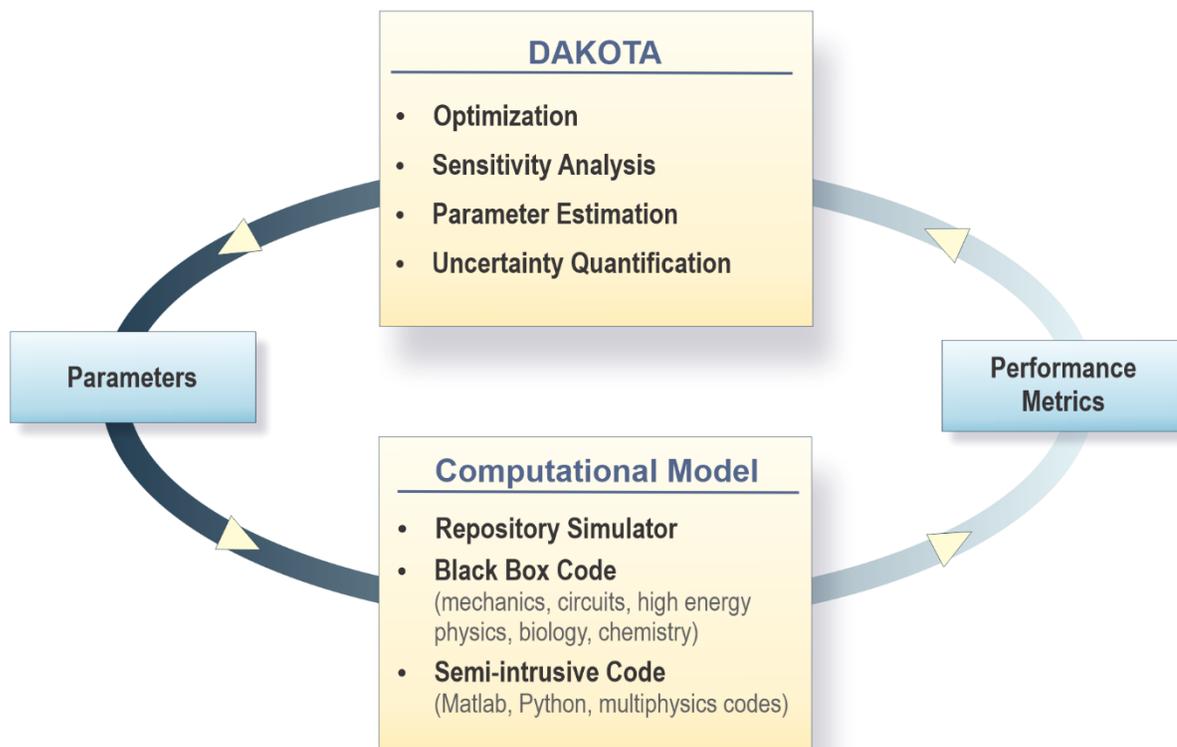
The Dakota software has a rich set of parametric analysis methods that enable design exploration, model calibration, risk analysis, and quantification of margins and uncertainty with computational models. Dakota provides a flexible, extensible interface between simulation codes and these iterative analysis methods, which include:

- optimization with gradient and nongradient-based methods;
- uncertainty quantification with sampling, reliability, stochastic expansion, and epistemic methods;
- parameter estimation using nonlinear least squares (deterministic) or Bayesian inference (stochastic); and
- sensitivity/variance analysis with design of experiments and parameter study methods.

These capabilities may be used on their own or as components within advanced strategies such as hybrid optimization, surrogate-based optimization, mixed-integer nonlinear programming, or optimization under uncertainty. Dakota is a C++ code and has been under development at Sandia since 1994 and has been primarily sponsored by DOE's Advanced Simulation and Computing (ASC) program. As such, it has a focus on interfacing to and running simulations which are computationally expensive, require high performance computing and parallel execution, and exhibit nonlinearities, non-monotonic and/or discontinuous responses, and often involve noisy responses and high-dimensional inputs. Thus, a focus of the algorithm development in Dakota has been on methods that are as efficient as possible and minimize the number of runs required of a high-fidelity simulation model. Such algorithms include surrogate or emulator models, adaptive sampling approaches, and multifidelity UQ methods which augment a small number of high-fidelity runs with many low-fidelity runs to obtain comparable accuracy in statistical estimators.

Dakota contains the uncertainty quantification and sensitivity analysis methods typically used in the U.S. repository program. Dakota implements Latin Hypercube Sampling (LHS) with correlation control on input parameters. It calculates moments on responses of interest as well as correlation matrices (simple, partial, and rank correlations) between inputs and outputs. Dakota also contains an algorithm for performing incremental LHS which allows one to double an initial LHS study such that the second LHS study is a Latin design and the combined initial and second LHS studies together form a Latin hypercube design. Dakota allows nested studies to perform an "outer loop" epistemic sampling and an "inner loop" aleatory sampling to generate ensembles of distributions. Dakota returns tables of input and output amenable to further processing and visualization with additional tools developed within *GDSA Framework* or by an individual user. Additional methods that have been implemented in Python for use in *GDSA Framework* include calculation of standardized regression coefficients via stepwise linear regression and calculation of partial correlation coefficients based on raw data or rank-transformed data.

A graphical depiction of Dakota interfacing with a computational model such as a repository simulation in PFLOTRAN is shown in Figure 1-2. Based on the type of study being performed (optimization, uncertainty quantification, etc.), Dakota chooses the next set of parameters at which to evaluate the simulator and runs the simulator, which returns the performance metrics of interest back to Dakota. Dakota then generates the next set of parameters according to the algorithm being used for the study and keeps iterating until the specified number of samples is reached.



**Figure 1-2. Dakota interfacing to a computational model such as a repository simulator**

The UQ/SA methods in Dakota have evolved as the standard of practice evolves. Over the past ten years, the Dakota team has invested in methods which calculate the Sobol’ variance-based sensitivity indices in an efficient manner. These indices estimate the proportion of variance in a quantity of interest that can be attributed to variance in each uncertain input parameter. Currently, a Dakota user can calculate these by extensive sampling of the simulation code, by using surrogate methods such as regression or Gaussian process models, and by the use of polynomial chaos expansions. These advanced methods are presented in more detail later as used in *GDSA Framework*. Dakota is an actively maintained and developed code with formal releases issued twice per year. Dakota uses formal software quality development processes including advanced version control, unit and regression testing, agile programming practices, and software quality assessment.

### 1.2.3 Visualization Tools

*GDSA Framework* employs ParaView and/or VisIT for visualization of results. ParaView is an open-source, multi-platform data analysis and visualization application developed by Sandia National Laboratories. As stated on the ParaView website (<https://www.paraview.org/>): “ParaView users can quickly build visualizations to analyze their data using qualitative and quantitative techniques. The data exploration can be done interactively in 3D or programmatically using ParaView’s batch processing capabilities. ParaView was developed to analyze extremely large datasets using distributed memory computing resources. It can be run on supercomputers to

analyze datasets of petascale size as well as on laptops for smaller data, has become an integral tool in many national laboratories, universities and industry, and has won several awards related to high performance computation.”

VisIT has been developed at Livermore National Laboratory. As stated on the VisIT website (<https://wci.llnl.gov/simulation/computer-codes/visit>): “VisIT is an open source, interactive, scalable, visualization, animation, and analysis tool. From Unix, Windows, or Mac workstations, users can interactively visualize and analyze data ranging in scale from small ( $<10^1$  cores) desktop-sized projects to large ( $>10^5$  cores) leadership-class computing facility simulation campaigns. Users can quickly generate visualizations, animate them through time, manipulate them with a variety of operators and mathematical expressions, and save the resulting images and animations for presentations. VisIt contains a rich set of visualization features to enable users to view a wide variety of data including scalar and vector fields defined on two- and three-dimensional (2D and 3D) structured, adaptive and unstructured meshes. Owing to its customizable plugin design, VisIt is capable of visualizing data from over 120 different scientific data formats.”

### **1.2.4 Automated Analysis Workflow Development for GDSA**

Uncertainty Quantification analysis workflows are not trivial to define and get running, even when using tools such as Dakota to generate nested studies involving sampling loops over both aleatory and epistemic samples. The analyst has to spend a significant amount of time writing scripts to interface the sample values to PFLOTRAN, extract the results, and put the entire workflow on a high-performance computing platform. Further, as the number of quantities of interest increases and many vectors of results are generated for each simulation, plotting and aggregating the results in a variety of ways (e.g., averaging over epistemic or aleatory slices as a function of time) becomes very involved. To address this, the *Crystalline Reference Case UA Nested Workflow* was developed. This workflow couples Dakota, PFLOTRAN, and NGW (the Next-Generation Workflow software) to present the user with a unified GUI where the actual workflow can be dictated and automated in an easy-to-use graphical format. This workflow also allows greater reproducibility and traceability of the actual files and scripts used for a particular study. This year, we demonstrated the Crystalline Reference Case UA Nested Workflow for a GDSA study involving both aleatory and epistemic sampling. This is further discussed in Chapter 5.

### **1.2.5 GDSA Software Strategy**

The software strategy for GDSA is to leverage and use open source software that is actively maintained and developed, whenever possible. That is why the *GDSA Framework* utilizes PFLOTRAN, Dakota, Paraview, and VisIT. Another goal is to support HPC computing, which is a primary focus for all of the software tools listed above. In addition, the *GDSA Framework* should have the flexibility to develop and adopt new capabilities as state-of-the-art hardware, software, and methodology evolves. Again, the codes chosen for GDSA exhibit this flexibility and are constantly evolving and adopting to utilize new software and hardware capabilities. Much of the interfacing between the codes is currently performed with Python scripts and other scripting tools, and there will be tighter integration as *GDSA Framework* and the GDSA workflow progresses.

### 1.3 References: Chapter 1

1. Swiler, L.P., E. Basurto, D.M. Brooks, A. C. Eckert, P. E. Mariner, T. Portone and E.R. Stein. "Advances in Uncertainty Quantification and Sensitivity Analysis Methods and Applications in GDSA Framework." SAND2020-10802 R. Sandia National Laboratories, Albuquerque, NM.
2. Swiler, L.P., J.C. Helton, E. Basurto, D.M. Brooks, P. Mariner, L.M. Moore, S. Mohanty, S.D. Sevougian, and E.R. Stein. "Status Report on Uncertainty Quantification and Sensitivity Analysis Tools in the Geologic Disposal Safety Assessment (GDSA) Framework." SAND2019-13835 R. Sandia National Laboratories, Albuquerque, NM.
3. Mariner, P. E., E. R. Stein, S. D. Sevougian, L. J. Cunningham, J. M. Frederick, G. E. Hammond, T. S. Lowry, S. Jordan, and E. Basurto (2018). Advances in Geologic Disposal Safety Assessment and an Unsaturated Alluvium Reference Case. SFWD-SFWST-2018-000509; SAND2018-11858R. Sandia National Laboratories, Albuquerque, NM.
4. Mariner, P. E., E. R. Stein, J. M. Frederick, S. D. Sevougian, and G. E. Hammond (2017). Advances in Geologic Disposal System Modeling and Shale Reference Cases. SFWD-SFWST-2017-000044 / SAND2017-10304R. Sandia National Laboratories, Albuquerque, NM.
5. Mariner, P. E., E. R. Stein, J. M. Frederick, S. D. Sevougian, G. E. Hammond, and D. G. Fascitelli (2016). Advances in Geologic Disposal System Modeling and Application to Crystalline Rock. FCRD-UFD-2016-000440 / SAND2016-9610R. Sandia National Laboratories, Albuquerque, NM.
6. Mariner, P. E., W. P. Gardner, G. E. Hammond, S. D. Sevougian, and E. R. Stein (2015). Application of Generic Disposal System Models. FCRD-UFD-2015-000126 / SAND2015-10037R. Sandia National Laboratories, Albuquerque, NM.
7. DOE (2008). Yucca Mountain Repository License Application Safety Analysis Report. DOE/RW-0573, Revision 1, <http://www.nrc.gov/waste/hlw-disposal/yucca-lic-app/yucca-lic-app-safety-report.html#1>. US Department of Energy, Washington, DC.
8. DOE (2014). Title 40 CFR Part 191 Subparts B and C Compliance Recertification Application 2014 for the Waste Isolation Pilot Plant. DOE/WIPP 15-3503. U.S. Department of Energy, Carlsbad, New Mexico.
9. Lichtner, P. C. and G. E. Hammond (2012). Quick Reference Guide: PFLOTRAN 2.0 (LA-CC-09-047) Multiphase-Multicomponent-Multiscale Massively Parallel Reactive Transport Code. LA-UR-06-7048. December 8, 2012. Los Alamos National Laboratory, Los Alamos, New Mexico.
10. Hammond, G. E., P. C. Lichtner, R. T. Mills, and C. Lu. (2008). Toward petascale computing in geosciences: application to the Hanford 300 Area. In R. L. Stevens (Ed.), Scidac 2008: Scientific Discovery through Advanced Computing (Vol. 125).
11. Hammond, G. E., P. C. Lichtner, and R. T. Mills (2014). "Evaluating the performance of parallel subsurface simulators: An illustrative example with PFLOTRAN". Water Resources Research, 50(1), 208-228. doi: 10.1002/2012wr013483.

12. Sevougian, S. D., G. E. Hammond, P. E. Mariner, E. R. Stein, J. M. Frederick, and R. J. MacKinnon (2018). "GDSA Framework: High-Performance Safety Assessment Software to Support the Safety Case." in Proceedings of the IGSC Safety Case Symposium 2018, Rotterdam, The Netherlands, October 10-11, 2018, SAND2018-9975C, Sandia National Laboratories. Albuquerque, NM.
13. Adams, B.M., Bohnhoff, W.J., Dalbey, K.R., Ebeida, M.S., Eddy, J.P., Eldred, M.S., Hooper, R.W., Hough, P.D., Hu, K.T., Jakeman, J.D., Khalil, M., Maupin, K.A., Monschke, J.A., Ridgway, E.M., Rushdi, A.A., Seidl, D.T., Stephens, J.A., Swiler, L.P., and Winokur, J.G., "Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.12 User's Manual," Sandia Technical Report SAND2020-5001, May 2020.
14. Helton J.C. and M.G. Marietta (Editors). 2000. Special Issue: The 1996 Performance Assessment for the Waste Isolation Pilot Plant. *Reliability Engineering and System Safety* 69:1-451.
15. Helton J.C., C.W. Hansen and P.N. Swift (Editors). 2014. Special Issue: Performance Assessment for the Proposed High-Level Radioactive Waste Repository at Yucca Mountain, Nevada. *Reliability Engineering and System Safety* 122:1-456.

## 2. UQ METHODOLOGY: MULTIFIDELITY METHODS

Repository performance assessment requires the repeated evaluation of complex multiscale, multiphysics models that simulate million-year time horizons. The high computational cost of these models limits the number of simulations that can be performed in an uncertainty analysis, which in turn limits the accuracy of summary statistics and statistical analyses using such few samples. Multifidelity uncertainty quantification methods mitigate this issue by exploiting an ensemble of models with varying fidelities and cost to achieve greater statistical accuracy at less computational cost. These methods rely on similarity in model responses to variations in uncertain model inputs, measured by correlations between the models, a rapid decay in the spectral content of differences between models, etc. Multilevel methods rely on the use of a hierarchy of models whose responses converge toward the highest-fidelity model's predictions the higher they are in the hierarchy. This enables the highest-fidelity model's response to be represented as a series of incrementally smaller corrections to the lower-fidelity model responses. Multifidelity methods encompass a broader class of approaches which relax this assumption of a convergent hierarchy of models.

This chapter details a study of the feasibility of deploying these methods to improve computational efficiency of GDSA UQ analyses. In [1], multilevel forward uncertainty propagation and multifidelity polynomial chaos expansion (MF PCE) for global sensitivity analysis (GSA) were successfully applied to a 1D problem simulating a tank leak at the Hanford nuclear waste site, achieving a marked improvement in computational efficiency. This year, a simplified version of the crystalline reference case was developed and MF PCE was applied to study the potential benefits of these methods for GSA on a more practical problem that mimics the characteristics of GDSA's UQ analyses using the full-scale crystalline reference case.

The remainder of this chapter proceeds as follows. In Section 2.1, multifidelity polynomial chaos expansions are briefly introduced. In Section 2.2 the simplified crystalline problem is defined, and in Section 2.3 a preliminary study using MF PCE is detailed.

### 2.1 Multifidelity Polynomial Chaos Expansion

Polynomial chaos expansions (PCEs) have gained popularity in uncertainty analyses in recent years and have been documented in detail in [1]. PCE is a stochastic expansion method whereby the output response is expanded in a series of polynomials that are orthogonal with respect to the densities of the random model inputs. One advantage of stochastic expansion methods such as PCE is that the response moments (means, variances, etc.), as well as Sobol' indices for global sensitivity analysis (GSA), can be computed analytically in terms of the coefficients of the expansion [2]. The ability to compute Sobol' indices from the PCE makes it an advantageous approach compared to other sampling-based methods for approximating the Sobol' indices (e.g. [3]), which typically require thousands of model evaluations for accurate estimates. For the PCE, the cost to compute Sobol' indices is negligible with respect to the cost of estimating its coefficients.

PCE coefficients are determined either by projecting model responses onto the basis or by solving a regression problem. The number of evaluations required to estimate the coefficients increases with increased number of input parameters and the order of the PCE, with the required

order of the PCE increasing with complexity of model responses. For instance, the number of model evaluations (samples) required to estimate the coefficients of a total-order PCE of power  $p$ , with number of input parameters  $n$ , using regression and a collocation ratio of  $r$  is computed as:

$$N_s = r \frac{(n+p)!}{n!p!}. \quad (2.1)$$

Further information about the many methods for constructing a PCE in Dakota and their associated costs is reported in [7]; coefficients were estimated using regression for this work. For complex model problems with many inputs and complex outputs requiring high-order PCEs, the cost of constructing an adequate surrogate model can quickly become intractable.

This challenge has motivated the development of multifidelity surrogate methods which exploit a hierarchy of model fidelities and costs, e.g. the same numerical model with a range of discretization refinements, or models with varying fidelities of physics models, to reduce the computational burden of constructing a surrogate for the highest-fidelity model's response. One example of a multifidelity surrogate is the multifidelity PCE (MF PCE) first introduced in a paper by Ng and Eldred in 2012 [4]. A benefit of the MF PCE is that once it is constructed it can be used in the same way as a single-fidelity PCE to analytically compute higher-order moments from its coefficients, as well as Sobol' indices. This makes it a prime candidate as an approach to multifidelity GSA.

Consider a two-model hierarchy, with the higher-fidelity model response denoted  $R_h$  and lower-fidelity model response denoted  $R_l$ , both of which take as inputs uncertain parameters  $\xi$ . The goal is to derive a PCE for the response of the high-fidelity model. In a single-fidelity framework, the high-fidelity model is evaluated multiple times to estimate the coefficients of the PCE  $\bar{R}_h$ :

$$R_h(\xi) \approx \bar{R}_h(\xi) = \beta_0 + \sum_{i=1}^p \beta_i \Psi_i(\xi), \quad (2.2)$$

where  $p$  is the order of the expansion; as  $p$  increases, more terms are included in the expansion, requiring more model evaluations to estimate the coefficients  $\beta_i$ .

A MF PCE can also be derived for  $R_h$  which uses evaluations from both models in the estimate. The concept is based on the fact that the high-fidelity model can be represented as the low-fidelity model corrected with a discrepancy term depending on the difference between the two models, which is hopefully easier to approximate than the high-fidelity model directly:

$$R_h = R_h + (R_l - R_l) = R_l + (R_h - R_l). \quad (2.3)$$

The MF PCE approximates  $R_h$  by constructing a PCE for  $R_l$  ( $\bar{R}_l$ ) and a PCE for  $R_h - R_l$  ( $\bar{R}_{h-l}$ ):

$$R_h(\xi) \approx \bar{R}_l(\xi) + \bar{R}_{h-l}(\xi) = [\beta_{0,l} + \sum_{i=1}^{p_l} \beta_{i,l} \Psi_i(\xi)] + [\beta_{0,h-l} + \sum_{i=1}^{p_{h-l}} \beta_{i,h-l} \Psi_i(\xi)]. \quad (2.4)$$

MF PCE is computationally more efficient if  $R_l$  is similar enough to  $R_h$  that the difference between the two responses is simpler than  $R_h$  alone. In cases where this is true, fewer terms are required to resolve the  $\bar{R}_{h-l}$  than to resolve  $\bar{R}_h$ , thereby requiring fewer evaluations of the more costly high-

fidelity model. This reduction in the number of terms may manifest as a more compact spectral content or in terms of simplified structure (sparsity/low rank).

Once the MF PCE is resolved to a desired level of accuracy, the coefficients of  $\bar{R}_l$  and  $\bar{R}_{h-l}$  can be aggregated by combining like terms, resulting in a single PCE representing  $R_h$ . For illustrative purposes assume  $p_l > p_{h-1}$ ; this combined PCE can thus be written

$$R_h(\xi) \approx (\beta_{0,l} + \beta_{0,h-1}) + \sum_{i=1}^{p_{h-l}} (\beta_{i,l} + \beta_{i,h-1}) \Psi_i(\xi) + \sum_{i=p_{h-1}+1}^{p_l} \beta_{i,l} \Psi_i(\xi). \quad (2.5)$$

This combined PCE can be used in any of the ways a PCE constructed in a single-fidelity fashion would be used, as a surrogate or to compute response moments or Sobol' indices. Generalizing to more than two models, a MF PCE can be constructed for a hierarchy of models, indexed from  $\ell = 0$  at the lowest fidelity to  $\ell = L$  at the highest fidelity, as

$$R_L \approx \bar{R}_0 + \sum_{\ell=1}^L \bar{R}_{\ell-(\ell-1)}, \quad (2.6)$$

where  $\bar{R}_{\ell-(\ell-1)}$  is the PCE expansion of  $R_\ell - R_{\ell-1}$ . The computational cost of constructing a MF PCE and the cost for a single-fidelity PCE can be compared by computing the number of equivalent high-fidelity model evaluations, denoted  $N_{equiv}$ . This is computed by defining the relative cost of the models in the hierarchy and multiplying the relative costs ( $C_\ell/C_L$ ) by the number of evaluations of each model ( $N_\ell$ ):

$$N_{equiv} = \sum_{\ell=0}^L \frac{C_\ell}{C_L} N_\ell. \quad (2.7)$$

This will be used to compare the performance of different PCE constructions in results. The remaining focus of this chapter is on the feasibility of using MF PCE to improve the computational efficiency of GSA for the crystalline reference case by applying the method to a simplified version of the problem which replicates the properties of the production problem as closely as possible.

## 2.2 Simplified crystalline problem

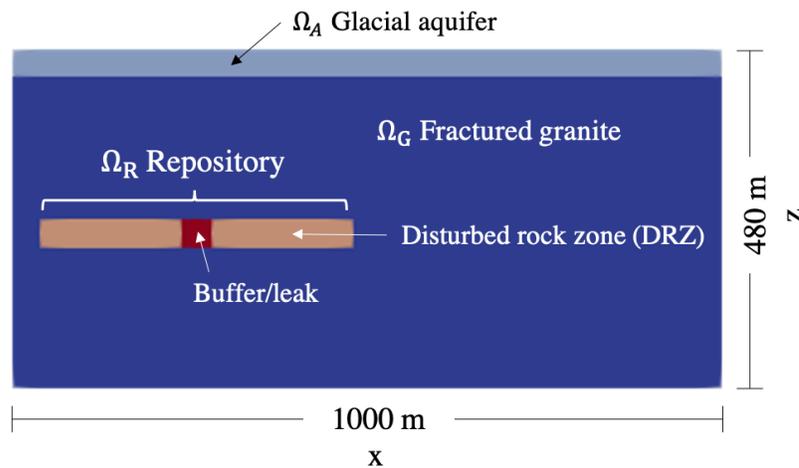
The eventual goal of this work is to deploy MF PCE to the crystalline reference case (CRC) used in production UQ/SA analyses for GDSA. To achieve this, a simplified crystalline problem was defined which is faster to evaluate but mimics the production problem as closely as possible. This will allow a robust framework for deploying MF PCE to the CRC to be created without the computational overhead of running the full model during the development phase. By defining simplified problem to mimic the properties of the CRC as closely as possible, potential challenges intrinsic to the problem can be identified and mitigated in advance. In 2.2.1 a description of the problem and the model hierarchy is provided, and in 2.2.2 a preliminary study of the model hierarchy to assess suitability for MF PCE is described.

### 2.2.1 Problem Description

The simplified crystalline problem was defined on a 1000 x 1000 x 480 m<sup>3</sup> domain and was implemented using PFLOTRAN. As with the CRC, several discrete fracture network (DFN) realizations were generated to represent spatial uncertainty. The parameterization for the DFNs was taken to be the [-200 m, -400 m] depth zone parametrization of the CRC, as defined in [1],

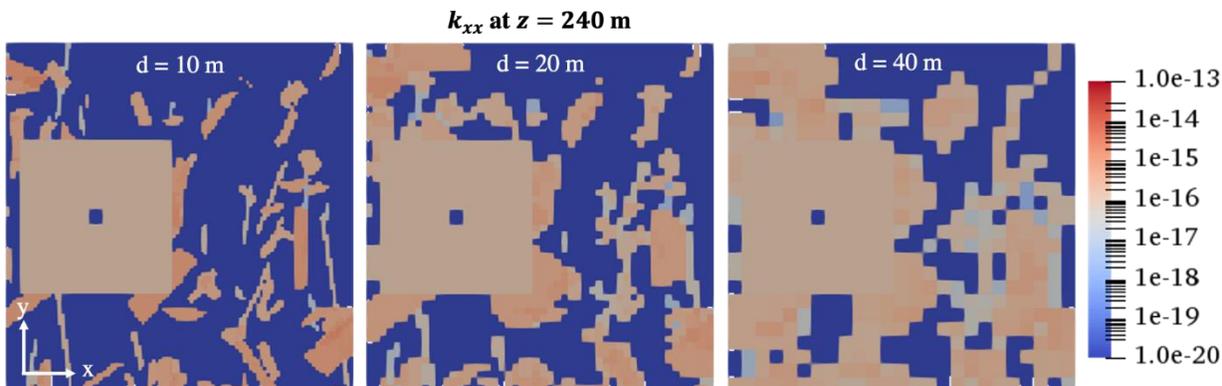
Table 3-1. Only one depth zone was used for this simplified problem because of the small size of the domain. The [-200 m, -400 m] depth zone was chosen because it is the parameterization that was used for the CRC previous to this year.

The CRC uses an unstructured mesh wherein the fractured granite is discretized using 15 m cubes, but the repository is refined to 5 m and 5/3 m cubes so that individual waste packages and drifts are resolved. For the initial formulation of the simplified problem a uniform mesh was applied to the entire domain, and the repository was assumed to simply be a homogeneous region of disturbed rock zone (DRZ), with a single cell in the center representing buffer material and a single waste package. A leak from the waste package occurs at some point in the simulation. To mimic the CRC a glacial aquifer region was placed at the top of the domain and  $^{129}\text{I}$ -concentration-dependent quantities of interest were tracked in the aquifer as a measure of repository performance. A 2D vertical slice of the domain is shown in Figure 2-1.



**Figure 2-1. A vertical slice of the simplified crystalline domain, taken at  $y = 500$  m**

A model hierarchy was designed in terms of the spatial discretization applied to the computational domain. Mesh sizes of 10, 20 and 40 m were used. The mesh size affects the continuum properties which are derived when converting from DFN to equivalent continuous porous medium (ECPM), described in [5]. The continuum values are explicitly scaled by the cell size. Additionally, when two fractures which do not intersect are mapped to the same cell, a false connection is created in the continuum property. The number of false connections increases with cell size, so it is expected that coarser meshes will exhibit increased flow compared to the finer meshes. This effect can be seen in Figure 2-2, which shows a horizontal slice of the permeability in the  $x$  direction ( $k_{xx}$ ), viewed from above, for the three mesh sizes. Note the increasingly large, connected regions of high permeability, especially around the repository region, which increases repository outflow.



**Figure 2-2. 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.**

The domain is defined in terms of four separate subdomains: the fractured granite, the repository (properties defined in terms of DRZ), the buffer/waste package, and the glacial aquifer region. Regions in PFLOTRAN are defined such that if the coordinates of a region lie within a cell, the entire cell is assigned to that region. Because of this, meshes and/or subdomain (regions) must be defined with care to ensure that the physical boundaries of the subdomains are not affected by mesh size. To ensure the physical sizes of the subdomains were consistent across meshes for this work, the minimum length scale for any of the domains was set to 40 m, and the boundaries of the regions were selected to coincide with cell boundaries that were common across all meshes. For the CRC, where subdomain sizes are chosen to reflect reality, adaptive/irregular meshing near boundaries may be necessary to ensure the physical size of subdomains is preserved.

To reduce the computational burden of the simplified problem, it is only run to a 100 kyr time horizon. Additionally, the system is assumed isothermal and only conservation of mass is solved, compared to the CRC, which solves the coupled mass and energy conservation equations. The uncertain parameters and quantities of interest were largely left the same as the CRC; one significant difference is in the treatment of uncertainty in the waste package breach time. For the simplified problem no degradation rate is assumed; rather, an uncertain waste canister breach time is assumed directly. The uncertain parameters and their distributions are identical to those used for the CRC in [1], except the two parameters related to waste package degradation rate were replaced with an uncertain waste package breach time. These parameters and their distributions are reported in Table 2-1.

**Table 2-1. Uncertain parameters in PFLOTRAN problem specification.**

Parameter	Description	Distribution	Crystalline Reference Case	Simplified Crystalline Case
$rateUNF$	Waste form bulk dissolution rate	$\log \mathcal{U}[10^{-8}, 10^{-6}]$	x	x
$kGlacial$	Glacial aquifer permeability	$\log \mathcal{U}[10^{-15}, 10^{-13}]$	x	x

<i>permDRZ</i>	DRZ permeability	$\log \mathcal{U}[10^{-19}, 10^{-16}]$	x	x
<i>permBuffer</i>	Buffer permeability	$\log \mathcal{U}[10^{-20}, 10^{-17}]$	x	x
<i>pBuffer</i>	Buffer porosity	$\mathcal{U}[0.3, 0.5]$	x	x
<i>wpBreachTime</i>	Waste package breach time [yr]	$\mathcal{U}[2500, 10000]$		x
<i>meanWPrate</i>	Mean of degradation rate normal distribution	$\mathcal{U}[-5.5, -4.5]$	x	
<i>stdWPrate</i>	Standard deviation of degradation rate normal distribution	$\mathcal{U}[0.15, 0.4]$	x	

GSA analyses were performed in this work, excluding and including spatial uncertainty in the problem. For the case excluding spatial uncertainty, a single DFN was generated and used as the modeled subsurface for all samples. The only inputs being varied for GSA were those in Table 2-1. For the case including spatial uncertainty, a different DFN realization was used for each uncertain sample to represent spatial uncertainty. To account for this spatial uncertainty in the GSA, graph quantities of interest were included, as proxy variables for the spatial uncertainty, in the set of input parameters over which the PCE was constructed. The procedure by which these quantities were included in the GSA is described in 3.3.

The relative costs of simulation for the different meshes was measured by comparing the core run time required for each (wall-clock run time  $\times$  number of cores). These are reported in Table 2-2.

**Table 2-2. Absolute and relative costs of models in hierarchy.**

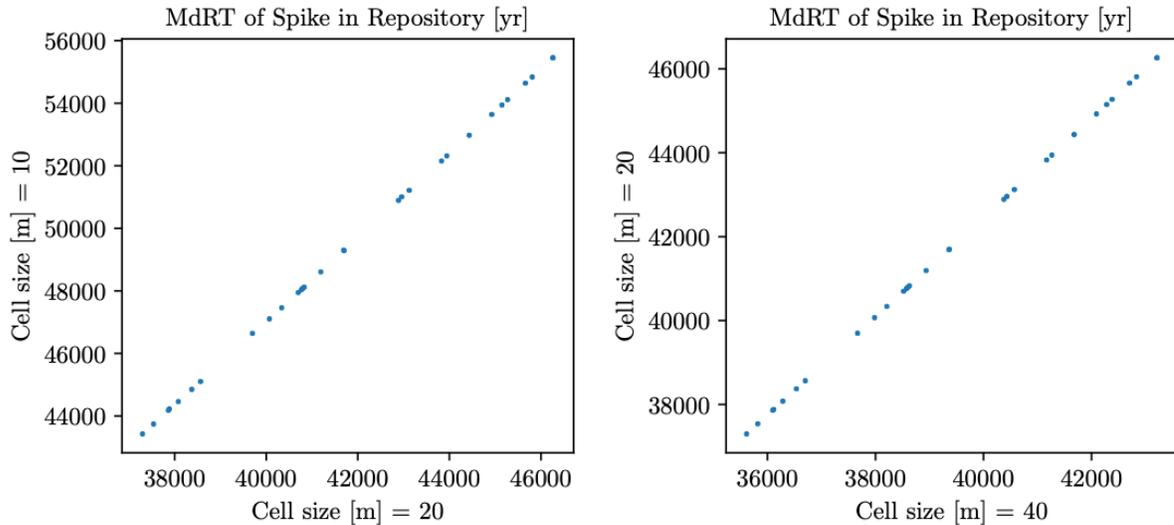
Mesh size [m]	Core time [s]	Relative cost
10	6181	1.0
20	137	$2.2 \times 10^{-2}$
40	36	$5.8 \times 10^{-3}$

## 2.2.2 Assessment of Model Hierarchy

An initial (pilot) assessment of the model hierarchy was performed to gain an understanding of which performance QoIs were expected to be amenable to a multifidelity PCE approach. This was judged based on 30 samples generated for each mesh size, which were used to produce scatterplots comparing the QoIs of the different meshes. This procedure was performed for the case with no spatial uncertainty and for the case with spatial uncertainty.

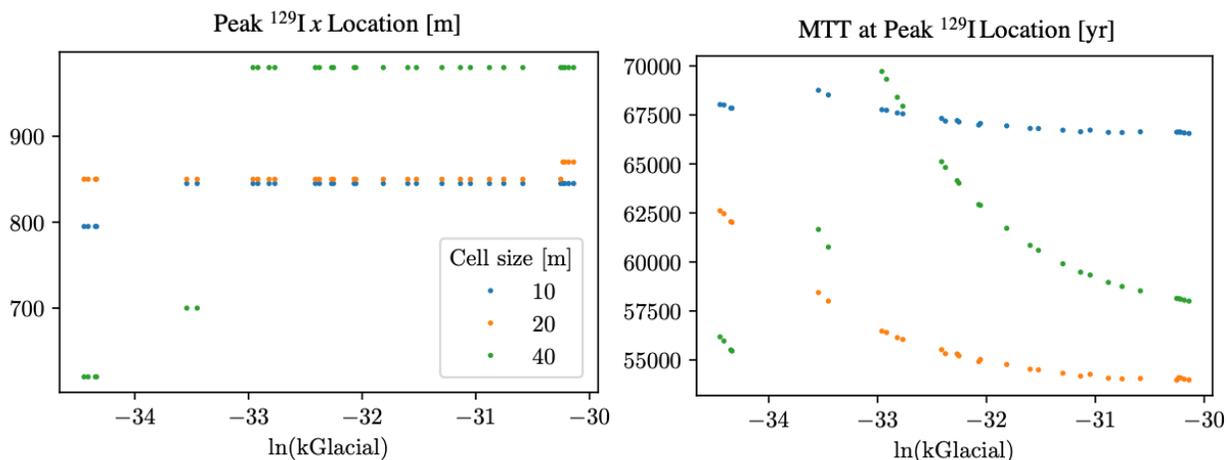
In the case of no spatial uncertainty, the subsurface was left fixed for all samples, so that the only differences were in the uncertain PFLOTRAN input variables. Scatterplots provided a great

deal of information about which performance QoIs were amenable to multifidelity methods and/or to GSA in general. Many of the performance QoIs showed a strong correlation between mesh cell sizes, as shown for example in Figure 2-3. All the mass flow rate QoIs and QoIs related to the residence time of the repository showed a similar significant dependence between meshes.



**Figure 2-3. Scatterplots for the median residence time for the repository for the 20 m vs 10 m mesh (left) and 40 m vs 20 m mesh (right)**

The peak  $^{129}\text{I}$  concentration itself admitted a strong dependence between meshes, but analysis of the QoIs related to the time/location at which the peak occurred, including the mean travel time (MTT) at the peak, indicated they would not be suitable for global-surrogate-based GSA. First, the time of the peak concentration was uniformly 100 kyr, the end of the simulation. Obviously, no sensitivity information can be recovered for QoIs that do not vary as a function of the input parameters. For this reason, the time of the peak was not included in GSA for this work. Additionally, quantities that vary in a discrete fashion can be challenging for surrogates which assume a continuous response to changes in input parameters, such as Gaussian processes and PCEs using continuous basis functions. Such discontinuous behavior was observed here for performance quantities of interest related to the location of the peak  $^{129}\text{I}$  concentration as a function of  $k_{\text{Glacial}}$ , as can be seen in Figure 2-4.



**Figure 2-4. Scatterplots of the  $x$  location of the peak  $^{129}\text{I}$  concentration as a function of  $k_{\text{Glacial}}$  (left) and of the mean travel time at the location of the peak  $^{129}\text{I}$  concentration as a function of  $k_{\text{Glacial}}$  (right), for the 10 m, 20 m, and 40 m meshes.**

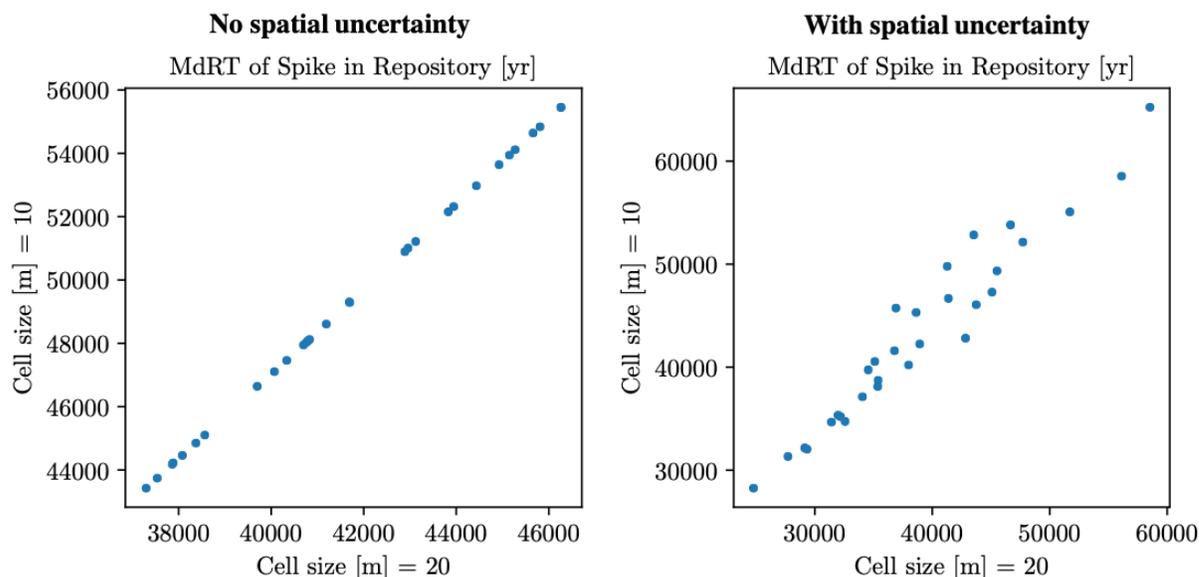
This discrete behavior coincides with very small values of permeability in the glacial aquifer ( $k_{\text{Glacial}}$ ), where the peak  $^{129}\text{I}$  concentration is tracked. The discrete nature of the DFNs means that, once meshed, there are discrete pathways of high permeability formed in the ECPM between the repository and the aquifer. This means that  $^{129}\text{I}$  likely enters the aquifer through the endpoints of these discrete pathways, at which point they are diffused throughout the region. A potential physical explanation for the observed discontinuous behavior is that for very small permeabilities, the extremely slow flushing of water through the aquifer maintains a high concentration at one of these discrete endpoints which is closer to the repository in the  $x$  direction. This subsequently affects the MTT, which is positively correlated with the distance between the repository and the peak location.

It is important to note that this discontinuous behavior is observed for the highest-fidelity model, the 10 m mesh, so even a single-fidelity PCE constructed only at this level would struggle to resolve responses related to the location of the peak. It is also important to note that while this behavior may have a physical explanation, there are still numerical effects related to false connections, since the coarser meshes produce more pathways between the repository and the aquifer, potentially resulting in other locations at which  $^{129}\text{I}$  enters the aquifer. Further analysis examining the 3D time evolution of  $^{129}\text{I}$  plumes and the entry points into the aquifer would be required to confirm this physical explanation for the discrete dependence on  $k_{\text{Glacial}}$  as well as the difference in dependence as a function of mesh. Additional accuracy and consistency between the mesh sizes may be achieved for these quantities of interest by using adaptive meshing, such as octree meshing provided by dfnWorks, to focus resolution around fractures and mitigate false connections. This will be a focus of future work.

To summarize, the QoIs related to mass flow rate and repository residence time are well suited to GSA using global surrogates such as PCE, but those related to the location/time of the peak  $^{129}\text{I}$  concentration are not.

In the case with spatial uncertainty, for each sample of the PFLOTRAN inputs in Table 2-1, a corresponding DFN realization was generated. These PFLOTRAN inputs and the DFN realization

was fixed across the models; the only difference between the models was the cell size of the mesh used to compute the ECPM. Similar behavior was observed in all the performance QoIs for this case, though there was more noise in the relationship between meshes compared to the case without spatial uncertainty. See, e.g. Figure 2-5. The suitability of the performance QoIs for the case with spatial uncertainty are thus expected to be the same as for the case without spatial uncertainty.



**Figure 2-5. Scatterplots for the median residence time for the repository for the 20 m vs 10 m mesh, with no spatial uncertainty (left) and including spatial uncertainty (right).**

## 2.3 Results

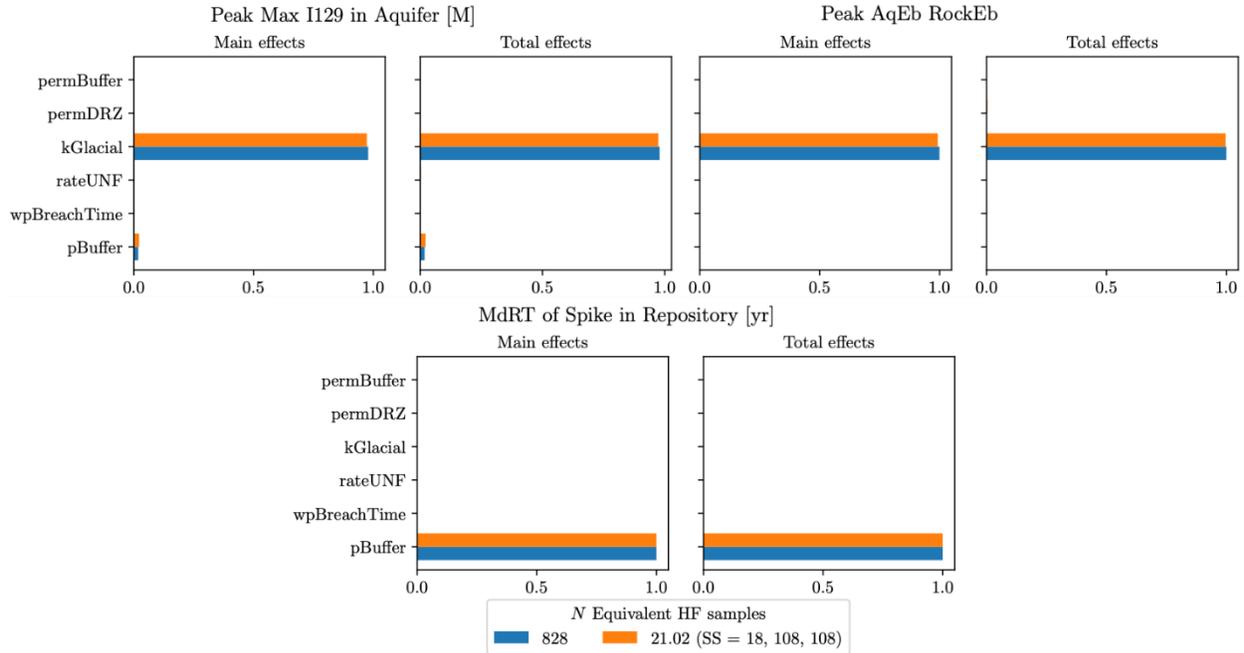
For both the case with no spatial uncertainty and the case with spatial uncertainty a reference single-fidelity PCE constructed using only evaluations from the highest-fidelity model (mesh size 10 m) to use for comparison with the MF PCE. The construction of PCEs was performed with Dakota [7]. The graph metrics that are used as proxy variables for spatial uncertainty are described in Table 3-3 and include average degree of the graph, the number of intersections with the repository, and relative shortest travel time. These graph descriptors cannot be specified *a priori* by a PCE algorithm but must be computed once a DFN has been generated. Because of this, it is necessary to augment an existing set of samples offline with their corresponding graph metric values before passing them to a PCE algorithm. It is known that the graph metrics are only proxies for the spatial uncertainty and likely to not fully capture its effects. This was illustrated in the analysis of correlations between the graph metrics and performance QoIs averaged over epistemic samples in [1], Chapter 4. This analysis indicated that 1) a significant amount of variation in the performance QoIs remained in the performance QoIs after averaging over the epistemic samples, indicating strong dependence on the spatial uncertainty and 2) the maximum correlation between the graph metrics and these averaged performance QoIs was 0.65, meaning no individual graph metric explained more than ~42% of that variation. Because of this it is expected that the performance QoIs will be noisy functions of the inputs. Dakota's cross-validation method for estimating coefficients was deployed to avoid spuriously fitting to this noise.

The goal is to assess whether the GSA results obtained from the single-fidelity PCE can be reproduced by a MF PCE with less computational cost, measured by the equivalent number of high-fidelity evaluations ( $N_{equiv}$ ). The need to augment samples with the graph metric inputs means adaptive algorithms cannot currently be deployed. Instead, an initial campaign of runs to generate 828 samples on all meshes was performed for the cases with and without spatial uncertainty, then convergence analyses were performed by hand-selecting subsets of these 828 samples to produce different sample profiles across the levels. 828 samples were chosen to achieve a collocation ratio of at least 1.15 for a 4<sup>th</sup> order PCE for the case with spatial uncertainty, which had 9 input parameters (samples were generated in batches of 36 to make full use of HPC nodes). For the case with spatial uncertainty, a range of possibilities for the order of PCEs at each level in the MF PCE were considered by modifying the maximum order allowed at each level and using a fixed collocation ratio of 3. For the case without spatial uncertainty the responses were sufficiently easy to resolve that this exploration in terms of order was unnecessary.

The number of samples at each level will be denoted the sample sequence (SS) and may be included in some plots along with  $N_{equiv}$ . The sample sequence will be reported as  $SS = (N_{10}, N_{20}, N_{40})$ . The samples are nested, i.e. the set of samples for the 10 m mesh are a subset of the samples for the 20 m mesh, etc. This means that  $N_{10}$  is equal to the number of samples used to estimate  $R_{10} - R_{20}$ ,  $N_{20}$  is equal to the number of samples used to estimate  $R_{20} - R_{40}$ , and  $N_{40}$  is equal to the number of samples used to estimate  $R_{40}$ .

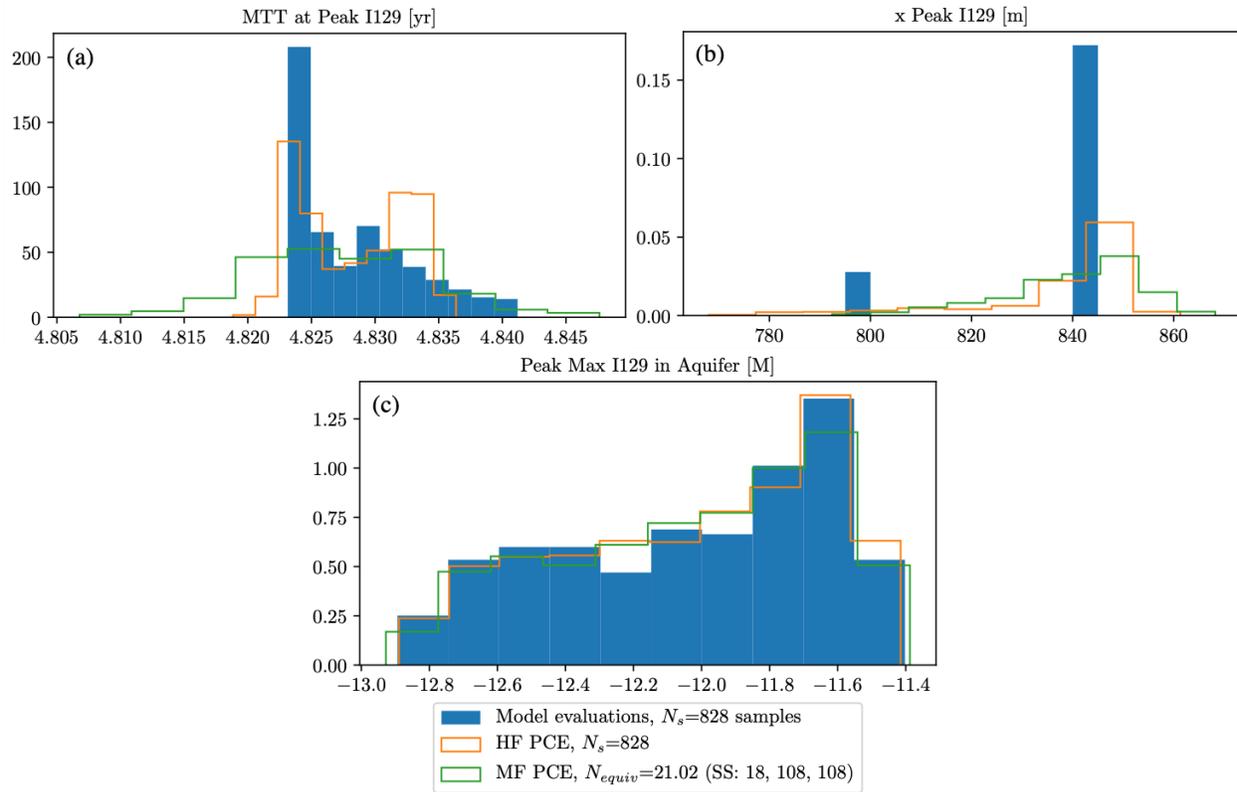
### 2.3.1 No Spatial Uncertainty

For the case with no spatial uncertainty, the reference single-fidelity PCE was constructed on all 828 samples that were generated for the 10 m mesh. The MF PCE's GSA results were compared to the GSA results from the reference PCE. Without the presence of spatial uncertainty, the responses were quite simple and easily resolved. Indeed, the Sobol' indices for the performance QoIs where MF PCE was expected to be suitably recovered those of the reference PCE with very few evaluations at each level ( $N_{equiv} = 21.02, SS = 18, 108, 108$ ). Figure 2-6 shows the comparison of Sobol' indices calculated from the full reference set of 828 high-fidelity samples with that calculated from MF PCE.



**Figure 2-6. Comparison of Sobol' indices from the reference PCE computed with 828 high-fidelity samples (blue) and for the MF PCE (orange).**

For the performance QoIs which were expected to perform poorly, a histogram of samples produced from each PCE compared to the 828 10 m model evaluations are shown in Figure 2-7. Note that, as expected, both the reference and MF PCEs fail to reproduce the nature of the variation of MTT and the  $x$  location of the peak  $^{129}\text{I}$  concentration, as shown in Figure 2-7 (a) and (b), respectively, as compared to a case where the variation is well captured, such as in Figure 2-7 (c). For this reason, neither of the PCEs should be considered reliable for the QoIs related to the time/location of the peak  $^{129}\text{I}$  concentration. These QoIs will be left out of analysis for the remainder of this chapter.



**Figure 2-7. Histograms of sample model (PFLOTTRAN) evaluations compared to samples from a single-fidelity PCE constructed using all 828 samples from the highest-fidelity (10 m mesh) model and samples from a multifidelity PCE constructed using ~21 equivalent high-fidelity evaluations.**

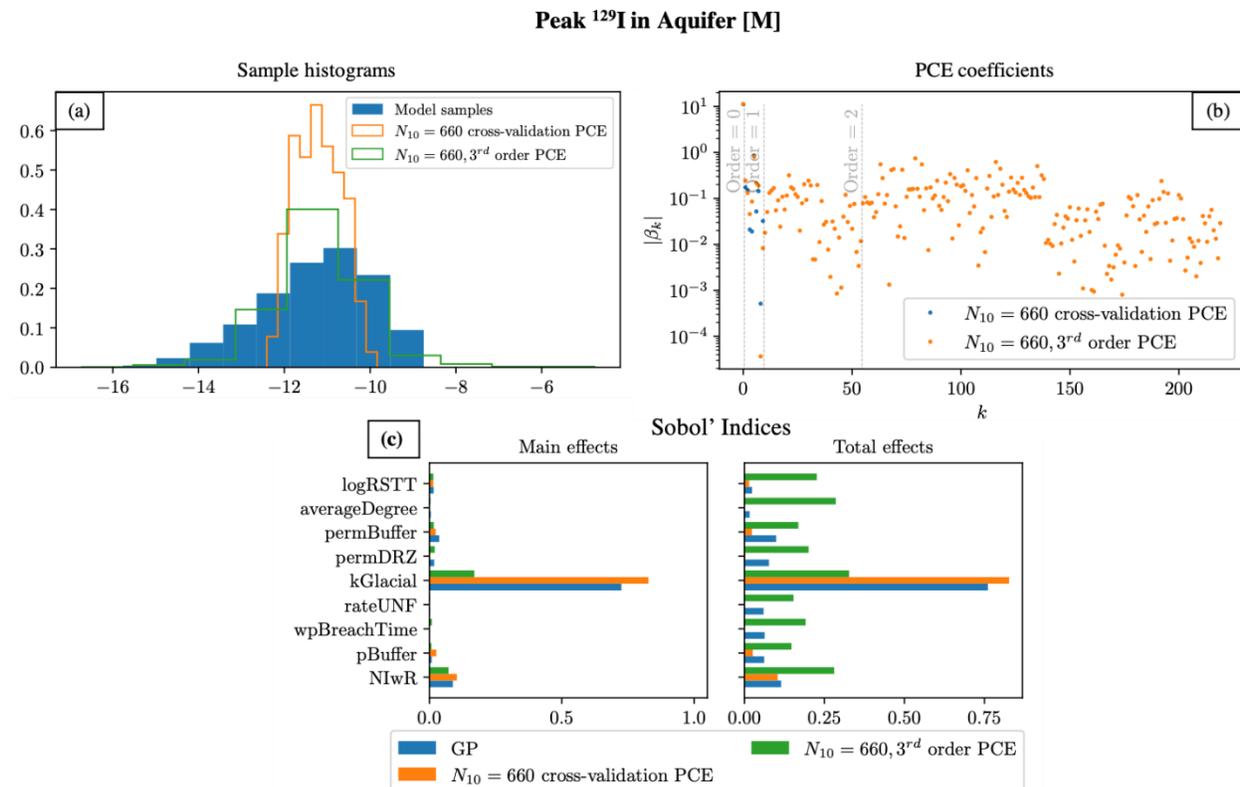
The ability to reproduce the reference PCE GSA results with a greatly reduced computation effort (~39 times less computational cost) with the MF PCE is a promising initial result. Additionally, preliminary scatterplot comparisons were able to identify in advance the performance QoIs for which MF PCE would not be suitable. However, it is expected that the introduction of spatial uncertainty and its proxy variables will make the GSA problem significantly more challenging. To this end, in 2.3.2 spatial uncertainty and its corresponding proxy variables are introduced and the performance of the MF PCE for the performance QoIs deemed suitable to MF PCE is considered.

### 2.3.2 With Spatial Uncertainty

For the case with spatial uncertainty, the subsurface (DFN) realization was varied for each sample. To attempt to account for this spatial uncertainty in the GSA, proxy variables derived from graph representations of the DFN realizations were introduced as input parameters, as discussed in Chapter 3 and defined in Table 3-3. Because it is expected that these proxy variables will capture some, but not all, of the variation induced in performance QoIs from the spatial uncertainty, K-fold cross validation was used in PCE construction to mitigate spuriously fitting to noise. An oversampling ratio of 3 was used to further mitigate this effect.

Using cross validation, the total order is only the maximum order the PCE might take.  $K$ -fold cross validation splits the samples used for regression and splits them into  $K$  sets of equal size (Dakota uses 10 sets). It constructs  $K$  sets of samples where one of the sets is left out in each. A PCE is constructed on each of the  $K$  sample sets and the prediction error of the PCE is computed against the set that was left out. The cross-validation error is computed as the average of these  $K$  prediction errors. The total order of the PCE is selected which minimize this cross-validation error. This procedure is meant to construct the most accurate regression PCE possible without fitting to noise. However, as noise increases, it becomes impossible to distinguish between the noise and genuine higher-order behavior in the response. Because of this, cross validation may choose a PCE which is lower order than the true order of the response.

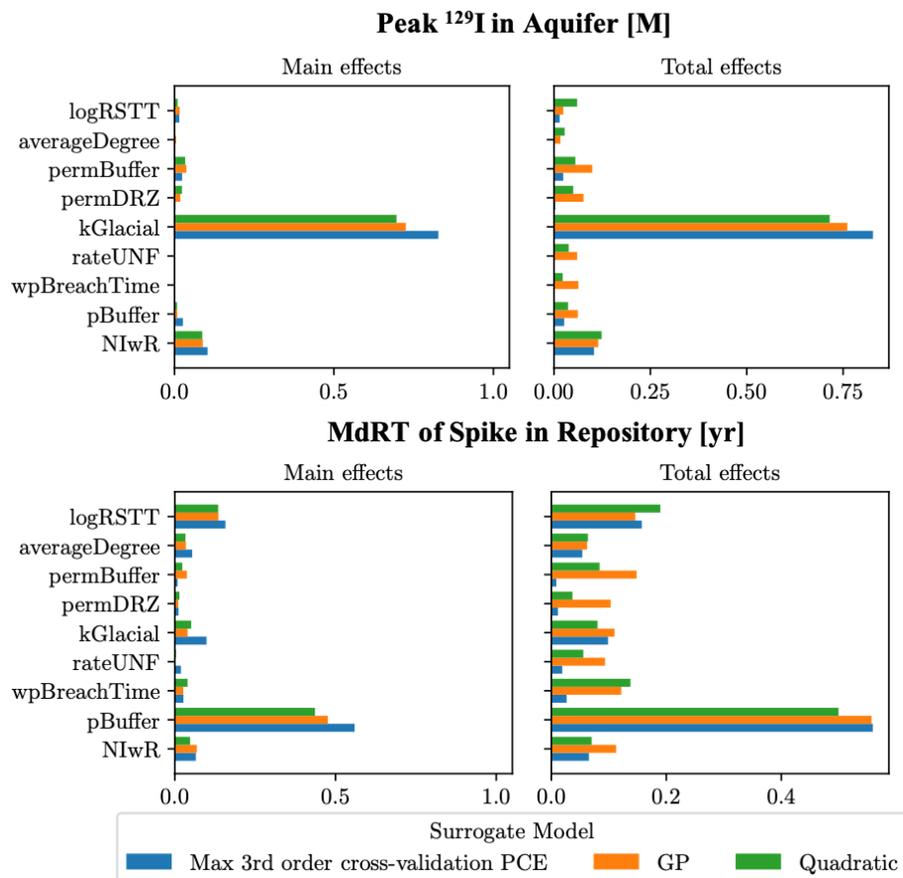
In this case, where the effect of spatial uncertainty is only being partially captured by the proxy graph metric variables, the noise level is quite high. Because of this, the cross-validation procedure selected a 1<sup>st</sup> order PCE for all performance QoIs except for one, the AqEb/RockEb mass flow rate ratio at 100 kyr QoI, for which it selected a 2<sup>nd</sup> order PCE. The PCE is capturing only about half the variance in the model samples while a 3<sup>rd</sup> order PCE without cross validation covers the support of the model samples, as shown in Figure 2-8 (a). However, Figure 2-8 (b) indicates that the coefficients in the 3<sup>rd</sup> order may be spuriously excited in order to achieve this additional variance. The effect of this spurious excitation in higher orders can be seen by comparing to the Sobol' indices computed using a Gaussian process (GP) surrogate and those of the cross-validation PCE with the 3<sup>rd</sup> order PCE, in Figure 2-8 (c). The Sobol' indices from the cross-validation PCE and GP agree quite well, which lends confidence to their values. However, the 3<sup>rd</sup> order PCE's artificially excited coefficients associated with higher-order mixed terms contribute to the computation of its total effects Sobol' indices, resulting in practically equal weight being assigned to all parameters and washing out the relative importance of *kGlacial* that is indicated by the other two surrogates. This illustrates how using too high of an order PCE in the presence of significant noise can produce misleading sensitivity results and motivates the use of the cross-validation PCE as the reference case for this work.



**Figure 2-8. A comparison of the reference max-3<sup>rd</sup> order cross-validation PCE with a 3<sup>rd</sup> order PCE without cross-validation, constructed using the same number of samples. In (a), sample profiles from each PCE is compared to those computed from PFLOTRAN (model).**

**In (b), the coefficients of each PCE (note that the cross-validation PCE coefficients are truncated after 1<sup>st</sup> order). In (c), the Sobol' indices for each PCE compared to those computed from a GP surrogate.**

To further test the accuracy of the Sobol' indices of the cross-validation PCE that was used as a reference for the multifidelity study, a quadratic polynomial and a gaussian process were both constructed using all 828 10-m-mesh samples generated for this case. These surrogates were each used to compute Sobol' indices, and they all were compared for consistency. This comparison is shown for the peak  $^{129}\text{I}$  concentration in the aquifer [M] and the median residence time for the repository [yr] in Figure 2-9. Although the Sobol' indices are not in perfect agreement, they agree on which inputs are most important to each of the QoIs. This lends confidence that the max-3<sup>rd</sup> order PCE computed using cross-validation is producing reasonable results and can be used as a reference for comparison with the MF PCE. However, it should be noted that the noise in the model samples has likely affected the quality of all the surrogates and their ability to capture higher-order behavior in model responses. Addressing the additional variation in performance QoIs caused by the spatial uncertainty will be critical to improving the quality and reliability of future sensitivity analyses and will be the subject of future work.



**Figure 2-9. Sobol' indices for peak <sup>129</sup>I concentration in the aquifer [M] and median residence time of the repository [yr], compared between the cross-validation PCE, the GP, and the quadratic polynomial surrogate used to compute them.**

To compare the max 3<sup>rd</sup> order cross-validation PCE (hereafter called the “reference PCE”) to a multifidelity PCE, cross-validation PCEs with a specified max order and oversampling ratio of 3 were constructed for each of the levels. The number of inputs, including the graph metrics, is 9. The number of terms and samples for a given total-order PCE are reported in Table 2-3. Because the graph metric proxy variables are added to the sample set offline prior to constructing the MF PCE it is currently not possible to use adaptive/greedy algorithms which would add evaluations and higher-order terms to different levels as necessary based on a convergence criterion. Instead, max 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order cross-validation PCEs were constructed at each level. The different orders at each level were assembled in different combinations in postprocessing to produce a range of MF PCEs generated with different numbers of samples at each level. This enabled an analysis of how the MF PCE’s Sobol’ indices converged to those of the reference PCE with different numbers of samples at each level.

**Table 2-3. The number of terms and samples associated with total-order PCEs over varying order with 9 input parameters.**

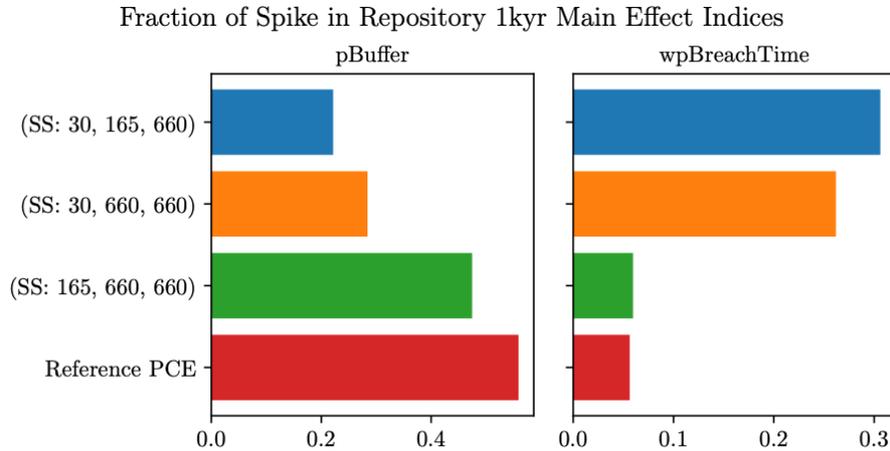
Total order	PCE Terms	Samples (oversampling ratio 3)
1	10	30
2	55	165
3	220	660

Recall that the MF PCE consists of a PCE constructed on the coarsest level, the 40 m mesh, then on recursive discrepancies between levels up to the finest level. For the given model hierarchy, the MF PCE is thus

$$\bar{R}_{10}^{MF} = \bar{R}_{10-20} + \bar{R}_{20-40} + \bar{R}_{40}. \quad (2.8)$$

Results will be reported in terms of the number of equivalent high-fidelity evaluations  $N_{equiv}$  and the sample sequence (SS) at each level, reported as  $SS: N_{10}, N_{20}, N_{40}$ , which indicates the number of samples evaluated on the 10, 20, and 40 m mesh respectively.

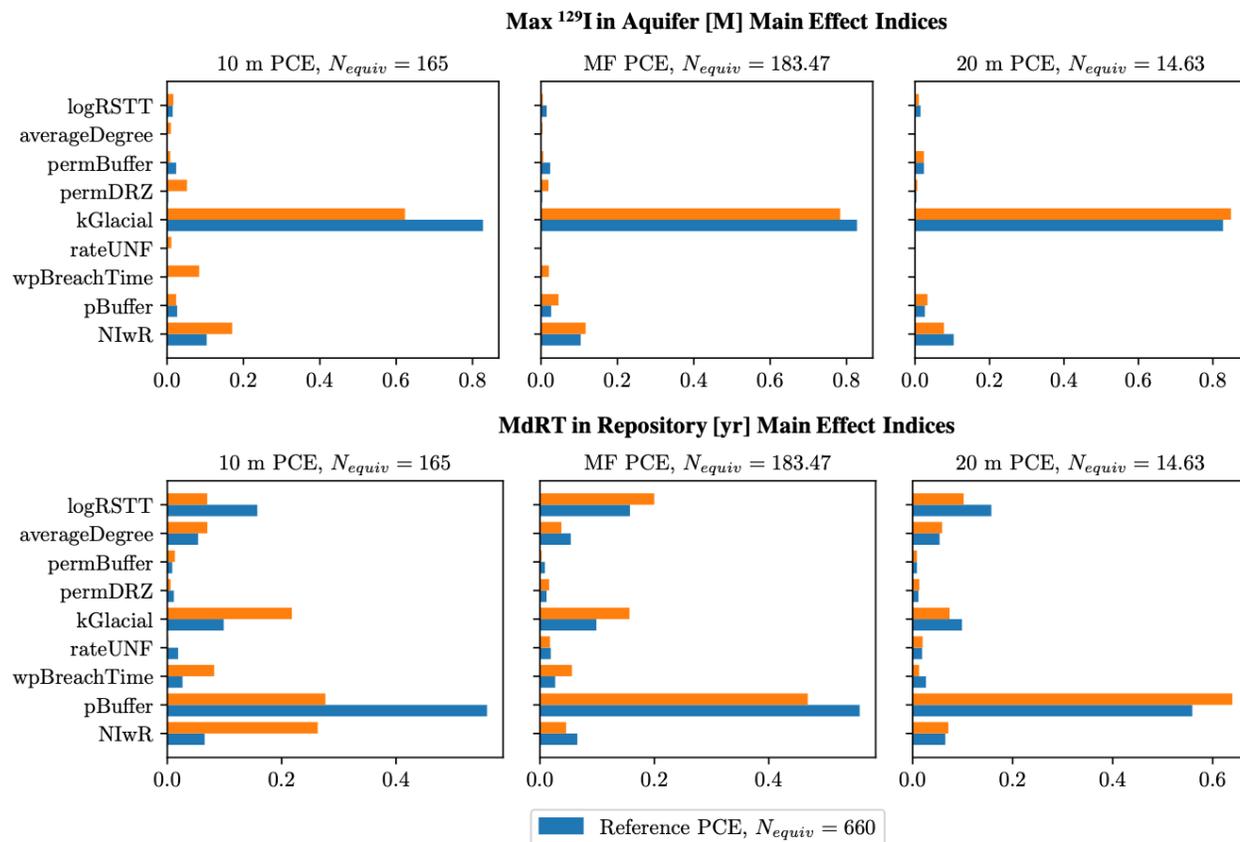
A convergence analysis for the MF PCE was performed by first generating an MF PCE with  $SS=30, 165, 660$  (corresponding to max order 1, 2, 3 over the levels), then incrementing the number of samples at the different levels until there was reasonable agreement in most of the Sobol' indices for most of the QoIs. An example of the convergence of the Sobol main effect indices of  $pBuffer$  and  $wpBreachTime$  as a function of the sample profile for one of the performance QoIs is shown in Figure 2-10. It was determined that the MF PCE constructed with  $SS=165, 660, 660$  was sufficiently in agreement with the reference PCE that it was appropriate to use it for performance comparisons.



**Figure 2-10. A comparison of the main effects indices from three MF PCEs with different sample profiles compared to the reference PCE.**

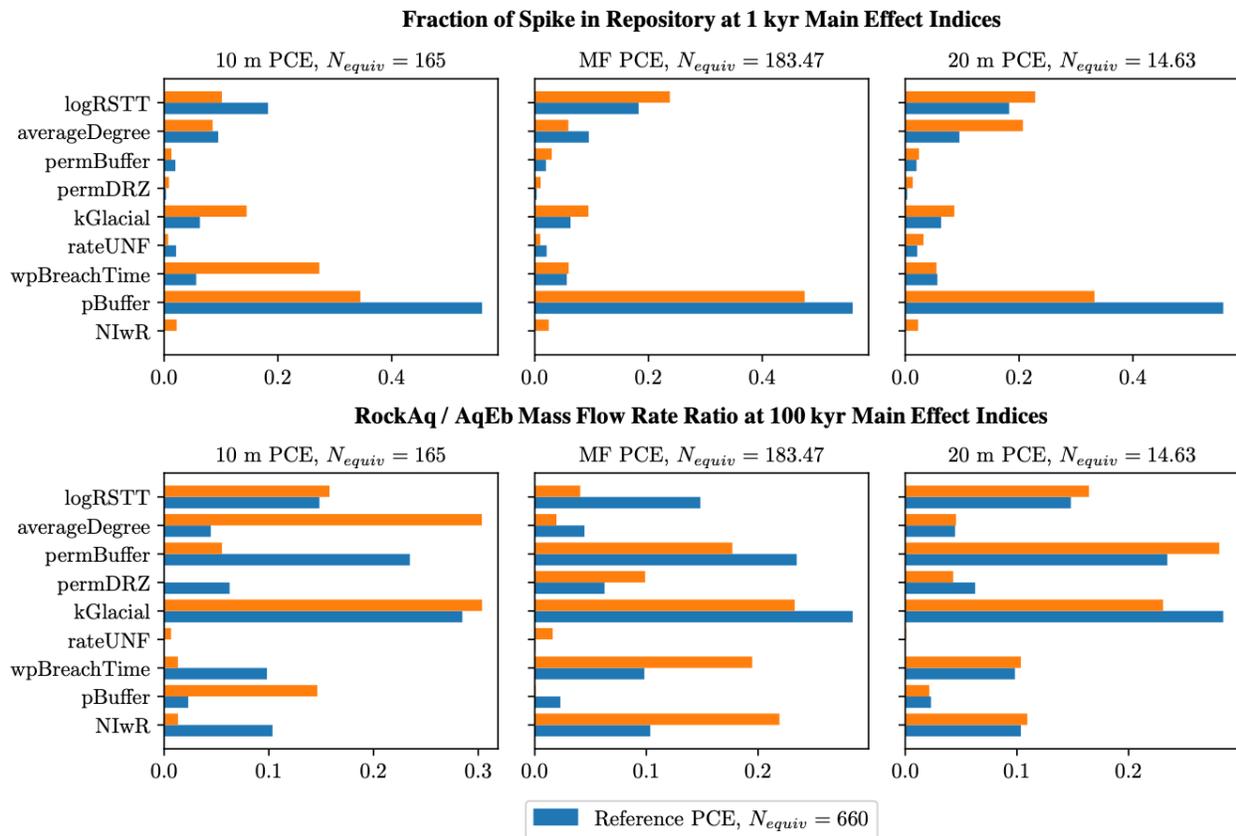
In some cases, a coarser or low-fidelity model is sufficiently converged with respect to the high-fidelity model such that its response to changes in inputs is the same as the high-fidelity model, although there may be a slight bias in the model’s output itself. When this occurs, it may be possible to attain similar sensitivity results to the highest-fidelity model using the coarser model, at a fraction of the cost. This behavior was observed for the 20 m model with respect to the 10 m model for several of the performance QoIs in this problem, so a max 3<sup>rd</sup> order cross-validation PCE constructed from 660 samples on the 20 m mesh was also compared to the reference PCE to explore this possibility. This PCE will herein be referred to as the “20 m PCE” for simplicity. Additionally, a cross-validation PCE was constructed using 165 evaluations on the 10 m mesh to study whether the MF PCE, which augments 165 10 m evaluations with evaluations at the coarser level, improves the sensitivity results. Given the few number of samples used to generate the 165-sample 10 m PCE and the noise in the response, it is expected that it will perform poorly for the GSA. Note that all comparisons of performance between the 20 m PCE, the 10 m PCE, and the MF PCE are *only in terms of the accuracy of their Sobol’ indices*. Comparisons of performance in terms of accuracy of first and second moments were not considered here.

For all the performance QoIs, both the 20 m and MF PCEs yielded significant improvements in the accuracy of the Sobol’ indices compared to the 10 m PCE constructed with only 165 samples, as shown in Figure 2-11. As can also be seen in Figure 2-11, neither the 20 m PCE nor the MF PCE significantly outperformed the other for most of the QoIs. Given the variation seen in comparing the Sobol’ indices between surrogates in Figure 2-9, it would be difficult to argue that either is “closer” to the true values of the Sobol’ indices.



**Figure 2-11. Main effects indices for two representative performance QoIs. Each subplot is a comparison of the Sobol' indices computed from the reference PCE (blue) with those computed from the PCE indicated in the title of the subplot (orange).**

In only two or three cases did either the 20 m PCE or the MF PCE significantly outperform the other. Two examples of this are shown in Figure 2-12. For the fraction of the spike of tracer remaining the repository at 1 kyr, the MF PCE clearly outperformed the 20 m PCE in terms of the relative rankings of the sensitivity of the parameters. Specifically, the 20 m PCE significantly underrepresented the sensitivity in *pBuffer* and overrepresented that of *averageDegree* relative to the reference PCE, while the MF PCE Sobol' indices agree reasonably well for all of the parameters. On the other hand, for the RockAq/AqEb mass flow rate ratio at 100 kyr, the MF PCE significantly underestimates the sensitivity of *logRSTT* and overestimates the sensitivity of *wpBreachTime* and *NIwR*, while the Sobol' indices for the 20 m PCE agree reasonably well overall. It should be noted that both PCEs in this case mis-rank the sensitivity of the parameters in this case. The MF PCE would indicate *wpBreachTime* and *NIwR* were more sensitive than *logRSTT*, and the 20 m PCE would rank *permBuffer* ahead of *kGlacial* in sensitivity. Again, in all cases, either the 20 m PCE or the MF PCE performs better than the corresponding 10 m PCE built with only 165 samples.



**Figure 2-12. Main effects indices for performance QoIs where either the MF PCE or the 20 m PCE outperformed the other. Each subplot is a comparison of the Sobol’ indices computed from the reference PCE (blue) with those computed from the PCE indicated in the title of the subplot (orange).**

This study shows that both a PCE constructed on a coarser mesh and a MF PCE which exploits a hierarchy of meshes have the potential to improve sensitivity analyses when evaluations at the finest level, such as for the production crystalline case, are limited. Whether one approach would be more beneficial for the crystalline case is still unclear. Convergence criteria, either on the surrogate reconstruction or on the Sobol’ indices themselves, would help to make more quantitative comparisons of performance between the two. The ability to assess convergence will also hinge on the ability to account for and quantify the variability in performance QoIs caused by the spatial variability which isn’t fully captured with the proxy graph metric variables. Addressing this challenge will be a focus of future work.

## 2.4 Conclusions and Future Work

- For the case with no spatial uncertainty, MF PCE was able to generate the global sensitivity indices for some performance QoIs (e.g. peak  $^{129}\text{I}$ ) at a fraction (2.5%) of the cost of calculating the sensitivity indices from a full set of high fidelity runs. This is a promising initial result demonstrating the improved computational efficiency from MF PCE. However, the PCE was not able to accurately represent the distribution for some of the performance QoIs which exhibited discrete/discontinuous behavior. This issue is not unique to PCEs; it will be problematic for all continuous surrogates. Different approaches to sensitivity analysis may be necessary for performance QoIs with this characteristic; if similar discontinuities occur for the crystalline reference case this will be a subject of future work.
- Both multifidelity PCE and a single-fidelity PCE constructed on a coarser mesh show potential to improve the efficiency of sensitivity analyses for the crystalline case. However, it should be noted that for multipronged analyses which additionally seek to predict first and second moments of performance QoIs, the single-fidelity PCE on a coarser mesh may be biased while the multifidelity PCE is constructed to minimize bias with respect to the highest-fidelity model. MF PCE methods are advantageous to predict the distribution of the response and associated statistics, not only global sensitivity indices.
- Spatial uncertainty is not fully captured by the input parameters (the graph metrics), but how much variation in outputs should be attributed to variations in inputs is unknown. This makes it challenging to assess surrogate quality. Future work will focus on addressing this challenge, which will in turn enable more rigorous analyses of performance vs. accuracy for the MF PCE.

## 2.5 References: Chapter 2

1. Swiler, L.P., Basurto, E., Brooks, D.M., Eckert, A.C., Mariner, P.E., Portone, T., & Stein, E. (2020). *Advances in Uncertainty and Sensitivity Analysis Methods and Applications in GDSA Framework*. (No. SAND2020-10802R). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States); Sandia National Laboratories, Las Vegas, NV.
2. Sudret, Bruno. "Global Sensitivity Analysis Using Polynomial Chaos Expansions." *Bayesian Networks in Dependability* 93, no. 7 (July 1, 2008): 964–79. <https://doi.org/10.1016/j.res.2007.04.002>.
3. Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer physics communications*, 181(2), pp.259-270.
4. Ng, L. W. T., & Eldred, M. (2012). Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA* (p. 1852).

5. Stein, Emily, Jennifer M Frederick, Glenn Edward Hammond, Kristopher L Kuhlman, Paul Mariner, and S David Sevougian. "Modeling Coupled Reactive Flow Processes in Fractured Crystalline Rock." In *Proceedings of the 16th International High-Level Radioactive Waste Management Conference*. Charlotte, North Carolina, 2017.  
<https://www.osti.gov/servlets/purl/1417242>.
6. Adams, B.M., Bohnhoff, W.J., Dalbey, K.R., Ebeida, M.S., Eddy, J.P., Eldred, M.S., Hooper, R.W., Hough, P.D., Hu, K.T., Jakeman, J.D., Khalil, M., Maupin, K.A., Monschke, J.A., Ridgway, E.M., Rushdi, A.A., Seidl, D.T., Stephens, J.A., Swiler, L.P., and Winokur, J.G., "[Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.14 User's Manual](#)," Sandia Technical Report SAND2020-12495, May 2021.
7. Dalbey, K.R., Eldred, M.S., Geraci, G., Jakeman, J.D., Maupin, K.A., Monschke, J.A., Seidl, D.T., Swiler, L.P., Tran, A., Menhorn, F., and Zeng, X., "[Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.14 Theory Manual](#)," Sandia Technical Report SAND2020-12496, May 2021.

### 3. DISCRETE FRACTURE NETWORKS

Discrete fracture network (DFN) modeling is an alternative to continuum approaches for simulating flow and transport through sparsely fractured rocks in the subsurface [1]. Continuum approaches use effective parameters to include the influence of fractures on the flow. In contrast, a DFN approach involves a network of fractures where the geometry and properties of individual fractures are explicitly represented as lines in two dimensions or planar polygons in three dimensions. These generated networks are then meshed for computation. *dfnWorks* [1] is an *open source* parallelized computational suite developed at Los Alamos National Laboratory to generate three-dimensional discrete fracture networks and simulate flow and transport. The work presented in this report uses the “*dfnGen*” capability of *dfnWorks* to generate the networks, with the flow and transport calculations run in PFLOTRAN. This capability was used to generate 25 discrete fracture networks for a crystalline reference case, which is based off properties at the Forsmark site in Sweden [3] but is not a direct replica of Forsmark. They were then converted to equivalent porous media (ECPMs) for flow and transport simulation with PFLOTRAN using a Sandia-developed open-source code, *mapDFN* 0. These 25 DFNs are used in the uncertainty analysis presented in Chapter 4.

The DFN specification used in Chapter 4 has the following changes with respect to the specification described in [2]:

- Converge on target  $P_{32}$  instead of number of fractures per fracture family;
- Corrected a bug in *dfnWorks* input which was incorrectly using radians instead of degrees for fracture Fisher distributions;
- Updated assumed relationship between fracture transmissivity and fracture radius to include depth-dependent parameterization.

#### 3.1 Transmissivity Relationships

Versions 2.3 and greater of *dfnWorks* support the specification of transmissivity relationships by fracture family. This enabled the transmissivity relationship to be varied by depth in this year’s study, compared to previous versions which assumed a single relationship for the entire computational domain. We present a comparison of the correlated constant transmissivity vs. the correlated depth-dependent transmissivity relationship in Section 3.4.

The fractures in the DFNs generated for this work, as well as for Forsmark, are assumed to be circular in shape. They are parameterized in terms of their radius and orientation, the fracture intensity  $P_{32}$  [ $\text{m}^2/\text{m}^3$ ], and the relationship between fracture transmissivity and fracture size (radius). The crystalline reference case parameterization is based on a subset of the fracture families and sets defined in [3]; the parameterization is reported in Table 3-1, save for the transmissivity relationship.

**Table 3-1. Crystalline reference case DFN parameterization**

Depth (meters below sea level) / Fracture Set Name	Orientation: Fisher Distribution for Poles			Size: Truncated Power Law for Radii			P <sub>32</sub>  P <sub>32</sub> , r ∈ [0.038,564] [m <sup>2</sup> /m <sup>3</sup> ]
	Mean Trend	Mean Plunge	κ	Power Law k	Min Radius r <sub>0</sub> [m]	Max Radius r <sub>upper</sub> [m]	
0-200 / NS	90°	0.0°	22	2.5	30	500	0.073
0-200 / NE	180°	0.0°	22	2.7	30	500	0.319
0-200 / HZ	360°	90.0°	10	2.4	30	500	0.543
200-400/ NS	90°	0.0°	22	2.5	30	500	0.142
200-400/NE	180°	0.0°	22	2.7	30	500	0.345
200-400 / HZ	360°	90.0°	10	2.4	30	500	0.316
>400 / NS	90°	0.0°	22	2.5	30	500	0.094
>400 / NE	180°	0.0°	22	2.7	30	500	0.163
>400 / HZ	360°	90.0°	10	2.4	30	500	0.141

Previous iterations of the crystalline reference case have assumed a single so-called “correlated” transmissivity relationship for all fracture families. The correlated relationship is defined as

$$T = ar^b, \quad (3.1)$$

where  $r$  is the fracture radius and  $T$  is the transmissivity. The use of one relationship for all fracture families was because previous versions of dfnWorks only supported a single relationship for the entire computational domain. However, different parameterizations for the correlated relationship are provided for each depth zone in [3]. The correlated transmissivity relationship for the three depth zones associated with fracture families 1 and 6, defined in [3], Table 2, were used to generate the DFNs. Specifically, the parameterization for the middle depth zone was previously used for the crystalline reference case.

The fracture transmissivity is used to determine the continuum permeability field when converting to ECPM, so it can significantly affect the flow and transport properties of the system. To understand the effect of adding depth dependence to the transmissivity relationship, this study compares performance quantities of interest (QoIs), permeabilities, and flow and transport properties for ECPMs generated from the same underlying DFN, where the only change is the

assumed transmissivity relationship. The assumed transmissivity relationships are reported in Table 3-2.

**Table 3-2. Transmissivity relationship parameterizations considered in this study**

Depth (meters below sea level)	Depth Zone (dz) Number	Transmissivity relationship	
		Correlated, $T = ar^b$	
		Constant over domain ( $a, b$ )	Depth-dependent ( $a, b$ )
0-200	dz1	(1.6e-9, 0.8)	(6.7e-9, 1.4)
200-400	dz2		(1.6e-9, 0.8)
>400	dz3		(1.8e-10, 1.0)

Each discrete fracture network is randomly generated, so the corresponding performance QoIs are also random variables depending on the random DFN. To account for this, multiple DFNs will be generated as part of the study and the ensemble of QoIs and other properties will be used in the analysis, rather than considering only one DFN.

### 3.2 Graph Metrics characterizing the DFNs

DFNs, by the nature of being networks, admit two types of graph representations: in the “fracture graph,” fractures are treated as nodes and intersections as edges; in the “intersection graph,” intersections are treated as nodes and fractures as edges. These representations can be used to derive metrics which represent the physical characteristics of a given DFN. These graph metrics were used in [2] to capture spatial uncertainties in GSA for the crystalline reference case. The relevant graph metrics are defined and described in Table 3-3. Based on observed correlations between performance QoIs and the graph metrics, as well as marked differences in GSA as a result of their incorporation, a process to integrate the graph metrics more fully into analyses was developed this year.

**Table 3-3. Graph Metrics used in crystalline reference case GSA.**

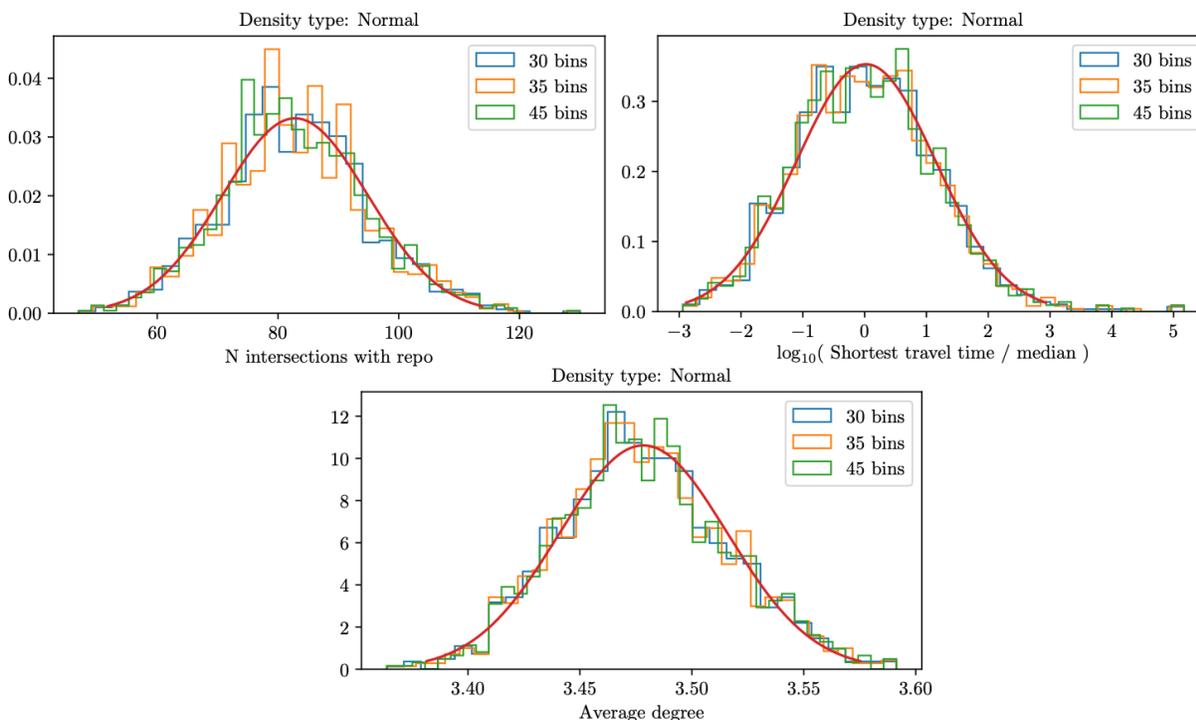
Graph Metric	Description
<i>logRSTT</i>	The log <sub>10</sub> -transformed relative shortest travel time between repository and aquifer. Relative shortest travel time computed by scaling shortest travel time for each DFN by the median. A measure of ease of flow between repository and aquifer.
<i>averageDegree</i>	Average number of intersections per fracture. A measure of how connected the network is over the entire domain.
<i>NIwR</i>	Number of fractures intersecting the repository. A measure of number of potential flow pathways out of the repository region.

To this end, two of the graph metrics which exhibited the most correlation with performance QoIs in [2] were the shortest travel time between repository and aquifer and the average degree of the fracture graph. The shortest travel time is a rough estimate computed using the dfnWorks “graph flow” capability and thus is only relevant as a relative comparison of travel times amongst DFNs. The average degree metric is the average number of intersections of which a fracture is part. Based on the significant correlations in the previous production studies, 5 of the 25 DFNs generated for this year were selected to achieve approximate 0.05, 0.5, and 0.95 quantiles in these two graph metrics. The 5 DFNs were selected as follows: 1080 DFNs were generated and their graph metrics were computed. These samples were used to compute 0.05, 0.5 and 0.95 quantiles for each of the graph metrics. The DFN which had the smallest  $l^2$  misfit between its graph metrics and both median values was selected as a representative median DFN. To select the DFN to represent the 0.05 quantile for average degree, the DFN whose average degree had the smallest  $l^2$  misfit with the computed 0.05 quantile was chosen. The same procedure was followed for shortest travel time and 0.95 quantiles. When choosing the representative DFNs for each of the quantiles, the value of the other metric was not taken into consideration. In future work it may be advantageous to select DFNs which enforce the other graph metric being near its median value, or to enforce that both metrics are near their quantile values. Whether these 5 DFNs exhibit any sort of bounding behavior is assessed in Chapter 4.

Besides shortest travel time and average degree, the number of intersections with the repository was computed from the DFN graph representation and used in GSA. In an advancement from the previous year’s approach, this year empirical densities for the graph metrics were computed using the 1080 samples, which will enable their incorporation into a polynomial chaos expansion (PCE) for GSA. This procedure is described in Section 3.3. Note that the graph metrics included in uncertainty analyses this year are based on the graph metrics which were correlated with the performance QoIs in the uncertainty analyses last year. However, the performance QoIs have been modified this year. A follow-up analysis was performed to confirm the continued correlations of the graph metrics with the updated set of performance QoIs.

### 3.3 Incorporating Graph Metrics into Polynomial Chaos Expansion

PCEs are stochastic expansions constructed over parameter space. Their basis functions are defined to be orthonormal with respect to the density of the uncertain parameters. Because of this, an analytical density must be provided for any input parameter over which the PCE will be constructed. However, the graph metric densities are unknown *a priori*; instead, they must be inferred from samples. To enable their incorporation in PCE-based global sensitivity analysis, empirical densities were fit to samples computed by generating a large number (>1000) of DFN realizations and computing their graph metrics. Appropriate densities and scalings of the graph metrics were chosen by visual inspection of histograms. The fitted densities for the crystalline reference case are shown in Figure 3-1.



**Figure 3-1. Normal density fits to the graph metrics for the crystalline reference case, along with histograms using a varying number of bins to show the goodness of fit over a range of bin sizes.**

The shortest travel time (STT) was highly skewed towards large values, and the values were large overall. Additionally, the STT was computed very approximately with an arbitrarily-chosen pressure differential between repository and aquifer—because of this it is only meaningful when considered as a relative (nondimensional) quantity. To mitigate these issues, the STTs were scaled by the median value computed over all samples, then the base-10 logarithm was taken. Let this scaled value be denoted *logRSTT*, for log-relative shortest travel time. A *logRSTT* of 0 can thus be interpreted as the median value, while a *logRSTT* of 3 corresponds to a STT which is three orders of magnitude greater than the median.

In this case it was clear by visual inspection that normal densities would be suitable representations of the underlying distributions for the graph metrics, however in general this may not be true. In future it may be necessary to develop a framework which automatically tests a range of transformations, scalings, and density types to find optimal analytical distributions for the graph metrics.

### 3.4 Analysis of Transmissivity Relationships

#### 3.4.1 Approach

In this section, we compare the correlated-constant transmissivity with the correlated depth-dependent transmissivity. To complete this analysis and understand the effect of adding depth dependence to the transmissivity relationship, 20 DFNs were generated. The two transmissivity

relationships were applied to each DFN and used to construct 40 ECPMs, 20 for each relationship. The ECPMs were then used in PFLOTRAN simulations of the crystalline reference case. Note that this analysis used the same crystalline reference case described in Chapter 4, but the epistemic uncertain parameters (which are varied in the Chapter 4 sensitivity analysis) were fixed at their mean values. This limited the scope to focus on the effect of the DFN transmissivity relationships. The crystalline QoIs and permeabilities for each DFN were then used for a comparative analysis. To allow for an easier interpretation of the results, the continuum permeability fields were first separated based on depth zone and then summarized into mean permeability tensor components in the x, y, and z directions, denoted  $k_{xx}$ ,  $k_{yy}$ , and  $k_{zz}$  respectively. Then, a geometric mean permeability ( $k_{gm}$ ) for each depth zone was determined using the following equation:

$$k_{gm} = \sqrt[3]{k_{xx}k_{yy}k_{zz}}. \quad (3.2)$$

To reiterate, the depth zones are with respect to the bottom of the problem domain and a summary is reported in Table 3-4 below. A glacial aquifer is positioned at the top of depth zone one and has a depth of 15 meters (ranging from 1245 m to 1260 m). The repository is located in depth zone three, centered at 675 m in z.

**Table 3-4. Depth zone definitions**

Depth Zone	Position with respect to z-axis
dz1	1060 m – 1260 m
dz2	860 m – 1060 m
dz3	0 m – 860 m

Additionally, there are a total of 15 QoIs that can be examined and compared in this study. For simplicity, only 6 QoIs were chosen to be included in this analysis. These are defined further in Section 4.2.7.

- Peak Total I129 [M]. Peak  $^{129}\text{I}$  concentration in aquifer in mol/L (M)
- MdRT of Spike in Repository [yr]. Median residence time of a tracer in the repository.
- Fraction of Spike in Repository 1Myr. The fraction of a tracer remaining in repository at one million years.
- Fractional Mass Flux from Repository [ $\text{yr}^{-1}$ ]<sub>3kyr</sub>. The instantaneous fractional loss rate of a tracer remaining in repository at 3000 years.
- AqEb RockEb 1Myr. The ratio of two water fluxes at one million years: the aquifer to east boundary flux and the rock to east boundary flux.
- RockAq RockEb 1Myr. The ratio of two water fluxes at one million years: the rock to aquifer flux and the rock to east boundary flux.

Python scripts were used to extract the QoIs for each of the PFLOTRAN results and to extract the  $k_{xx}$ ,  $k_{yy}$ , and  $k_{zz}$  permeability tensor and calculate the mean values per depth zone. This allowed the statistical comparisons presented in the following section.

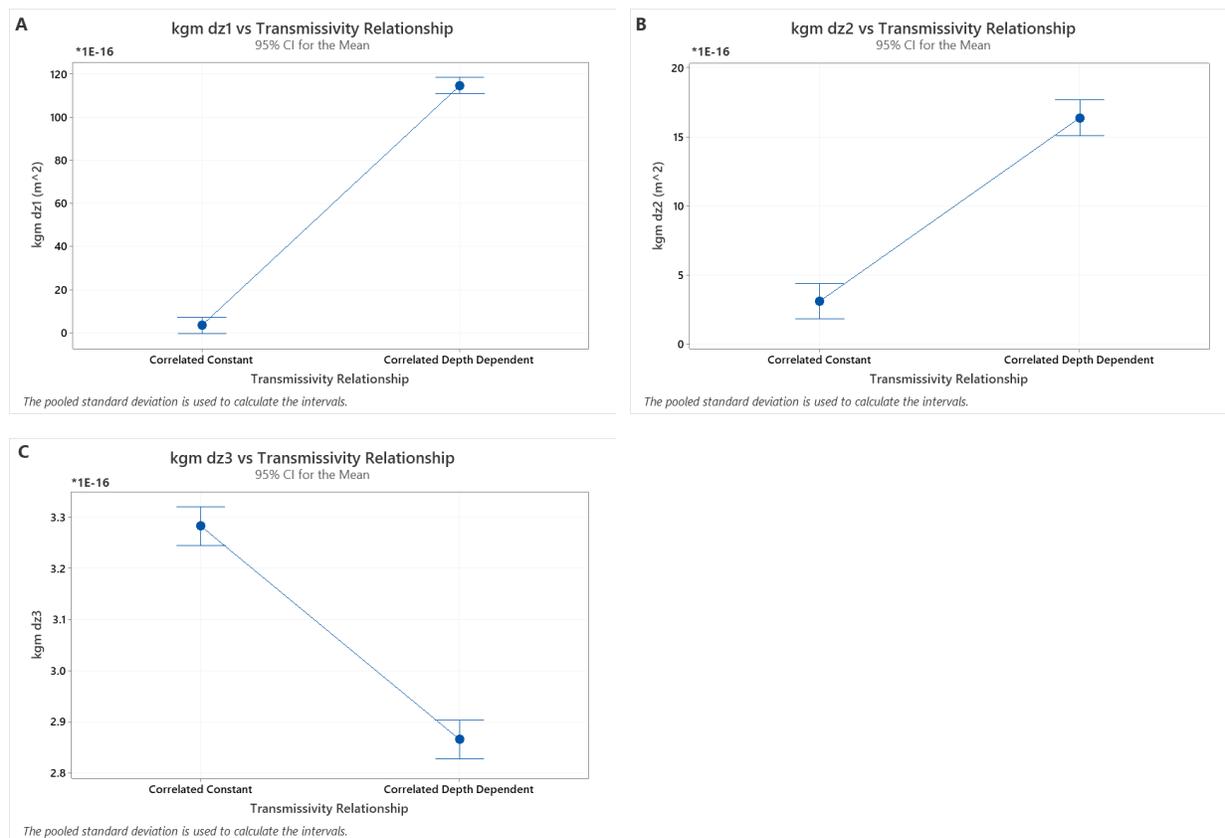
### 3.4.2 Results

To determine the effect of the different transmissivity relationships, the geometric mean permeabilities and specified QoIs were visually compared using interval plots, boxplots, and scatterplots. The primary goal was to determine if the results using the 20 DFNs with the correlated constant transmissivity relationship were significantly different from the results using the 20 DFNs with the depth-dependent transmissivity relationship.

The information shown in the subsequent plots is representative of all 40 cases: outliers were not removed. Interval plots were used to examine the mean values for each type of data and the associated 95% confidence interval on the means. When examining the interval plots in Section 3.4.2.1, one sees the sample mean (and its associated confidence interval) for the correlated constant case and for the depth-dependent case. Determining if the confidence intervals for the means of the two groups are overlapping gives a good indication of whether the depth-dependent transmissivity relationship influenced the results significantly. For example, one sees that the interval plots for the geometric means of the permeability fields for all three depth zones are statistically significantly different using the two types of DFN transmissivity relationships.

To give a better interpretation of what the information in the interval plots (also called “main effects” plots) is indicating, boxplots were constructed as well to show the actual spread of the data for each transmissivity relationship. These plots are treated similarly to interval plots and the level of overlapping between the groups indicates the amount of influence the transmissivity relationship had on the results. Lastly, scatterplots were used to further investigate some specific relationships between the permeabilities and QoIs.

### 3.4.2.1 INTERVAL PLOTS

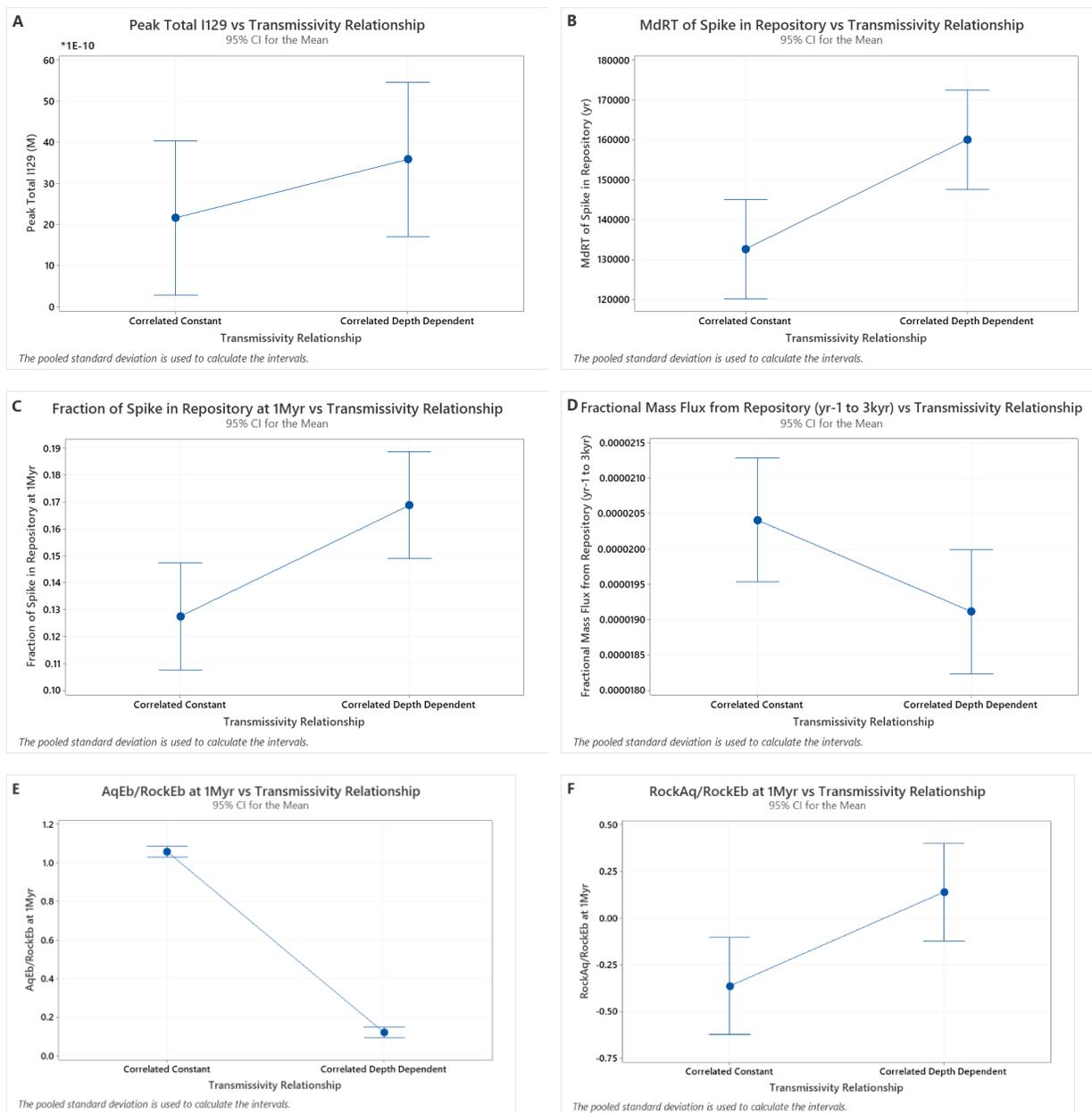


**Figure 3-2. Interval plots for the geometric mean permeabilities of each depth zone.**

As can be seen in Figure 3-2 above, the geometric mean permeabilities for all three depth zones show that there is a considerable difference between transmissivity relationships. Note that the means on these plots are scaled: they have been divided by 1.E-16 for ease of comparing the differences. Also note that the mean values of the  $k_{gm}$  for dz1 and dz2 for the correlated, depth-dependent DFNs are significantly higher than for the correlated constant DFNs. In the case of dz1, for example, the mean  $k_{gm}$  for the depth dependent relation is two orders of magnitude more than the mean  $k_{gm}$  for the constant correlated relationship (115E-16 vs. 3E-16). This difference in  $k_{gm}$  for dz2 was unexpected, since the parameterization did not change between the two relationships. The difference arises because the transmissivity relationship for each depth zone is assigned to an entire elliptical fracture based on which depth zone its centroid occupies. This means that very large fractures whose centroids are near the boundary between dz1 and dz2, but whose centroid is in dz1, are assigned a much larger transmissivity than those in dz2. However, these large fractures extend into dz2, and therefore their large transmissivities and corresponding continuum permeabilities contribute to the domain average in dz2. Methods to identify and filter out the contributions from fractures in dz1 could mitigate this effect.

In contrast, the mean correlated constant  $k_{gm}$  is higher than the depth-dependent  $k_{gm}$  for dz3. However, the absolute magnitude of this difference is quite small for dz3: about 0.5E-16. Additionally, the small confidence intervals for all datasets indicates little variance from the sample mean further supporting the hypothesis that the geometric mean permeabilities for each

depth zone are statistically different for the two transmissivity relationships. This observation is to be expected considering the difference between the transmissivity relationships detailed in Table 3-4. A final observation is that the correlated constant transmissivity relationship had permeability values of the same order of magnitude in the X, Y, and Z direction. In contrast, for the depth-dependent transmissivity relationship, the X and Y directions of the permeability tensor were larger than the Z direction by two and one orders of magnitude, respectively, for both depth zones one and two. We would expect to see an increase in downstream flow and little increase in vertical flow towards the aquifer for the depth-dependent transmissivity relationship.



**Figure 3-3. Interval plots for the specified quantities of interest.**

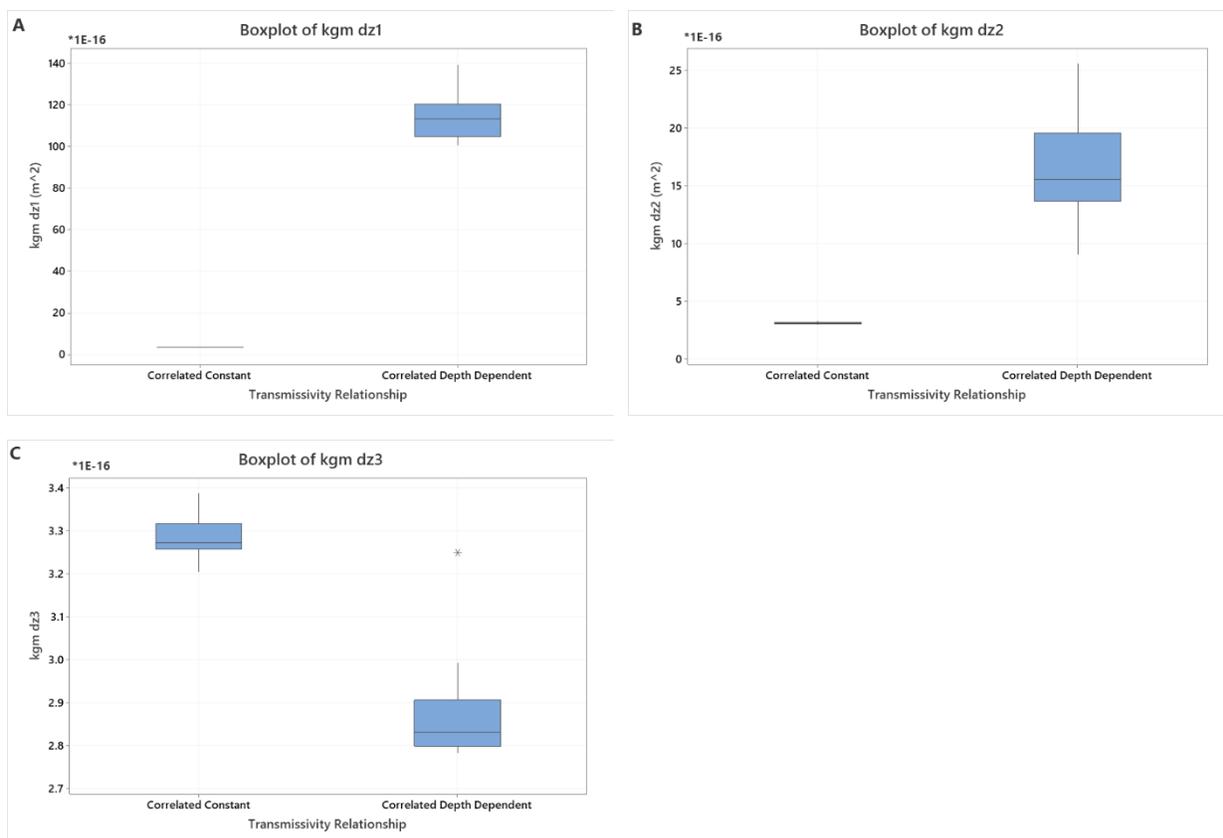
The interval plots for the 6 specified quantities of interest shown in Figure 3-3 above are not as definitive as the interval plots for the geometric mean permeabilities. Aside from Figure 3-3E,

the confidence intervals for the two transmissivity relationships are overlapping or are very close to overlapping for each QoI. Additionally, the confidence intervals are quite large, indicating high variance in the mean estimates of these quantities. Based on this information, it is not accurate to say the transmissivity relationship is statistically significant for five of the QoIs.

Figure 3-3E, on the other hand, shows that the transmissivity relationship is statistically significant for this one QoI specifically, given the large difference in sample means and the very small confidence intervals. This difference is to be expected since the mass flow rate from the aquifer to the east boundary is likely minimally affected by the DFN while the mass flow rate from the rock to the east boundary was increased significantly for the depth dependent transmissivity relationship due to the large increase in permeability of the rock in depth zone 1.

As for Figure 3-3F, while it is expected that the flow from the rock to the east boundary increases for the depth-dependent transmissivity relationship, there is little difference in the mean flux ratios between the two relationships. This implies that the flow from the rock to the aquifer increased as well. A further investigation of the unscaled flow rates would be needed to know for certain.

### 3.4.2.2 BOXPLOTS

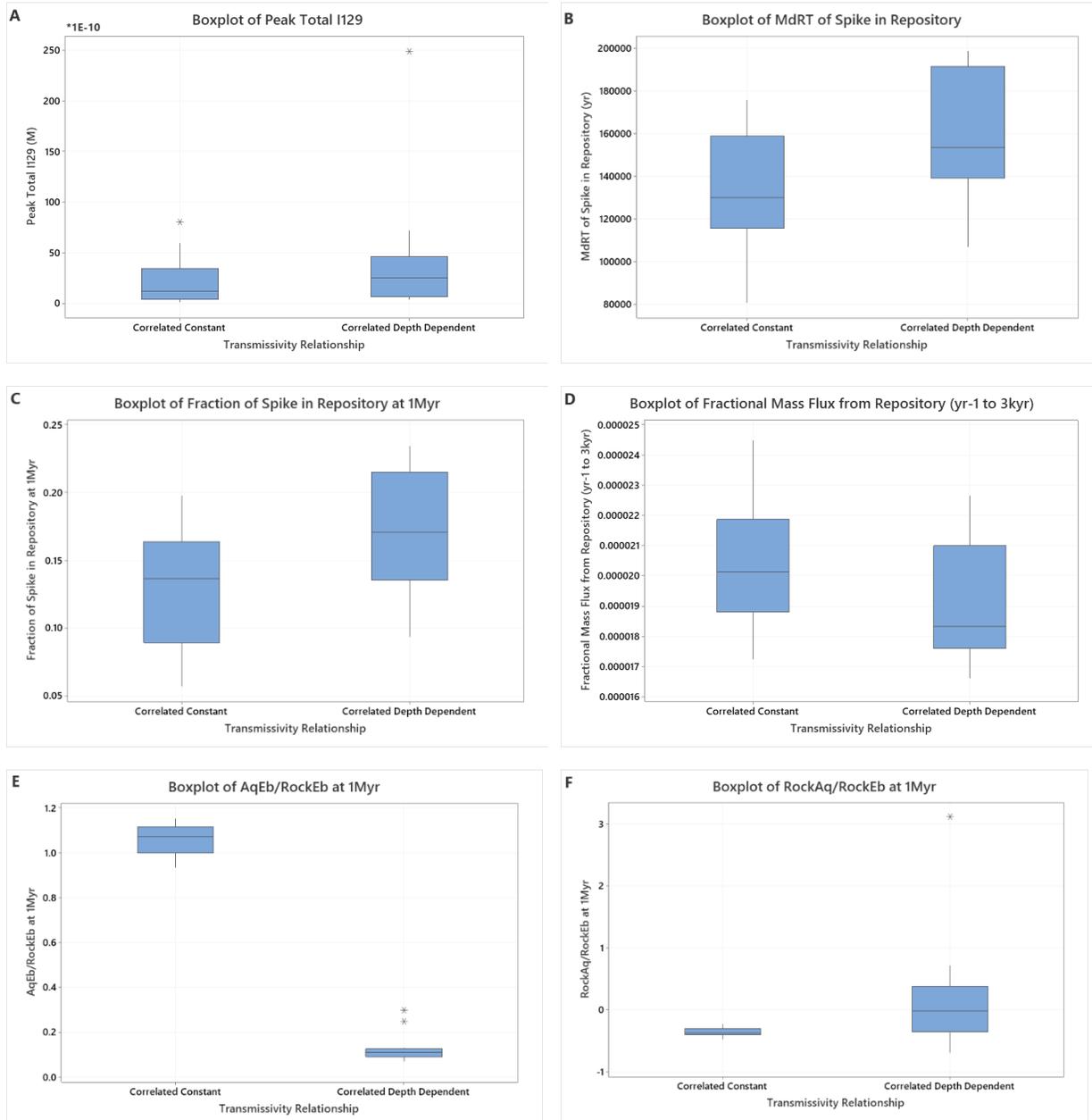


**Figure 3-4. Boxplots for the geometric mean permeabilities of each depth zone.**

The results shown in Figure 3-2 are further supported when examining the information shown in the boxplots in Figure 3-4. In the boxplots, the middle line denotes the median value, the 25<sup>th</sup> and 75<sup>th</sup> percentiles are shown in the shaded boxes, and the whiskers encompass most of the data

(technically, whiskers cover the maximum and minimum data points within 1.5 times the box height). The stars designate outliers: these are data points that stand apart from the rest of the data.

All three geometric mean  $k_{gm}$  permeability data sets show a considerable difference between the two transmissivity relationships. However, it is worth noting the outlier in Figure 3-4C for the depth dependent transmissivity relationship. This one value is overlapping with the spread of data shown for the constant transmissivity relationship.

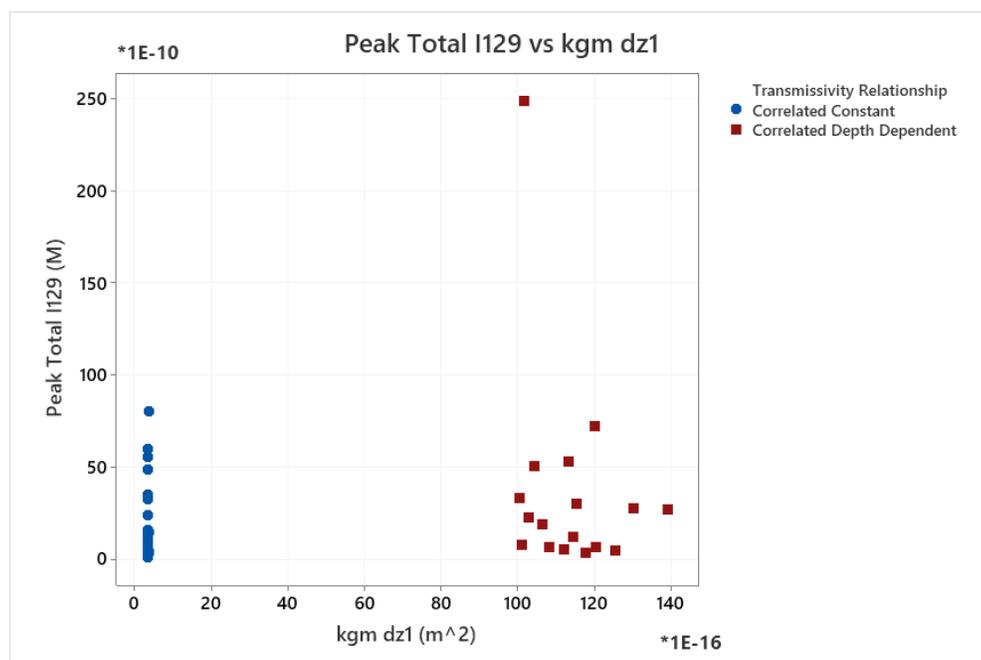


**Figure 3-5. Boxplots for the specified quantities of interest.**

The results shown in Figure 3-5 above further solidify the conclusions made from the results in Figure 3-3. We note that a few of the DFNs result in outlier QoI values that are quite different from the others. For example, in Figure 3-5A, we see that range of the Peak total  $^{129}\text{I}$  data is fairly

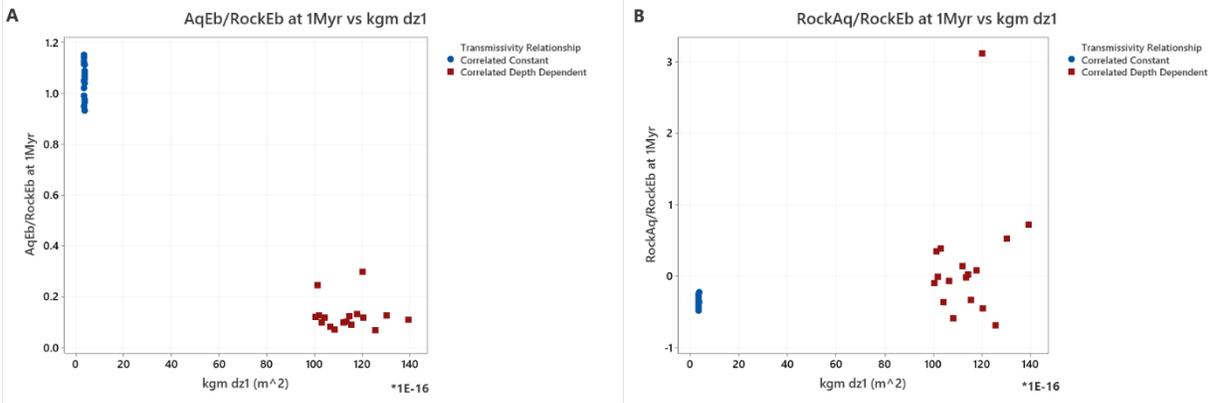
consistent between the correlated constant and the correlated depth-dependent cases. The depth-dependent case has a slightly higher median, and its distribution is not quite as skewed as the correlated constant case. However, there is one outlier in the depth-dependent case in Figure 3-5A with a value of  $250E-10$  M, where the other results are mostly below  $75E-10$  M. This outlier can significantly increase and influence the mean estimate and shown in Figure 3-3A. Without the outlier, the means of these two datasets would be closer. Similarly, the outlier in Figure 3-5F for the depth-dependent transmissivity relationship could account for part of the difference between the two transmissivity relationships shown in Figure 3-3F.

### 3.4.2.3 SCATTERPLOTS



**Figure 3-6. Scatterplot displaying the relationship between the total maximum <sup>129</sup>I concentration in the aquifer and the geometric mean permeability for depth zone one.**

The relationship shown in Figure 3-6 is consistent with the results seen thus far. Due to the small differences between DFN relationships shown in Figure 4-A (except for the outlier), it was decided to further investigate the relationship between maximum <sup>129</sup>I concentration in the aquifer and the geometric mean permeability for depth zone one since this depth zone had the largest difference in permeability for the two transmissivity relationships. As can be seen in the plot, besides the one outlier, the increase in permeability showed no significant effect on the <sup>129</sup>I concentration. However, it is noted that large concentrations were observed for this quantity of interest in Chapter 4; further evaluations to better characterize the distributions would help to determine whether the depth-dependent relationship had a significant effect on the peak <sup>129</sup>I concentration.



**Figure 3-7. Scatterplots displaying the relationship between (A) the ratio of the mass flow from the aquifer to the east boundary to the mass flow rate from the rock to the east boundary (at 1 million years) and (B) the ratio of the mass flow rate from the rock to the aquifer to the mass flow rate from the rock to the east boundary (at 1 million years) with respect to the geometric mean permeability for depth zone one.**

The relationship between the geometric mean permeability for depth zone one and the QoIs was further assessed for the ratios of mass flow rates to the east boundary for the rock and aquifer at 1 million years. Again, the relationships shown in Figure 3-7 are consistent with the results seen thus far.

### 3.4.3 Conclusions from correlated-constant vs. depth-dependent transmissivity study

In conclusion, the effect of a depth-dependent transmissivity relationship on permeability and various QoIs when compared to a constant transmissivity relationship was examined using 20 DFNs resulting in 40 total ECPMs. The results indicated a statistically significant difference for the permeabilities of each depth zone and the ratio between the mass flow rate from the aquifer to the east boundary and the mass flow rate from the rock to the east boundary. The five other QoIs examined showed no statistically-significant difference in means between the two transmissivity relationships. The depth-dependent transmissivity relationship did not strongly influence a change in these quantities. This may be due to the small differences in the geometric mean permeability in dz3 with the depth-dependent transmissivity compared with the constant correlated transmissivity: we hypothesize the QoIs relating to repository performance did not show much change given the small change in dz3 permeability. Based on the results of this study, it is concluded that repository performance characteristics are mostly independent of the transmissivity relationship used for these 20 DFNs specifically. Further work will investigate additional DFN samples and comparison of quantile estimates.

### References: Chapter 3

- [1] Hyman, J. D., S. Karra, N. Makedonska, C. W. Gable, S. L. Painter and H. S. Viswanathan (2015). "dfnWorks: A discrete fracture network framework for modeling subsurface flow and transport," *Computers & Geoscience*, **84**:10-19.
- [2] Swiler, L.P., Basurto, E., Brooks, D.M., Eckert, A.C., Mariner, P.E., Portone, T., & Stein, E. (2020). *Advances in Uncertainty and Sensitivity Analysis Methods and Applications in GDSA Framework*. (No. SAND2020-10802R). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States); Sandia National Laboratories, Las Vegas, NV.
- [3] Steven Joyce, Lee Hartley, David Applegate, Jaap Hoek, and Peter Jackson. Multi-scale groundwater flow modeling during temperate climate conditions for the safety assessment of the proposed high-level nuclear waste repository site at Forsmark, Sweden. *Hydrogeology Journal*, 22(6):1233–1249, 2014.
- [4] Stein, Emily, Jennifer M Frederick, Glenn Edward Hammond, Kristopher L Kuhlman, Paul Mariner, and S David Sevougian. "Modeling Coupled Reactive Flow Processes in Fractured Crystalline Rock." In *Proceedings of the 16th International High-Level Radioactive Waste Management Conference*. Charlotte, North Carolina, 2017. <https://www.osti.gov/servlets/purl/1417242>.

## 4. CRYSTALLINE REFERENCE CASE: UPDATES

For a nuclear waste repository located in crystalline rock, a major source of uncertainty in performance assessment is the spatial heterogeneity of potential fracture flow paths through the host rock. Conceptually, a long-lived radionuclide (or radionuclides produced within radioactive decay chains) released from a waste package (WP) will initially migrate through the buffer material and into the surrounding damaged rock zone (DRZ). From there it will migrate along the DRZ until it enters a fracture that takes it farther into the host rock, where connected fractures can provide a path to a nearby fracture zone. It might then migrate along this fracture zone and through connected fracture zones to the biosphere. Along the flow path, the radionuclide will undergo radioactive decay and ingrowth and diffuse into and out of dead-end pores and fractures. Additionally, depending on its properties and the environmental conditions along the flow path, it will adsorb and desorb from colloids and immobile mineral surfaces, chemically react with aqueous species, possibly change oxidation state, and, if solubility-limited, precipitate and dissolve.

This chapter presents a new uncertainty analysis of a reference case repository for commercial spent nuclear fuel in fractured crystalline rock. The reference case is identical to that in Stein et al. [1], Mariner et al. [4], and Sevougian et al. [9] except for improvements to the fracture network implementation (discussed in Chapter 3), the design of the uncertainty analyses (Section 4.1), and modifications to the model domain (Section 4.2). Note that this analysis builds extensively on the crystalline reference case uncertainty analyses performed in 2019 and 2020 and documented in Chapter 8 of [14] and Chapter 4 of [22].

### 4.1 Uncertainty Analysis (UA)

The uncertainty analysis includes multiple types of uncertainty: spatial uncertainty, parameter epistemic uncertainty, and aleatory uncertainty. Uncertainties described as spatial uncertainty relate to the unknown structure of the repository and surrounding rock. The spatial uncertainty is not strictly categorized as epistemic or aleatory, in part because it incorporates some of both types of uncertainties. Uncertainties described as parameter epistemic uncertainty relate to inputs with fixed but unknown scalar values.

Section 4.1.1 discusses different perspectives on the categorization of spatial uncertainty and Section 4.1.2 summarizes the structure and implementation of the UA.

#### 4.1.1 Uncertainty Categorization

The categorization of spatial uncertainty in this analysis is an issue of debate. In this report, spatial heterogeneity in the discrete fracture networks was separated from most of the parameter uncertainty because one of the goals of the analysis was to understand the relative contributions of the spatial heterogeneity and parameter uncertainties to total uncertainty.

The categorization of the spatial uncertainty depends on the definitions of aleatory and epistemic uncertainty, the intent of the crystalline reference case, the difference between theoretically and practically reducible uncertainty, and historical usage. With respect to definitions, epistemic uncertainty is lack-of-knowledge uncertainty, also called reducible uncertainty, and aleatory uncertainty is stochastic or irreducible uncertainty. Aleatory uncertainty

can also be described with respect to time; it is uncertainty about future occurrences, not the present state [20]. However, outside of PA, this connection to time is not always included in the definition [21].

One perspective on spatial uncertainty is to treat it as an epistemic uncertainty. This assumes that the underlying fracture network at the crystalline reference case site is fixed, but we do not know what it is. Our knowledge of the fracture network is imperfect: we cannot perfectly characterize the fractures due to limitations of geophysical imaging capabilities. As discussed in Chapter 3, we have used the *dfnWorks* software [3] to generate 25 realizations that are consistent with field data and with what we understand about crystalline formations and their properties. We cannot say that any one of the 25 DFNs we have generated is the exact real fracture network, but they are possible realizations each intended to represent a realistic possibility.

An alternative perspective on spatial uncertainty is to treat it as an aleatory uncertainty. This argument is based on two concepts: 1) predicting the distribution of random fractures through space is analogous to predicting the occurrence of random events through time, and 2) uncertainties that may be reducible philosophically are not necessarily practically reducible. The *dfnWorks* software produces random fracture networks that are consistent with a set of fixed characteristics specified for the site (this is done by generating random fractures from user specified distributions on fracture size, orientation, etc. and placing them in the fracture network, see Chapter 3 for details). Analogous to future random events, there are infinitely many fracture networks that can occur in space which are consistent with these fixed characteristics. This can be perceived either as inherent randomness or model uncertainty: the characteristics are insufficient to define a fixed fracture network.

The perspective that spatial uncertainty should be treated as aleatory also relates to the practical motivation for separating uncertainty. Epistemic uncertainty can be reduced, so separating epistemic uncertainty from aleatory uncertainty identifies where uncertainty could be reduced, as well as how much uncertainty cannot be eliminated. Even with the most advanced measurement technologies, it is not feasible to measure and model the true fracture network at a potential repository site. Though the fracture network is fixed, and so the uncertainty is hypothetically reducible, it cannot be reduced practically. If the goal behind separating out epistemic uncertainty is to identify uncertainty that can actually be reduced, it makes more sense to include uncertainty in the fracture network as an aleatory uncertainty.

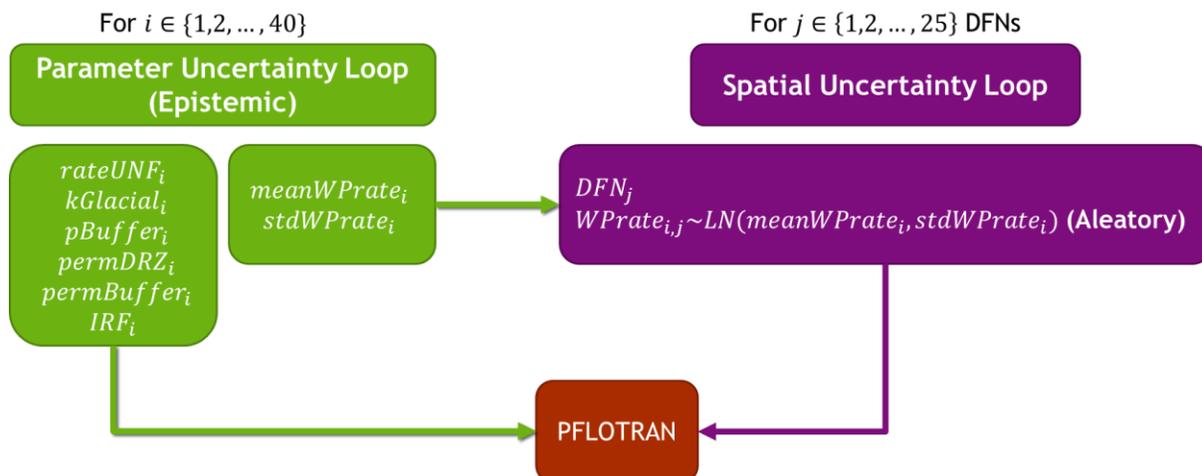
Finally, the categorization of spatial uncertainty is further complicated by the introduction of another type of spatial heterogeneity that clearly mixes epistemic and aleatory components: uncertainty in the order in which waste packages breach in the repository (e.g., this is a spatial variability in the sense that waste package 37 may breach first, followed by waste package 483, etc. These waste packages are at different locations in the repository). Uncertainty in waste package corrosion rate is introduced by sampling from a truncated log-normal distribution for the base corrosion rate in a temperature-dependent rate expression for general corrosion [4]; this log-normal distribution for the (spatially distributed) waste package corrosion rate is sampled within PFLOTRAN [6, 7] and is considered an aleatory uncertainty (random, stochastic, or irreducible uncertainty). Both the mean (*meanWPrate*) and standard deviation (*stdWPrate*) of this log-normal distribution, however, are treated as epistemically uncertain inputs.

This analysis also contains epistemic uncertainty in parameter values, which were categorized without debate. Epistemic uncertainties include porosity and permeability of the bentonite buffer (*pBuffer* and *permBuffer*, respectively), permeability of the DRZ (*permDRZ*) and of the overlying sedimentary unit (*kGlacial*), the rate of spent (used) nuclear fuel dissolution (*rateUNF*), and the instant release fraction (*IRF*) upon WP breach. Note that the epistemic parameter uncertainty is simpler to treat than the fracture network uncertainty. For epistemic parameter uncertainty, we assume that the parameters have a fixed but unknown value for this case study. That is, there is an assumed value for porosity of the bentonite buffer, but we do not know what it is, and the distribution characterizes our uncertainty about that value. Thus, we generate possible realizations of a number for this porosity by sampling the distribution. Though the categorization of these parameters as epistemic was unanimous, we acknowledge that there is nuance to the types of uncertainties inherent to these properties; there will be some spatial variability to these properties which is currently assumed not to exist because it is eclipsed by state-of-knowledge uncertainty.

We refrain from making an explicit statement about the categorization of spatial uncertainty in this report. Spatial heterogeneity of the fracture networks is treated as aleatory in the analysis because it is insightful to separate out the influence of the fracture networks from the epistemic parameter uncertainties. This treatment of spatial uncertainty in the sampling structure is motivated by utility, not by the type of uncertainty. The implementation of spatial heterogeneity includes epistemic and aleatory uncertainties, and there are philosophical arguments for either categorization. Regardless of how it is characterized, the spatial uncertainty was separated from the parameter uncertainty in a two-loop structure for this UA. To avoid implying a categorization, the sampling loop with spatial heterogeneity is referred to simply as the spatial loop.

#### 4.1.2 Uncertainty Implementation

The UA comprises a spatial loop of sample size 25 fracture network realizations, and a parameter loop of sample size 40 for a total of 1000 simulations. Note that both loops involve epistemic uncertainties, but the spatial uncertainty loop involves the spatial heterogeneity exhibited by the 3-D variability in the fracture networks and the epistemic parameter loop involves the parametric uncertainty associated with the epistemic parameters. Because the aleatory uncertainty is also spatial, it is included in the spatial loop. For each DFN realization in the spatial loop, a different random seed is used when sampling the waste package corrosion rate distribution, so that the order of waste package breach associated with each DFN is different (see Section 4.2.3). The sampling loops are shown in Figure 4-1.



**Figure 4-1. Sampling loops used in the Uncertainty Analysis of the Crystalline Reference Case presented in this Chapter.**

Latin hypercube sampling of epistemically uncertain parameters is performed using Dakota [8]. The uncertainty analysis includes  $pBuffer$ ,  $permBuffer$ ,  $permDRZ$ ,  $kGlacial$ ,  $rateUNF$ , and  $IRF$  as epistemically uncertain inputs, and both  $meanWPrate$  and  $stdWPrate$  are also included as epistemically uncertain inputs in the parameter loop. The uncertain parameters are listed in Table 4-1. The outputs include maximum  $^{129}\text{I}$  concentration in the aquifer over time and location as well as quantities such as mean travel time, median residence time, and flow rate ratios as defined in Table 4-4. These Quantities of Interest (QoIs) have been implemented and are the subject of this study (see Table 4-4 and the discussion in Section 4.2.7 below).

In previous UAs for the crystalline reference case, the same set of epistemic samples (e.g., same 40 sample vectors for the epistemic variables) were run for each of the spatial realizations [14, 22]. This sampling structure was chosen to allow explicit separation of effects from the epistemic uncertainty and from spatial heterogeneity. However, the repeated use of the same epistemic samples for each spatial realization comes at the cost of covering less of the epistemic sample space and interferes with surrogate model fitting for global sensitivity analysis due to the repeated values. This sampling approach was changed this year to sample a different set of epistemic samples for each spatial realization, so the UA consists of 40 different epistemic sample vectors for each of the 25 spatial realizations, resulting in 1000 unique sample vectors for the epistemic variables. This increases coverage of the epistemic sample space and reduces interference from repeated values in the global sensitivity analysis. The cost of this change is that it is now more difficult to quantify how much of the QoI uncertainty is due to spatial heterogeneity versus epistemic parameter uncertainty. The change in sampling structure is not necessarily, therefore, an improvement upon the design from last year. Rather, it is an alternative structure targeted towards answering a different question. The epistemically uncertain parameters and their distributions are listed in Table 4-1.

**Table 4-1.** Uncertainty distributions propagated in crystalline reference case UA.

Input	Description	Range	Units	Distribution	Sampling Loop
<i>DFN</i>	Particular realization of a discrete fracture network	1-25			Spatial loop
<i>Aleatory_I</i>	Aleatory sampling to generate a fixed ordering of waste package degradation for each DFN	0-1		Uniform	Spatial loop
<i>rateUNF</i>	Fractional dissolution rate of spent (used) nuclear fuel	$10^{-8} - 10^{-6}$	yr <sup>-1</sup>	log uniform	Parameter loop
<i>kGlacial</i>	Glacial till permeability	$10^{-15} - 10^{-13}$	m <sup>2</sup>	log uniform	Parameter loop
<i>pBuffer</i>	Buffer porosity	0.3 – 0.5	-	Uniform	Parameter loop
<i>permDRZ</i>	DRZ permeability	$10^{-19} - 10^{-16}$	m <sup>2</sup>	log uniform	Parameter loop
<i>permBuffer</i>	Buffer permeability	$10^{-20} - 10^{-17}$	m <sup>2</sup>	log uniform	Parameter loop
<i>meanWPrate</i>	Mean of the truncated log normal distribution on base normalized general corrosion rate ( <i>R</i> )	-5.5 – (-4.5)	log(yr <sup>-1</sup> )	Uniform	Parameter loop
<i>stdWPrate</i>	Standard deviation of the truncated log normal distribution	0.15 – 0.4	log(yr <sup>-1</sup> )	Uniform	Parameter loop
<i>IRF</i>	Instant release fraction	0.038 – 0.156		Uniform	Parameter loop

### 4.1.3 Instant Release Fraction

Instant release fractions (IRFs) are the fractions of radionuclide inventories in a waste package assumed to be released instantly upon waste package breach. Spent nuclear fuel in the form of pelletized uranium oxide is encapsulated in assemblies of Zircaloy fuel rods inside the waste package. Fission during nuclear reactor operation results in the accumulation of an assortment of radionuclides in the fuel rods. Depending on the element, a subset of these radionuclides is volatile or leachable and can partially collect in spaces between pellets, in cracks in the pellets, and between the pellets and the rod wall. It is this highly mobile subset of radionuclides that is released instantly upon failure of the Zircaloy cladding, which is assumed to fail when the waste package is breached.

In addition to instant release, the other major mechanism of waste form release typically simulated in performance assessment calculations is kinetically limited degradation. Kinetic degradation is typically defined in terms of either mass degraded per exposed surface area per time or fraction per time (i.e., mass degraded per total mass per time) like *rateUNF*. This type of degradation releases radionuclides congruently, i.e., in proportion to radionuclide mole fractions in the evolving degraded waste form.

The radionuclide most often responsible for the largest contribution to dose to a member of the public in the accessible environment in reference case simulations is <sup>129</sup>I. For spent fuel, <sup>129</sup>I has an IRF on the order of 10%. An important question in performance assessment is the importance of instant release of radionuclides relative to kinetic degradation.

To answer this question, IRF uncertainty was added to the PA calculations for the crystalline reference case via an uncertain parameter *IRF*. The idea was to assess the effect of *IRF* on the performance metric (i.e., peak <sup>129</sup>I concentration in the aquifer) over the range of *IRF* uncertainty and extrapolate the effect to an *IRF* of 0% to estimate what the mean peak <sup>129</sup>I would have been without instantaneous release. Several sets of *IRF* data were available (e.g., Espriu-Gascon 2020 [26]; Johnson 2005 [27]; Lemmens 2017 [28]). Previously, an *IRF* of 10% for <sup>129</sup>I was assumed for the disposed spent fuel, consistent with a review by Sassani et al. 2012 [29]). That value came from Johnson et al. 2005 [27]. For this study, statistics from two PWR sample sources reported in Lemmens et al. [28], KKG-UO<sub>2</sub> and D05, were chosen. Those sample types were chosen because they have the same enrichment (4.3%) and nearly the same burnup (50.5 to 65.1 GWd/MTHM) as previously assumed for the crystalline reference case. These *IRF* measurements are shown in Table 4-2.

**Table 4-2. I-129 IRF data used to bound IRF uncertainty for spent fuel in crystalline reference case**

Sample	Form	IRF (%)
KKG-UO <sub>2</sub>	Clad fuel segment	8.9 ± 1.3
	Fragment	3.8 ± 1.9
D05	Clad fuel segment	10.8 ± 2.3
	Open segment	15.6 ± 3.8

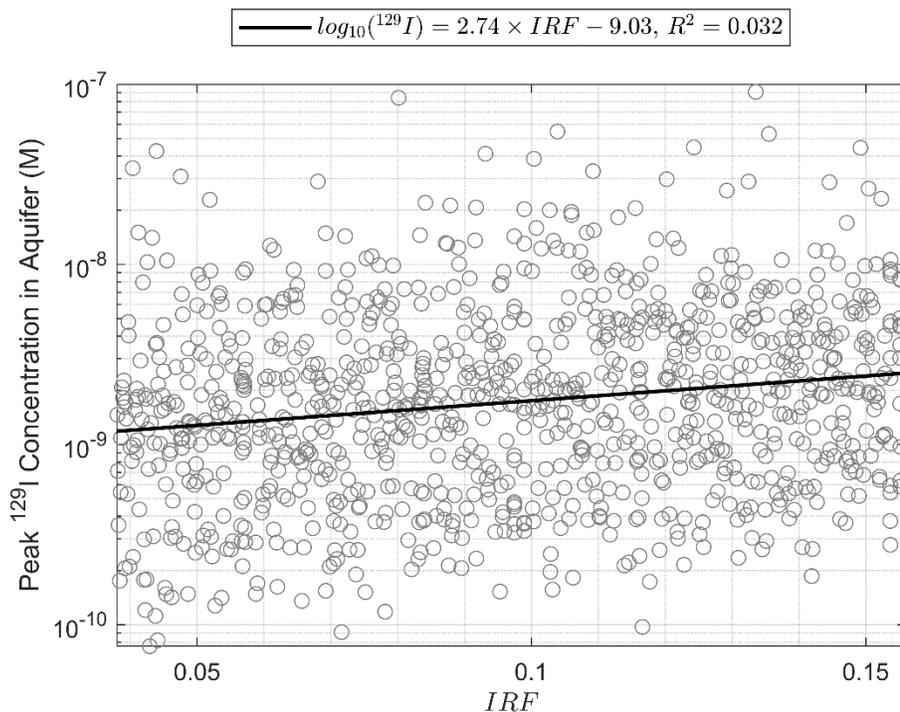
Lemmens et al. 2017, Table 3

The uncertainty provided in Table 4-2 reflects measurement error and intra-sample variation. If it is assumed that there is no single true value but rather some variation in *IRF* from one waste package to the next, as would be expected due to differences in burnup, enrichment, and linear power rating, then each waste package in a realization would have its own *IRF* value. Pursuing that approach for the reference case simulations would require careful characterization and individual rendering of each waste package in the simulation. While that approach is certainly possible, the current study simply assumes that there is no variation in *IRF* among the waste packages; rather, there is only uncertainty about what the *IRF* value is for all waste packages. Accordingly, it selects a single *IRF* value from an *IRF* uncertainty distribution for all waste packages in a realization. The uncertainty in *IRF* across realizations was assumed to be constrained by the range of mean values in Table 4-2. Specifically, the *IRF* uncertainty across realizations was uniformly distributed between 3.8 and 15.6%.

An additional reason for not applying uncertainty in the *IRF* across waste packages in a single realization is that PFLOTRAN is not factored to do so. Adding that capability is doable but not trivial. Because *IRF* is foremost a property of the species, there would need to be an *IRF* uncertainty distribution for each species and a way to assign that uncertainty across waste packages. If it becomes clear that this capability is important, its addition will be requested.

Figure 4-2 shows a scatter plot of peak <sup>129</sup>I concentration in the aquifer versus *IRF*. This figure indicates that *IRF* has an important effect on peak <sup>129</sup>I concentration. A simple linear regression was added to the plot to summarize this trend. Though the *R*<sup>2</sup> value for the regression model is low due to high variance in the data, the F-test for the regression indicates that the relationship is significant (*p* = 1.4E-08). The y-intercept of -9.03 implies that the mean peak <sup>129</sup>I would be approximately 10<sup>-9.03</sup> M (or 9.3E-10 M) in the absence of an instant release fraction and

$10^{-9.03+2.74(0.1)}$  M (or  $1.75\text{E-}9$  M) for an *IRF* value of 10%. Thus, including an *IRF* of 10% causes an average increase in the peak  $^{129}\text{I}$  concentration of 88% compared to an *IRF* of zero.



**Figure 4-2** Scatter plot of peak  $^{129}\text{I}$  concentration (M) in the aquifer versus *IRF*.

The same type of regression analysis appears to be inappropriate for *rateUNF*, the kinetic spent fuel degradation rate. Below a value of about  $5 \times 10^{-8}$ , *rateUNF* and peak  $^{129}\text{I}$  concentration appear uncorrelated (Figure 4-3). This is consistent with the idea that at very low values of *rateUNF*, *IRF* dominates the contribution to peak  $^{129}\text{I}$  concentration.

Clearly, *IRF* (like *rateUNF*) has a substantial effect on the expected value (first moment) of the peak  $^{129}\text{I}$  distribution. Some sensitivity analysis results for peak  $^{129}\text{I}$  are shown in Figure 4-4; these results are discussed in depth in Section 4.3.1 but are useful to preview here to understand the effect of including *IRF* in the analysis. Based on results of this sensitivity analysis, uncertainty in *IRF* appears to have some effect, but it is much smaller than the uncertainty contribution of *rateUNF*. Part of the reason *rateUNF* has such a large effect in the sensitivity analysis is because the analysis uses Sobol' indices, which attribute variance (second moment) to uncertain parameters. The range of uncertainty in *rateUNF* is two orders of magnitude. In comparison, the *IRF* uncertainty range is much less than one order of magnitude.

In summary, both *IRF* and *rateUNF* make substantial contributions to the peak  $^{129}\text{I}$  performance metric within the uncertainty ranges of the current crystalline reference case. Continued R&D on both types of release, instant and kinetic, could therefore potentially provide data and reveal relationships that will enhance the accuracy and confidence of future crystalline repository performance assessment.

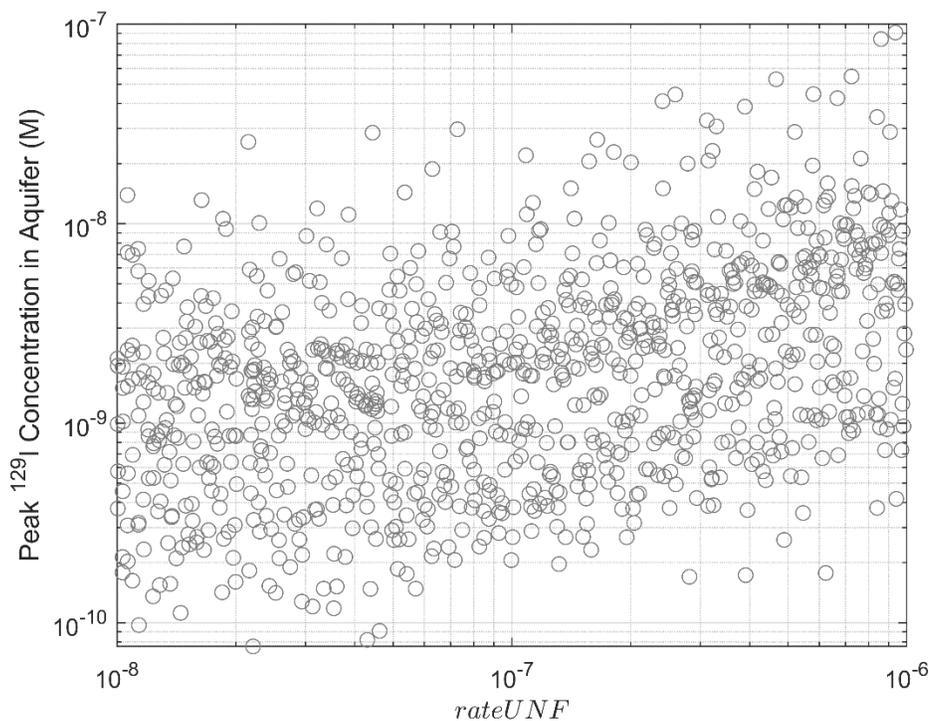


Figure 4-3 Scatterplot of the peak  $^{129}\text{I}$  concentration in the aquifer (M) versus  $\text{rateUNF}$ .

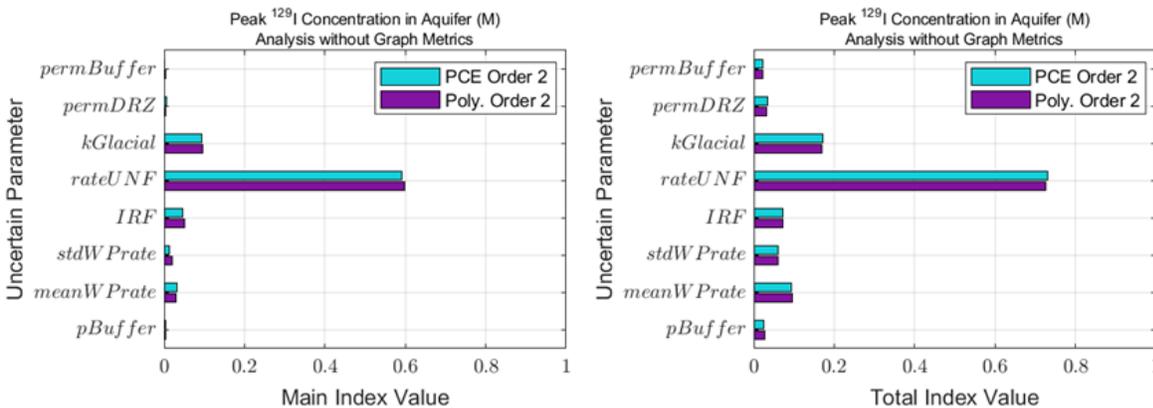


Figure 4-4 Sensitivity of peak I-129 concentration to various inputs.

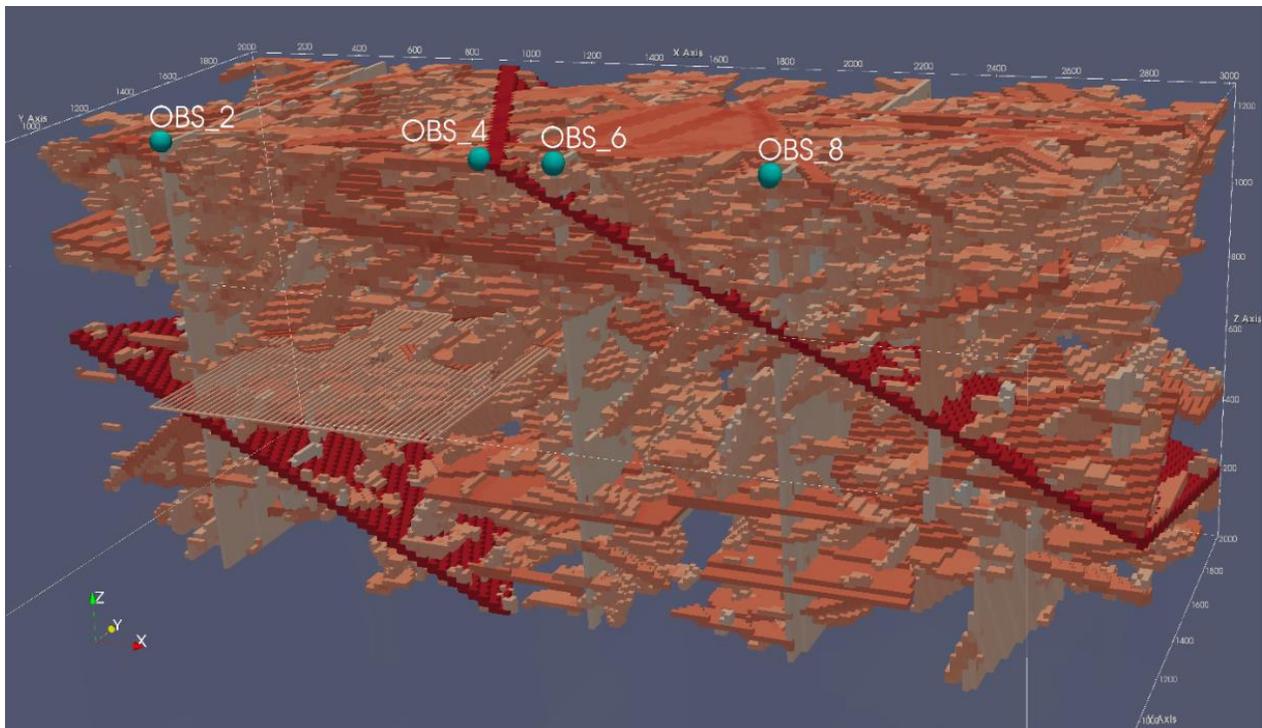
## 4.2 Model Set-up

### 4.2.1 Model Domain

The making of a cooler repository required some modification to the 2020 case. The main motivation behind the cooler repository is due to the Swiss concept [31], which allows maximum buffer temperatures in the 120 – 130°C range. If the WP temperature can be kept in this range, then it can be argued that temperature is low enough and the bentonite buffer properties can be

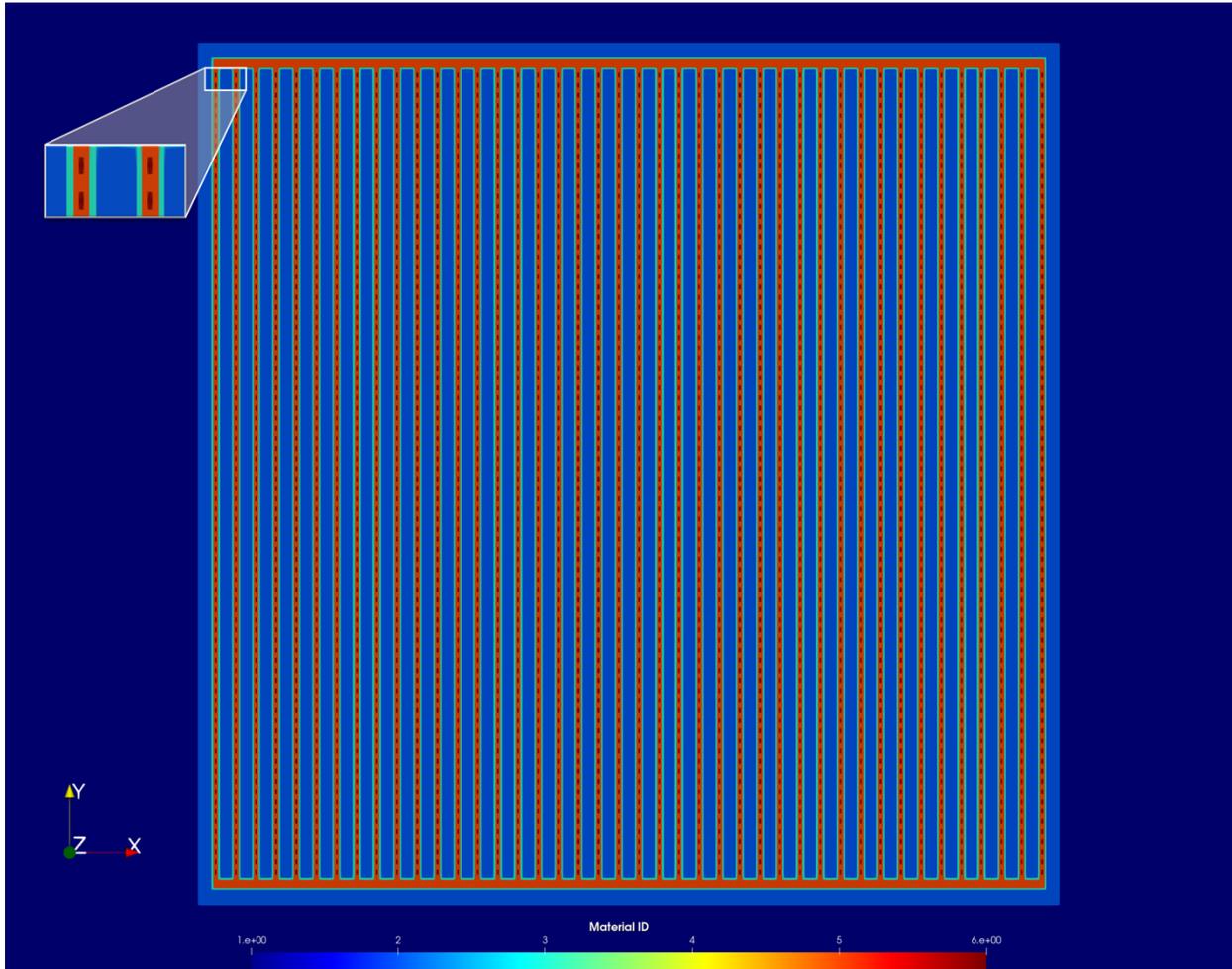
preserved. The necessary modifications needed to accomplish this cooler repository include (1) rearrangement of WPs at the repository level, and (2) recalculation of WP heat source data. Modifications were also made to the PFLOTRAN input deck this year to generate more relevant data for use in the analysis.

The current model domain which has been modified from Stein et al. [1] is 3015-m in length, 2025-m in width, and 1260-m in height, partially depicted in Figure 4-5. Overlying the host rock is a 15-m thick overburden of glacial sediments (not shown). The repository is located at a depth of 585 m. Forty-two disposal drifts now contain 40 12-PWR waste packages each (1680 12-PWR waste packages in total), which is half of the total from the FY20 case. Drifts are backfilled with bentonite buffer and are surrounded by a 1.67-m thick DRZ. Within the repository, grid cells are as small as 1.67-m on a side; elsewhere grid cells are 15-m on a side. The model domain contains 4,848,260 cells; of these, approximately 2.5 million are the smaller cells in and around the repository that allow representation of individual waste packages with surrounding buffer materials. Additional information on the grid and dimensions may be found [1] available for download at <https://pa.sandia.gov>.

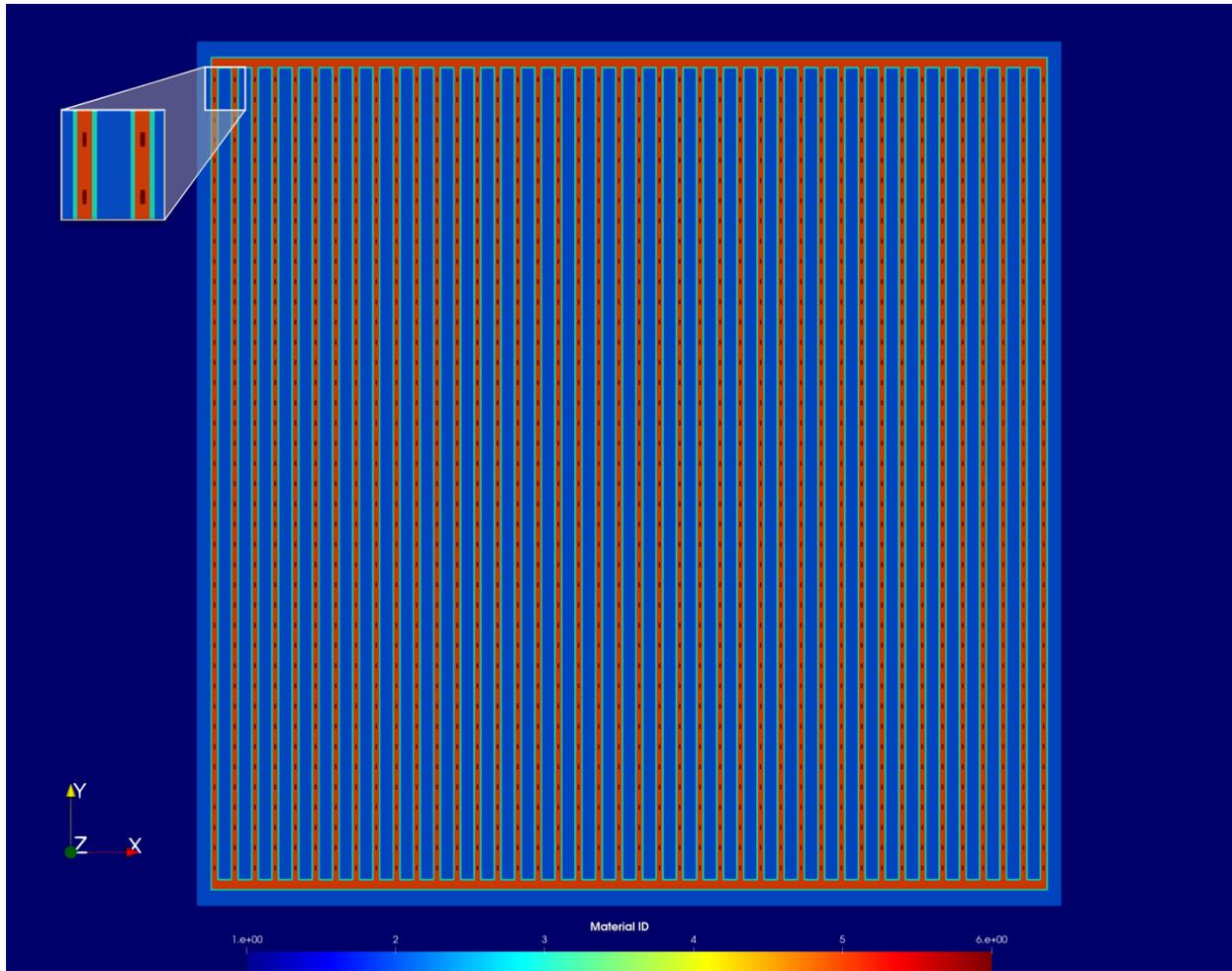


**Figure 4-5. Cut-away of DFN 1 realization mapped to porous medium grid, showing the full repository and the far half of the model domain.**

The first modification to the model domain involves a rearrangement of WPs. Previously, as seen in Figure 4-6, WPs were spaced 10 m apart within a drift. They are now spaced 20 m apart, as shown in Figure 4-7. The total number of WPs was reduced by half as a result of this wider spacing.

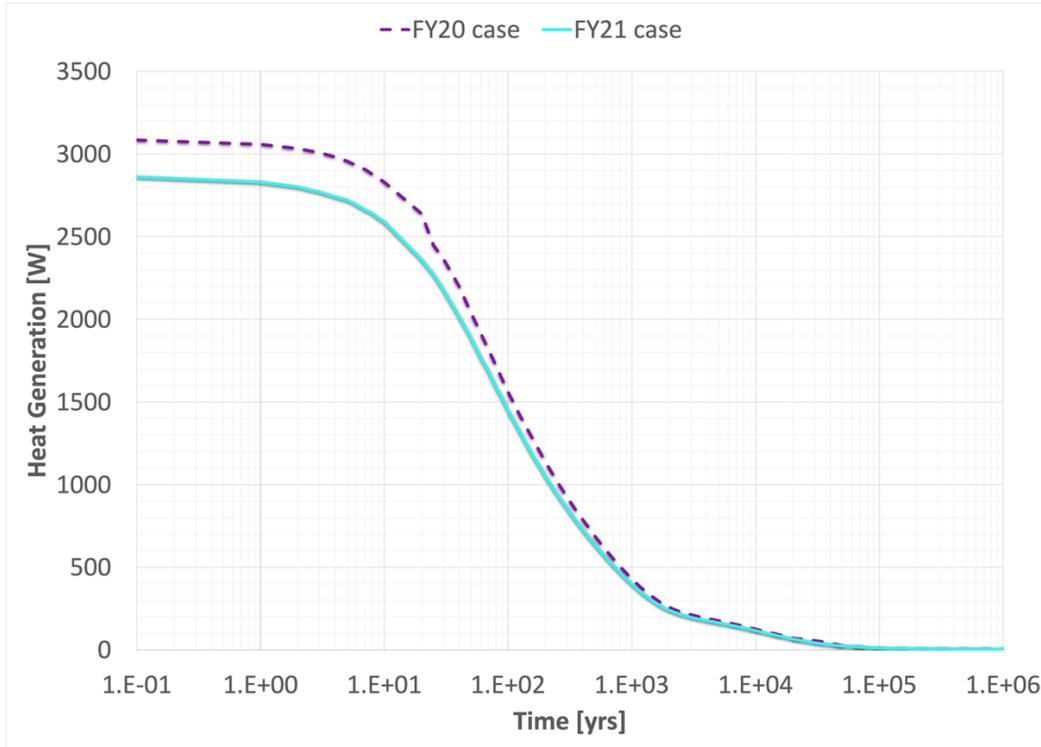


**Figure 4-6. FY20 model domain XY slice through the repository colored by material: dark blue, granite; light blue, granite at the repository level; red, buffer; maroon, WPs. Spacing between WPs is 10 m, 80 WPs/drift, 3360 total WPs. Inset zoom on the top left of the figure shows four WPs.**



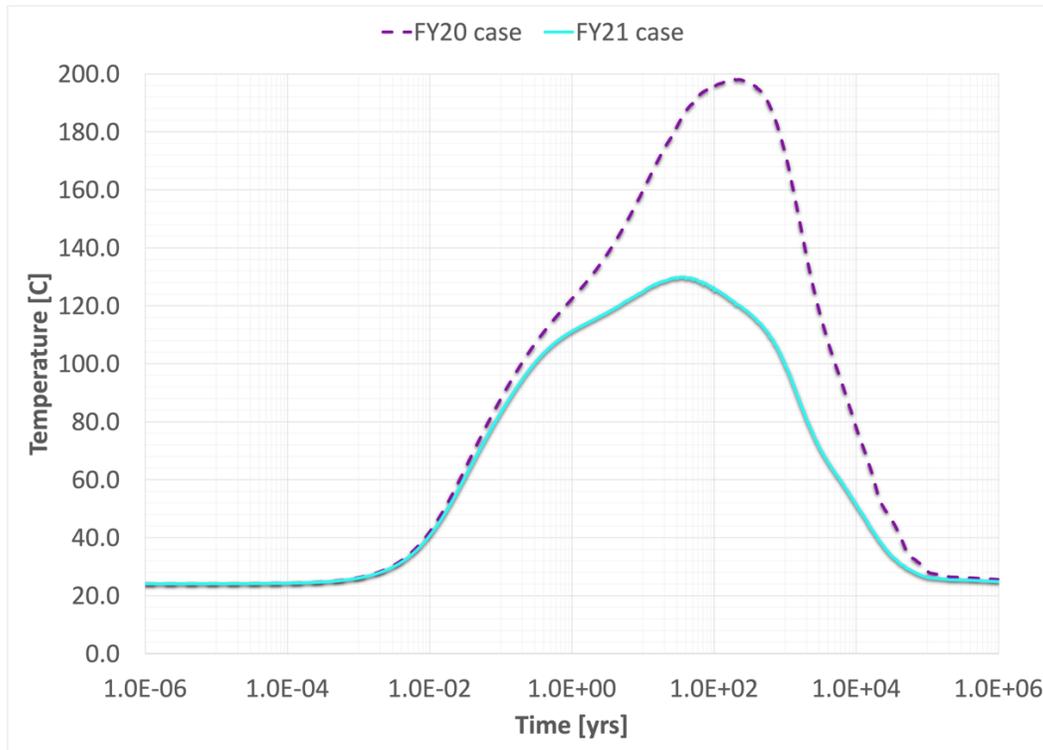
**Figure 4-7. FY21 model domain XY slice through the repository colored by material: dark blue, granite; light blue, granite at the repository level; red, buffer; maroon, WPs. Spacing between WPs is 20 m, 40 WPs/drift, 1680 total WPs. Inset zoom on the top left of the figure shows four WPs.**

The second modification required for the cooler repository involved the recalculation of the WP heat source. The FY20 heat generation per waste package was copied from early simulations that used heat output at 80 years OoR to approximate heat output at 100 years OoR. In FY21, heat generation per waste package was recalculated using correctly indexed time out of the reactor. For the FY21 simulations, 12-PWR, 100 yrs OoR, 60 GWd/MTU burn-up was recalculated from an existing 37-PWR, 100 year OoR, 60 GWd/MTU burn-up data set. Figure 4-8 shows the comparison between the heat generation per waste package used in the FY20 simulations and the recalculated data set used this year, which is several hundred watts lower initially.



**Figure 4-8. Plot of heat generation per waste package over time for the previously used 12-PWR WPs in the FY20 case (dashed purple) and the recalculated 12-PWR WPs for the FY21 case (solid turquoise).**

As a result of the modifications made to the model domain this FY, the repository was indeed much cooler. In the FY20 model, WPs reached temperatures of around 200°C. Now the maximum temperature of WPs is just over 100°C, between 120 – 130°C. A comparison of the temperature history for the centermost WP for the FY20 and cooler case for this year is shown in Figure 4-9.



**Figure 4-9. Plot comparing the temperature (°C) of the centermost WP of the repository between the FY20 (dashed purple) and FY21 case (solid turquoise).**

Modifications made to the PFLOTRAN input deck include (1) the implementation of the Leijnse soil compressibility function [24], (2) correction of previously implemented integral flux definitions, and (3) addition of several more observation points on the edges and center of the repository as discussed in Section 4.2.5.

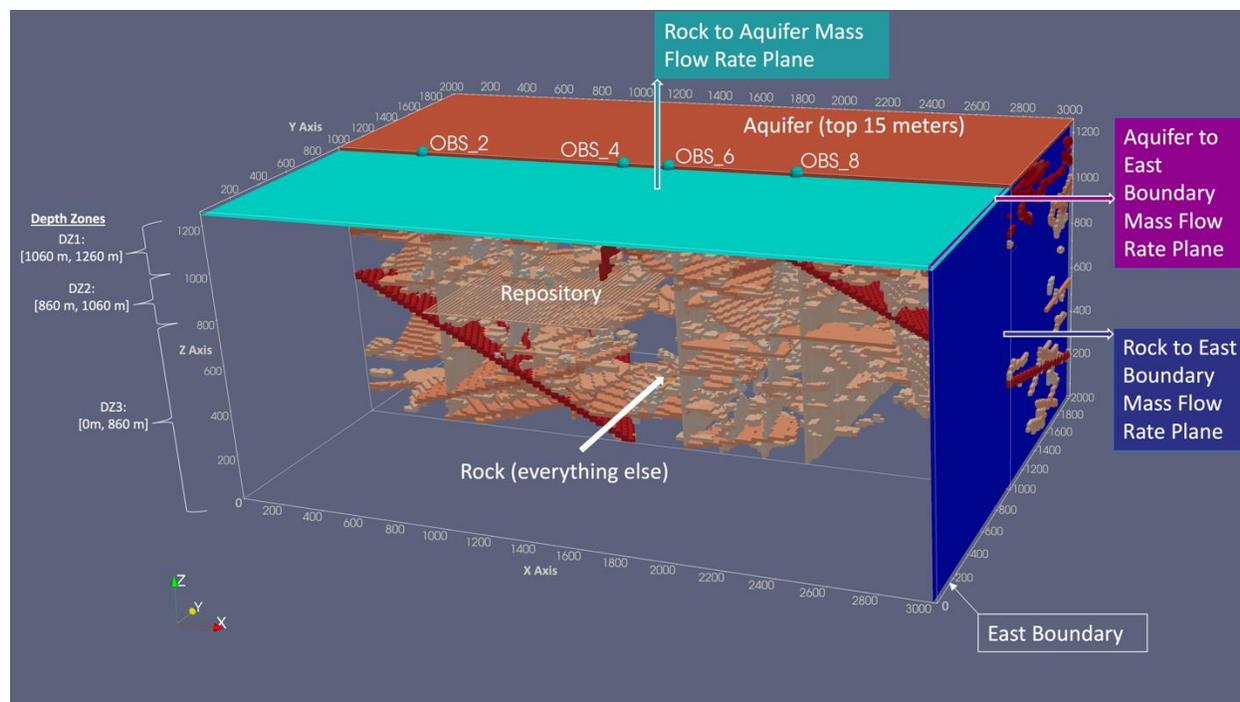
The implementation of the Leijnse soil compressibility function in the PFLOTRAN model for all material properties is important because a compressible soil absorbs some of the shock of a pressure wave. As pressure increases due to the heating of the repository, some of that increase in pressure is absorbed by the compressible soil instead of it all going into driving force for fluid flow. The soil compressibility and soil reference pressure are two values that needed to be specified for each material used in the model. The soil compressibility values used are listed in Table 4-3. The soil reference pressure is set to use the initial pressure at each grid cell for all materials.

**Table 4-3. Implemented values of soil compressibility for all material properties.**

Material ID	Material Property	Soil Compressibility [1/Pa]
1	granite	$2.2 \times 10^{-11}$ <sup>a</sup>
2	rep_granite	$2.2 \times 10^{-11}$
3	drz	$2.2 \times 10^{-10}$
4	glacial	$2.2 \times 10^{-8}$
5	buffer	$1.6 \times 10^{-8}$
6	wp	0.0

<sup>a</sup> Stevenson et al. (1996, Section 3.3)

In FY20, PFLOTRAN’s integral flux card was used in the model to set up three surfaces through which fluxes of all primary dependent variables could be calculated. In the first, three coordinates were used to define a plane between the granite (rock) and glacial (aquifer) materials within the model. The second surface attempted to define a 2D outflow region from the aquifer through the east boundary (downstream of the repository). The third attempted to define a 2D outflow region from the rock to the east boundary. The first surface was correctly defined but the second and third were not. This is because a plane defined in PFLOTRAN is assumed to extend infinitely, passing through the coordinates specified in the input file. As a result, the second and third surfaces, which were both parallel to the east boundary, resulted in the same plane through the model, yielding identical integral fluxes across them. This integral flux was thus redefined as rock and aquifer to east boundary for the FY20 study. For FY21 the second and third surfaces have been redefined as polygons in the PFLOTRAN input, along with specific coordinates defining the surfaces as originally intended. Figure 4-10 shows the updated PFLOTRAN integral flux planes within a cut-away of a single epistemic run mapped to a porous medium grid showing the full repository and the far half of the model domain. Within this figure, the rock to aquifer plane is colored in turquoise, the aquifer to east boundary plane in purple, and the rock to east boundary plane in blue. Also included in the figure are labels for the location of the repository, rock, aquifer, four observation points represented by turquoise spheres, and the location of three depth zones.



**Figure 4-10. FY21 PFLOTRAN integral flux planes and depth zones visualized within the model domain.**

#### 4.2.2 Discrete Fracture Networks

As described by Mariner et al. [4], the representation of fractured crystalline rock in the GDSA reference case is based on the well-characterized, sparsely fractured metagranite at Forsmark, Sweden [2]. At Forsmark, volumes of relatively undeformed rock (termed a “fracture domain”) are bounded by large-scale mappable features of concentrated brittle and/or ductile deformation (termed “deformation zones”). The fracture domains are sparsely fractured, and the fractures within each domain can be described in terms of “fracture sets,” which are distinguished from each other on the basis of fracture orientation. Six fracture domains are defined at Forsmark, each containing five fracture sets. Three depth zones are defined (<200 m below sea level (mbsl), 200-400 mbsl, and >400 mbsl) to account for the decrease in fracture density and fracture transmissivity with depth. We note that the repository is near the top of depth zone 3 (or 675m above the bottom of the domain as shown in Figure 4-10.) The parameters describing the DFNs for this case study are described in Chapter 3 of this report.

The crystalline host-rock reference case analyzed here [4], based on the Forsmark data set, is modeled using some fixed features and an uncertain fracture network. The case contains five fixed fracture zones and three depth intervals. Twenty random networks were generated with dfnWorks, one for each realization of the spatial uncertainty loop. Additionally, five networks were generated based on their values for average degree and shortest travel time, two of the graph quantities of interest (QoIs) that admitted high correlations with multiple performance QoIs in previous studies [22]. These five networks were chosen so that one had approximately median values for both graph

metrics, two were chosen to have a 0.05- and 0.95-quantile average degree, and two were similarly chosen for shortest travel time. The goal was to inspect whether these five DFNs would admit bounding cases for some of the performance QoIs with which they were correlated.

The deterministic, user-defined fracture zones represent large, mappable features, such as faults, and are common to all realizations. There are three subvertical fracture zones (Figure 4-5, in gray) and two fracture zones with a dip of approximately 30 degrees (in red). The discrete fracture networks, two-dimensional planes distributed in the three-dimensional model domain, are generated using dfnWorks [3], and mapped to the equivalent continuous porous medium domain using mapDFN.py, a code that approximates hydraulic fracture properties by calculating and assigning permeability and porosity on a cell-by-cell basis [1].

The fracture set properties and deterministic fracture zones employed in this study provided sufficient fracture connectedness such that each DFN realization resulted in direct fracture pathways from the repository to the top boundary of the fractured crystalline host rock. The existence of connected fracture pathways was determined by dfnWorks.

#### 4.2.3 Waste Package Corrosion Model

The waste package corrosion model implemented in PFLOTRAN (Mariner et al. [4], Section 3.2.1) calculates normalized thickness of the waste package wall at each time step as a function of a base waste package corrosion rate, a canister material constant, and temperature. Waste package breach occurs when the normalized thickness reaches zero. The normalized thickness is initialized to 1, and is reduced at each time step as a function of the effective waste package corrosion rate  $R_{eff}$  defined by

$$R_{eff} = R \cdot e^{C\left(\frac{1}{333.15} - \frac{1}{T(t,\bar{x})}\right)} \quad (4-1)$$

where  $R$  is the base corrosion rate at 60°C in units of  $\text{yr}^{-1}$ ,  $T$  is the local temperature (in Kelvin), and  $C$  is the canister material constant. This equation assumes that reaction rates are a function of temperature as described by the Arrhenius equation. It is also assumed that waste package corrosion occurs via general corrosion.  $R$  represents the normalized general corrosion rate at 60°C in units of  $1/T$  (time) (i.e., units of  $L/T$  normalized by the thickness (in units of length  $L$ ) of the waste package wall). In PFLOTRAN,  $R$  is in units of  $\text{yr}^{-1}$  and is entered as a lognormal distribution across waste packages.

PFLOTRAN assigns a base normalized general corrosion rate ( $R$ ) for each waste package by sampling on a truncated log normal distribution whose mean, standard deviation, and upper truncation limit may be treated as epistemic uncertain inputs sampled by Dakota [8]. In UA the mean of the distribution is sampled over the range  $10^{-5.5}$  to  $10^{-4.5}$  (in units of  $\text{yr}^{-1}$ ). The standard deviation is sampled over the range  $10^{0.15}$  to  $10^{0.4}$  (also in units of  $\text{yr}^{-1}$ ). These samples are drawn by sampling uniformly on the exponent, so they are uniform in log-space. The uncertainty analysis uses a constant upper truncation limit of  $10^{-3.5} \text{yr}^{-1}$ . A mean of  $10^{-5}$  for  $R$  results in a mean waste package breach time of 100,000 years if waste packages are held at a constant temperature of 60°C.

Sampling on both the mean and standard deviation implements uncertainty in the normalized general corrosion rate distribution. This high level of uncertainty is appropriate because the state

of knowledge is low and, since the distribution is sampled for each waste package, it makes sense to have a different corrosion rate distribution for each realization. Unlike the other uncertainty distributions included in the UA, the normalized general corrosion rate distribution is sampled many times within each realization. First, the mean and standard deviation are sampled once per realization to establish the corrosion rate distribution for that realization. Then this distribution is sampled independently for each waste package within that realization. Sampling of the distribution parameters implements the epistemic uncertainty in the corrosion rate and sampling each waste package independently from that distribution implements spatial aleatory uncertainty. The distributions used for the mean and standard deviation in this study are likely wider than would be used in a PA because this study focused on simulating and demonstrating the conceptual model over the large range of possible WP materials and thicknesses. These distributions would likely be narrower when applied for a specific WP material and thickness, as would be the case in a PA. In that case, there would likely be less variability in WP failure times.

#### 4.2.4 Initial and Boundary Conditions

Initial conditions specified are pressure and temperature. Nominal nonzero radionuclide concentrations are also specified as initial conditions, but this is for numerical necessity, not to represent reality. Initial pressures and temperatures throughout the model domain are calculated by applying a liquid flux of 0 m/s and an energy flux of 60 mW/m<sup>2</sup> to the base of the domain and holding temperature (10°C) and pressure (approximately atmospheric) constant at the top of the domain and allowing the simulation to run to 10<sup>6</sup> years. Pressure at the top of the domain decreases from west (left) to east (right) with a head gradient of -0.0013 (m/m). This technique results in initial conditions that represent a geothermal temperature gradient and hydrostatic pressure gradient in the vertical direction, and a horizontal pressure gradient that drives flow from west to east.

The initial concentration of <sup>129</sup>I in all cells is 10<sup>-22</sup> mol/L. A non-zero value is necessary, because PFLOTRAN transport equations are formulated in terms of the log of concentration. A concentration of 10<sup>-22</sup> mol/L is approximately 60 atoms of <sup>129</sup>I per liter of water.

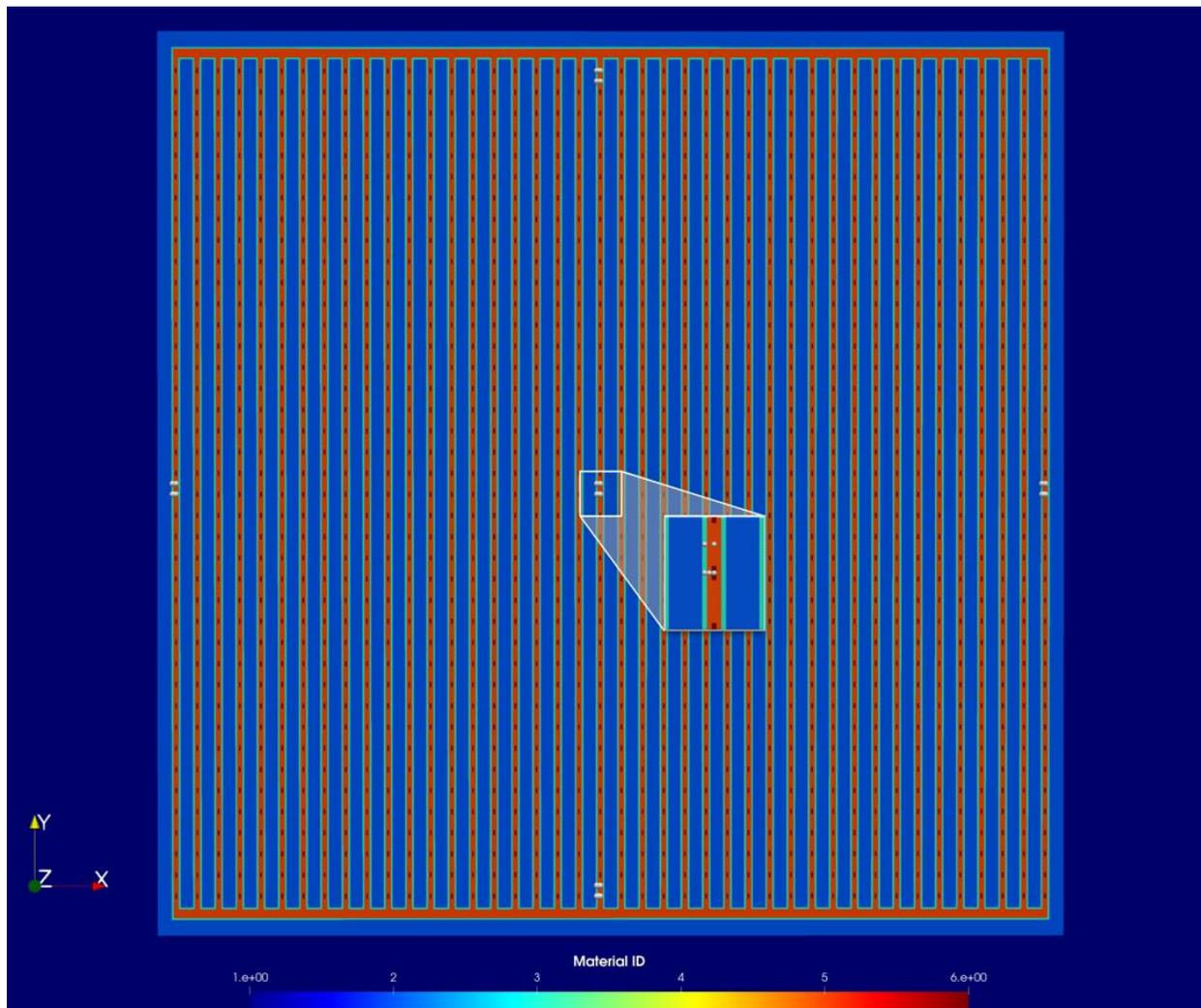
At all six faces of the model domain, pressures and temperatures are held constant at initial values. Concentration of <sup>129</sup>I is held at the initial concentration at inflow boundaries. At outflow boundaries, the concentration gradient is set to zero.

#### 4.2.5 Observation Points

Previous analysis of the crystalline reference case analyzed uncertainty in <sup>129</sup>I concentrations at specific observation points [14]. Rather than focusing on each spatial location individually as in that analysis, this UA uses the maximum <sup>129</sup>I concentration calculated by taking an element-wise maximum over all the cells in a region such as the aquifer. This QoI is assumed to act as a surrogate for repository and geosphere performance in this analysis as if there is a groundwater protection standard and a maximum concentration limit for the aquifer. The performance metric for a PA is established legally and, though laws change, future requirements will likely limit dose to a reasonably maximally exposed individual. The current crystalline reference case model does not yet include dose modeling, so until the model reaches that level of maturity, the peak <sup>129</sup>I concentration is treated as a performance measure to guide understanding of the model behavior

and facilitate development of analysis tools that will later be applied to a dose-based performance measure.

This year, twenty-seven observation points were added to the repository that may be useful for future studies. These newly added observation points consist of a trio of points in regions of interest such as the buffer, DRZ, and WPs within the repository and sets of two points between neighboring WPs along a given drift. Out of the total number of points, twenty are located along the four edges of the repository, while seven are located in its interior. Two of the interior observation points are located in the buffer on either end of the centermost WP in the y direction. Figure 4-11 shows an XY slice of the repository with all newly added observation points represented by white dots.



**Figure 4-11. XY slice through the repository colored by material ID shows all twenty-seven newly added observation points represented by white dots**

#### 4.2.6 Timestep Size

During previous analysis of the crystalline reference case [14], oscillations in  $^{129}\text{I}$  concentrations which could impact estimates of  $^{129}\text{I}$  breakthrough times were found at observation points. An analysis was conducted to identify the cause of the oscillations and determine how the effect could be minimized in future studies. In [14], we document a study showing the tradeoff between maximum timestep size and computational time: Oscillations in the  $^{129}\text{I}$  concentrations become less frequent with smaller timesteps but this results in an increased computational time. From the study, we determined that a maximum timestep of 5,000 years was acceptable for our quantities of interest. For the PFLOTRAN runs presented in this report, a maximum timestep size of 5,000 years was also used.

#### 4.2.7 Quantities of Interest (QoIs)

In FY20, several quantities of interest were generated for the crystalline reference case. These are described in [22] and we are still tracking them. We continue to develop additional QoIs that may be better indicators of repository performance, repository leakiness, retention of conservative tracers in the repository, relative flow rates and directions through the rock, and flow connections between the repository and aquifer. Leakiness describes water flux out of the repository. The QoIs used for the sensitivity analysis of the updated crystalline reference case are listed in Table 4-4.

The main performance metric capability is the peak  $^{129}\text{I}$  concentration, as in previous studies. First, the entire aquifer is monitored at each time step for  $^{129}\text{I}$  concentration and the maximum  $^{129}\text{I}$  concentration is recorded. Then, the maximum value over all time points is determined. This is called the peak  $^{129}\text{I}$  concentration in the aquifer over the course of the simulation. The location of the peak  $^{129}\text{I}$  concentration (in terms of x and y coordinates in the aquifer layer) is also recorded. The x- and y-location of peak  $^{129}\text{I}$  concentration indicate the spatial variability of the system. This metric enables analysis of the model and development of PA tools and techniques; as the model matures to include dose to a member of the public, the peak  $^{129}\text{I}$  concentration will be replaced by a dose-based performance metric.

The DFN-related characteristics that are potentially relevant to performance include the mean travel time (MTT) of a conservative tracer from the repository to the aquifer, the median residence time (MdRT) of an initial conservative tracer within the repository, fraction of a conservative tracer remaining in the repository at certain times, fractional mass fluxes of that tracer at certain times, and rock boundary water mass flow rates.

A more detailed description of MTT may be found in Mariner et al. (2020) [15] and Swiler et al. (2020) [22]. The MTT measures the mean travel time of a conservative (non-decaying) tracer from the repository to an observation point beyond the repository. For the equations defining MTT, see equations 32-35 in [15]. Mean travel time can be directly measured using tracers. Identical concentrations of two conservative tracers are artificially and continuously injected at a constant rate at the starting point. The only difference between the two tracers is that one decays or ingrows exponentially over time from injection. Because the movement of these tracers within the domain is identical, the difference in concentration at a distant location is solely due to the mean time since tracer injection. This measure of time is considered the mean travel time (MTT) of a conservative tracer. Note for the QoI studied in this report (especially in Chapter 2), we focus on the MTT evaluated when the peak  $^{129}\text{I}$  concentration occurs.

The median residence time of a conservative tracer initially present in the repository (MdRT) is the time at which half of the original tracer amount remains in the repository. We note a difference between MdRT and MTT: for MdRT, there is one initial concentration (“spike”) of the conservative tracer at the beginning of the simulation and no more of it is injected into the region, then MdRT is obtained by tracking the tracer mass remaining in the repository until half remains. In contrast, for MTT, two tracers are continuously injected and the differences in their concentrations at locations beyond the injection region can be used to calculate the MTT of injected tracer. Note additionally that for FY20 the mean residence time (MRT) was tracked, which is also computed by tracking the mass of the “spike,” but the estimated MRT is valid only if the mass of the tracer has decreased nearly to zero in the repository. It was found that the fraction of mass of the “spike” remaining in the repository by the end of the simulation was still around 0.25, so the MdRT was deemed a more useful measure of repository retention for this work.

Additional metrics relating to the tracer concentrations include the fraction of the tracer remaining in the repository at 3000 years or at 1 million years. These two times are of interest since 3000 years is a time point during the thermal pulse from the repository and 1 million years is the final performance time. The final tracer-related metric is the fractional mass flux of the tracer from repository. This is also assessed at 3000 and 1 million years. It is an instantaneous fractional loss rate of the tracer remaining in the repository at a particular point in time, in units of (1/yr). MdRT, the fraction of tracer remaining at certain times, and the fractional mass flux of the tracer from the repository are all indicators of repository retention.

The final group of metrics involves ratios of water fluxes. There are two ratios which we study in this sensitivity analysis. The first is the ratio of the aquifer-to-east-boundary flux to the rock-to-east-boundary flux. This is the ratio of two water fluxes: the flux from the aquifer to the east boundary normalized by the flux from the rock to the east boundary. It indicates the multiplication factor on aquifer dominance of the East side effluent. We evaluate this flux ratio at 3000 years during the thermal pulse and also at one million years, where the conditions are nearly undisturbed.

The other flux ratio is the ratio of the rock-to-aquifer flux to the rock-to-the-east-boundary flux. This indicates a multiplication factor on upward vs. horizontal flow. It is also evaluated at 3000 years during the thermal pulse as well as at 1M years in nearly undisturbed conditions.

Fluxes and flux ratios were also tracked over time so time-dependent sensitivity analysis could identify the uncertainties driving changes in flow over time. This also allows for a more continuous comparison between the pressure pulse and thermal pulse phases, rather than comparing only two points in time. Time-dependent sensitivity analyses for select fluxes are presented in Section 4.3.2.

**Table 4-4. FY2021 QoIs for crystalline reference case sensitivity analysis**

#	Quantities of Interest (QoIs)	Description
1	Peak <sup>129</sup> I concentration in aquifer in mol/L (M)	Scalar with (X,Y,Z) location of peak. This is an indicator of repository performance.
2	MTT for Peak in years	Scalar, MTT as calculated from the tracer concentration ratio, evaluated at the time of Peak <sup>129</sup> I concentration
3	Median Residence Time MdRT of Spike in Repository	Median residence of a tracer in the repository. This is the time when half the tracer remains in the repository. It is an indicator of repository retention
4	Fraction of Spike in repository at 3000 yrs	The fraction of a tracer remaining in repository at 3000 years. It is an indicator of repository retention.
5	Fraction of Spike in repository at 1 million years	The fraction of a tracer remaining in repository at 1 million years. It is an indicator of repository retention.
6	Fractional Mass Flux of Tracer from Repository at 3000 yr (1/yr)	The instantaneous fractional loss rate of tracer remaining in repository at 3000 years. It is an indicator of repository retention.
7	Fractional Mass Flux of Tracer from Repository at 1 million years (1/yr)	The instantaneous fractional loss rate of tracer remaining in repository at one million years. It is an indicator of repository retention.
8	Aquifer to East / Rock to East at 3000 years	This is the ratio of two water fluxes: the flux from the aquifer to the east boundary normalized by the flux from the rock to the east boundary. It indicates the multiplication factor on aquifer dominance of East side effluent at 3000yr during thermal pulse.
9	Aquifer to East / Rock to East at 1 million years	This is the ratio of two water fluxes: the flux from the aquifer to the east boundary normalized by the flux from the rock to the east boundary. It indicates the multiplication factor on aquifer dominance of East side effluent at 1 Myr near undisturbed conditions.
10	Rock to Aquifer / Rock to East at 3000 year	This is the ratio of two water fluxes: the rock to the aquifer vs. the rock to the east boundary at 3000 years. It indicates a multiplication factor on upward vs horizontal flow at 3000 years during the thermal pulse.
11	Rock to Aquifer / Rock to East at 1 million years	This is the ratio of two water fluxes: the rock to the aquifer vs. the rock to the east boundary at 1 million years. It indicates a multiplication factor on upward vs horizontal flow at 1 M years in near undisturbed conditions.
12	Rock to aquifer/ Rock to East Boundary	This is the vector of the ratio of two water fluxes: the rock to the aquifer vs. the rock to the east boundary. The vector is used in the time-varying sensitivity analysis.
12	x Location of Peak I-129 (m)	The ensemble indicates spatial variation in the x location of peak I-129.
13	y Location of Peak I-129 (m)	The ensemble indicates spatial variation in the y location of peak I-129
14	Rock to Aquifer Flux	Water flux from rock to aquifer (kg/yr). This is a vector of values calculated at different time points.
15	Rock to east flux normalized by the flux at 1 Ma	Water flux from rock to the east boundary, normalized by the flux at 1 Ma. This is a vector of values calculated at different time points (note it has a value of 1 at 1Ma).
16	Aquifer to east flux normalized by the flux at 1 Ma	Water flux from aquifer to the east boundary, normalized by the flux at 1 Ma. This is a vector of values calculated at different time points (note it has a value of 1 at 1Ma).

## 4.3 Results

This uncertainty and sensitivity analysis focused on exploration of the FY21 quantities of interest and their utility for increasing understanding of the repository and its performance. Sensitivity analysis results, time series plots, and scatter plots are presented for the QoIs that proved illustrative. As in previous analyses of the crystalline reference case, sensitivity analysis was performed using various surrogate models to estimate main and total effect Sobol' indices using Dakota [8, 14, 16]. Results are presented for second-order polynomial chaos expansion (PCE) surrogate models as well as second-order polynomial surrogate models. Analyses were also performed with multivariate adaptive regression splines and Gaussian process models, but these surrogate models performed poorly for some QoIs so the final analysis includes results from only the PCE and polynomial surrogate models. Though it would be valuable to study why some surrogates performed better than others and how to improve surrogate performance, the new QoIs were really the focus of this study; choosing the best performing surrogates allows the analysis to focus on these QoIs.

Two versions of each sensitivity analysis are also provided for comparison to each other. One version includes both parameter uncertainty and the spatial uncertainty characterized using graph metrics in the surrogate construction; the second version excludes the spatial uncertainties and does not account for them in the surrogate construction. Both versions of the analysis attempt to explain the same variance; the difference is whether the surrogate is allowed to attribute any of the variance to the spatial heterogeneity. It is illustrative to perform both versions of the sensitivity analysis because the sampling structure does not enable a complete separation of the spatial and epistemic uncertainties. Comparison of the two analyses highlights the influence of spatial uncertainties and the importance of developing summary statistics of the spatial uncertainty that can be applied in variance decomposition based global sensitivity analysis. However, interpretation of these analyses and their comparison is also complicated by the spatial uncertainties. Because of the dual-loop sampling structure, the spatial uncertainties are repeated for 25 epistemic realizations each, whereas the epistemic samples are all unique. This may over-emphasize the spatial uncertainties and presents a challenge for the surrogate models. Scatter plots and time series plots are included with the sensitivity analyses to support interpretation.

Section 4.3.1 presents sensitivity analysis results for QoIs that involve some aggregation over time (such as a maximum over time) or are defined at a specific point in time. Section 4.3.2 presents sensitivity analysis results for QoIs that are tracked over time with corresponding sensitivity analysis results at each time step.

### 4.3.1 Sensitivity Results for Scalar QoIs

This section presents analysis results for QoIs that are either aggregated over time or correspond to a specific point in time and so are scalars. Results are provided using four plots, which contain the main and total Sobol' index values for the sensitivity analyses that were performed with and without the graph metrics. Recall that the main index value for a parameter indicates the proportion of the variance in the QoI accounted for by that parameter alone in the surrogate model. The total index value for a parameter indicates the proportion of the variance in the QoI accounted for by that parameter and any interactions it has with other parameters in the surrogate model. In this analysis, both the PCE surrogate and the polynomial surrogate are second order, so the total index for a parameter includes its main effect and effects from two-way

interactions. When the total index is greater than the main index, this indicates significant interactions.

The sensitivity analysis results for peak  $^{129}\text{I}$  concentration are plotted in Figure 4-12. The top subplots show the results when variance cannot be attributed to the graph metrics representing the spatial uncertainty and the bottom subplots show the results when variance can be attributed to these graph metrics. In the top plots, which do not account for spatial heterogeneity in the surrogate model, the spent fuel dissolution rate (*rateUNF*) and glacial till permeability (*kGlacial*) account for most of the variance. The total index for *rateUNF* is around 0.1 higher than its main index and the total index for *kGlacial* is also higher than the main index, which suggests interaction.

The analysis in the bottom subplots of Figure 4-12 includes three graph metrics to represent spatial heterogeneity. With these metrics included, the indices for *kGlacial* and *rateUNF* are substantially reduced. This also contrasts with the FY20 analysis because that analysis included only the relative shortest travel time. With the addition of two more graph metrics, more of the variance in peak  $^{129}\text{I}$  is attributed to the spatial loop. These results, however, should be interpreted together and through comparison with scatter plots since the repeated values in the spatial loop may bias the surrogate model to over-attribute variance to the graph metrics. Note that in Figure 4-13, the trend between peak  $^{129}\text{I}$  and *rateUNF* is distinct and there are apparent minor trends with respect to *kGlacial* and *IRF*. Additionally, the scatter plots show trends between the graph metrics and peak  $^{129}\text{I}$  that correspond to the high main indices for the graph metrics. The importance of the graph metrics makes sense because the number of intersections influences  $^{129}\text{I}$  transport out of the repository and the average degree summarizes the densities of transport paths through the host rock. Interpreting both analyses and the scatter plots together, it is clear that uncertainties in both loops are important.

These results contrast with the FY20 analysis of the crystalline reference case. In that analysis, the only parameter with a clear influence was *kGlacial*. The improved clarity in sensitivity analysis results in the FY21 analysis may be due to the change in how the parameter uncertainty loop is sampled. Previously, there were only 40 unique samples in the parameter uncertainty loop, which were repeated for each spatial sample. Now the parameter uncertainty loop is re-sampled for each iteration of the spatial loop, giving denser coverage of the parameter input space and thus, the sample size is effectively larger. Additionally, the temperatures in the repository are lower (see Section 4.2.1), which causes small changes in flows and degradation rates, and a new set of DFNs was used. These changes could also affect simulation results and change the relative influence of uncertain parameters.

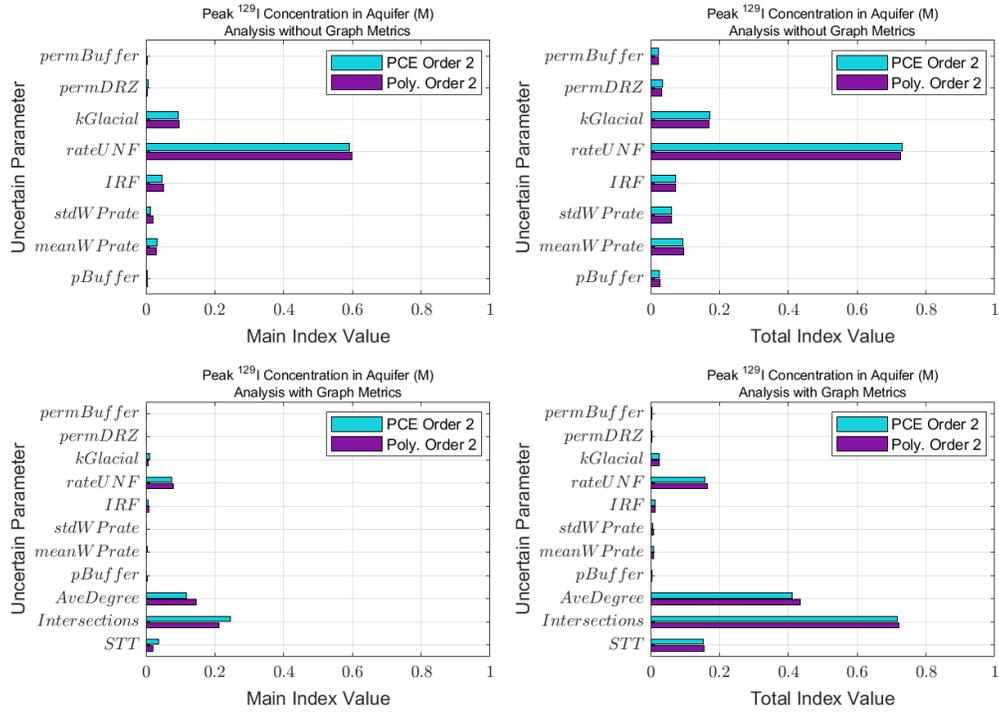


Figure 4-12. Global sensitivity analysis results for peak <sup>129</sup>I.

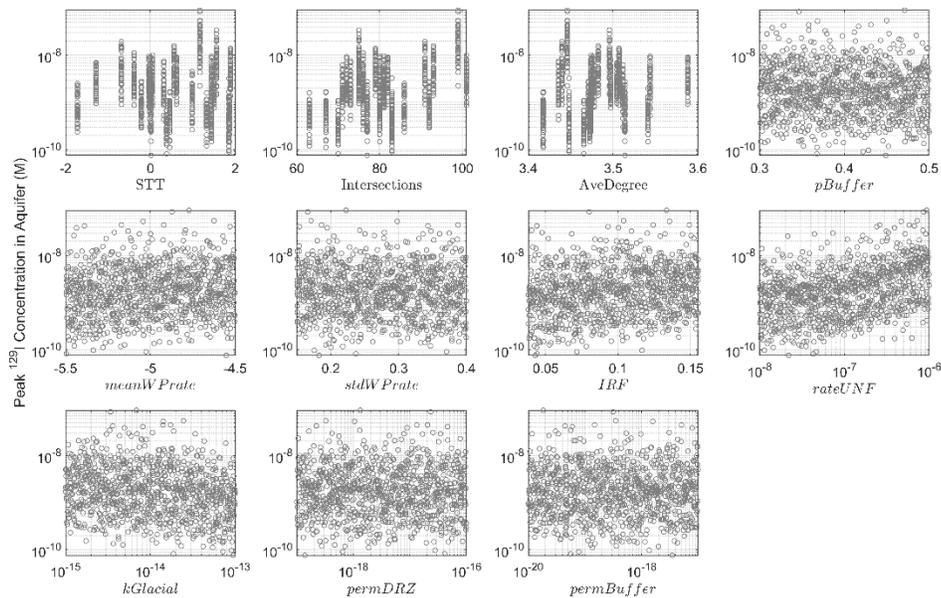


Figure 4-13. Scatterplots of peak <sup>129</sup>I versus uncertain parameters and graph metrics.

Sensitivity analysis results for the mean travel time at peak  $^{129}\text{I}$  are shown in Figure 4-14. Similar to the peak  $^{129}\text{I}$  results, inclusion of the graph metrics drastically reduces the indices for the parameter uncertainties. Prior to inclusion of the graph metrics, *rateUNF* and *meanWPrate* appear important with significant interactions. With inclusion of the graph metrics, they both appear completely unimportant. The scatter plots in Figure 4-15 show clear trends with the graph metrics and with *meanWPrate*, so these results appear to inflate the importance of graph metrics over parameter uncertainties. The trend with *rateUNF* in the scatter plot is less distinctive but still apparent.

Sensitivity analysis results for the median residence time of the tracer in repository are plotted in Figure 4-16. This is one of the new QoIs for the crystalline reference case and is an indicator of repository leakiness, so it makes sense that buffer porosity (*pBuffer*) would dominate the sensitivity results when graph metrics are not included. This is confirmed in the scatter plots in Figure 4-17, where the trend between buffer porosity and the median residence time is distinctly positive. The initial tracer is established as a uniform aqueous concentration throughout the repository at time zero. The repository domain is defined as a tightly fitting rectangular box around the repository DRZ that includes the DRZ and rock in the box. By far the largest fraction of the water in this domain is in the buffer cells because buffer has much higher porosity than the DRZ and rock in the domain. Thus, a 10% increase in buffer porosity will nearly cause a 10% increase in the total mass of initial tracer in the repository domain. This is important, because although transport in the DRZ and rock may be dominated by advection, transport from the buffer to the DRZ is dominated by diffusion, which is a very slow process by comparison. Therefore, an increase in the buffer porosity of 10% will significantly increase the proportion of tracer whose migration in the repository domain is diffusively limited, thereby increasing the mean and median residence times of the tracer in the repository.

For median residence time, including the graph metrics still reduces the index values for *pBuffer*, but not as dramatically as for some of the other QoIs, and *pBuffer* still dominates. Of the graph metrics, the number of intersections with the repository has the largest main index. Though the main index is somewhat low (just below 0.2), the total index is not much higher, suggesting the importance of the number of intersections is predominantly a main effect, not an interaction effect. There is a weak trend in the scatterplot in Figure 4-17 suggesting more intersections with the repository are associated with lower median residence times. This makes sense because the additional intersections with the repository represent additional cracks, which should increase outward transport.

Sensitivity analysis results for the fractional mass flux from the repository at 3000 years are plotted in Figure 4-18. These results are very similar to the median residence times. The primary difference is just that *pBuffer* maintains slightly higher indices for this QoI when the graph metrics are included in the analysis. The reason these results are so similar is because the trends are essentially the opposite of those for the median residence time. Scatter plots for the fractional mass flux from the repository at 3000 years are plotted in Figure 4-19. The mass flux (kg/yr) is essentially constant. However, in the fractional mass flux calculation, larger porosities (*pBuffer*) lead to larger tracer masses. So the calculation of this fractional QoI has an increasing denominator with respect to *pBuffer* and a constant numerator, resulting in the fractional mass flux decreasing with respect to *pBuffer* due to the effect of *pBuffer* on the tracer mass. The number of intersections with the repository is the most significant of the graph metrics and demonstrates a positive trend.

Results for the ratio between the aquifer to east boundary flux and the rock to east boundary flux are plotted in Figure 4-20. There is little difference between the analysis that included graph metrics and the analysis that excluded them for this quantity of interest. This makes sense because the trend with *kGlacial* in Figure 4-21 is strong and obvious trends are not apparent for any of the graph metrics. However, note that the trend with *kGlacial* appears in nearly parallel lines. These lines are different spatial realizations. So, while the general trend is strongly determined by *kGlacial*, that trend varies up or down due to spatial heterogeneity. None of the graph metrics have very significant sensitivity index values, so they do not appear to describe the spatial heterogeneity that drives this variation. This example may motivate exploration of additional graph metrics in the future; the current graph metrics clearly improve understanding but are not sufficient to fully characterize the effects of spatial heterogeneity.

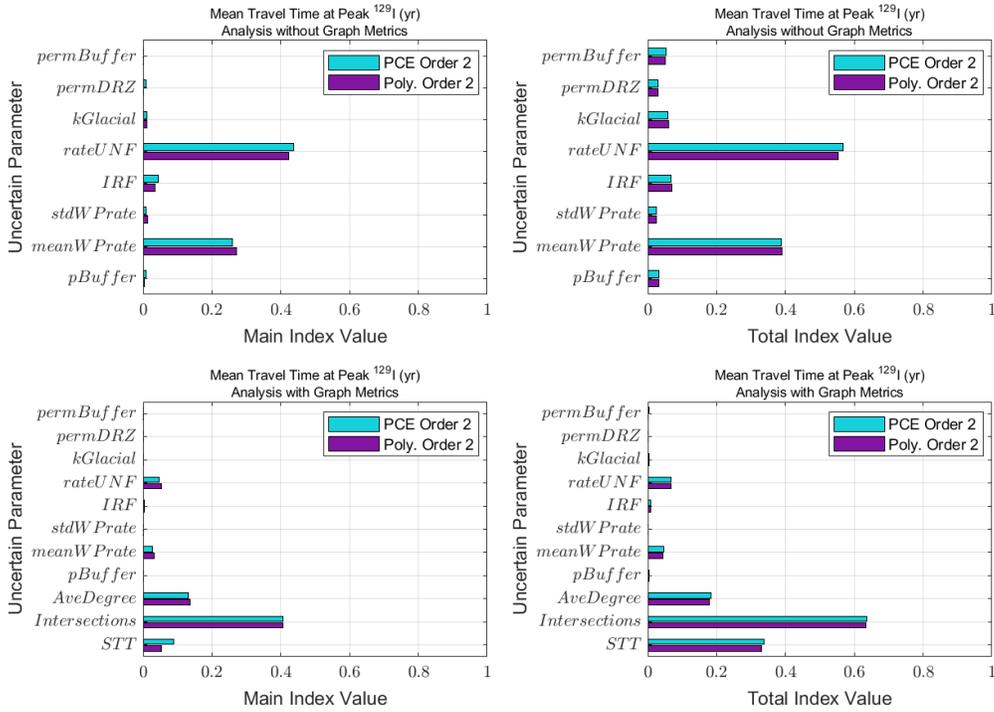


Figure 4-14. Global sensitivity analysis results for the mean travel time at peak  $^{129}\text{I}$ .

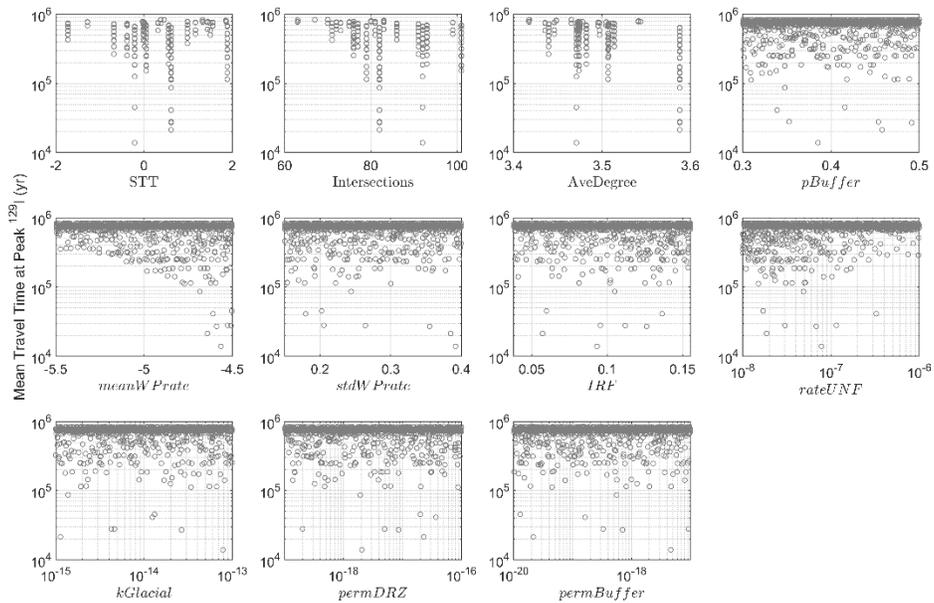
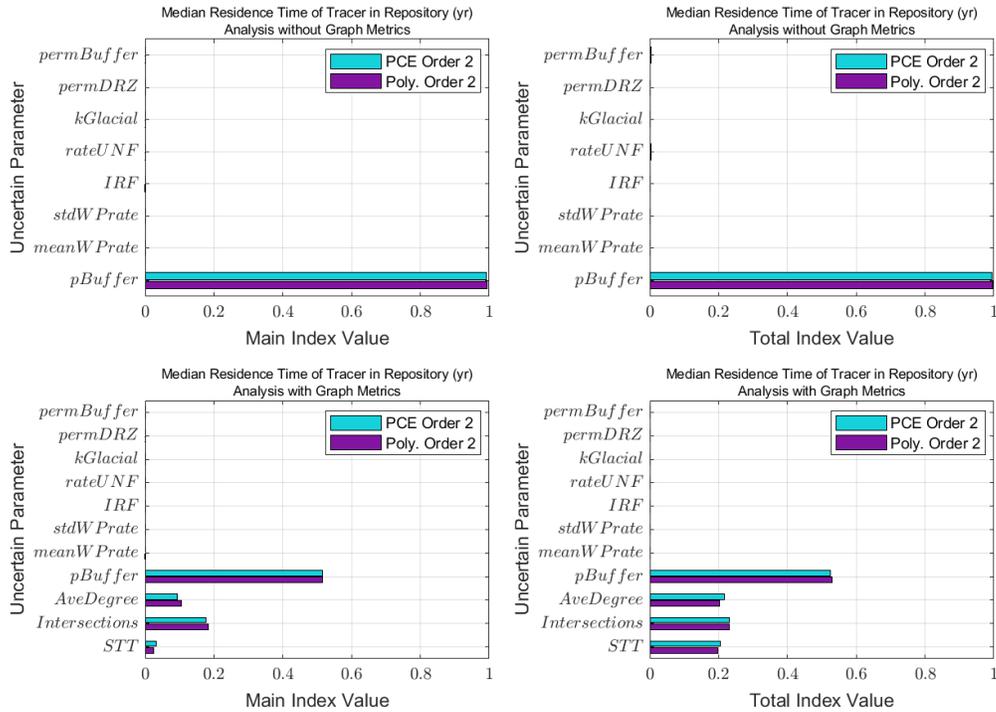
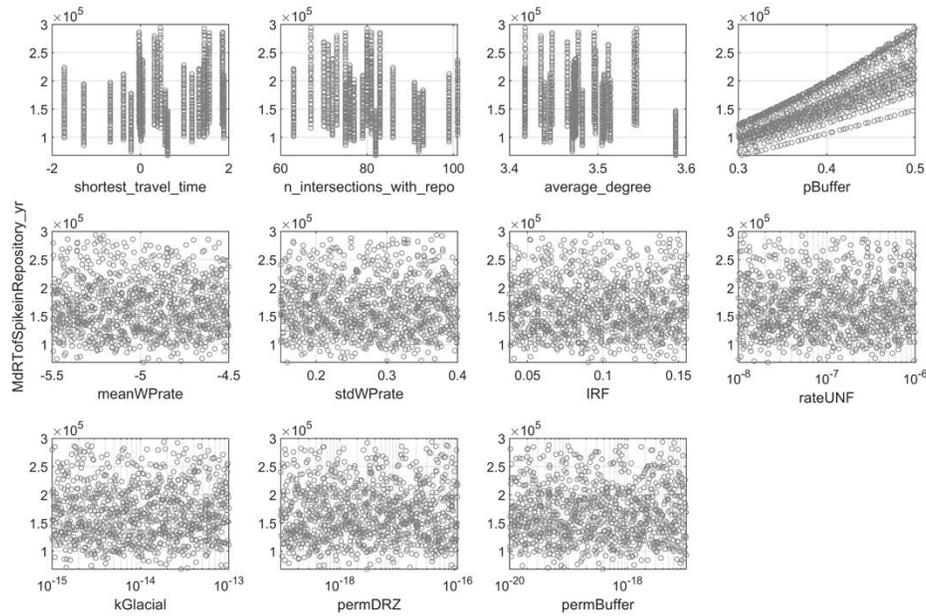


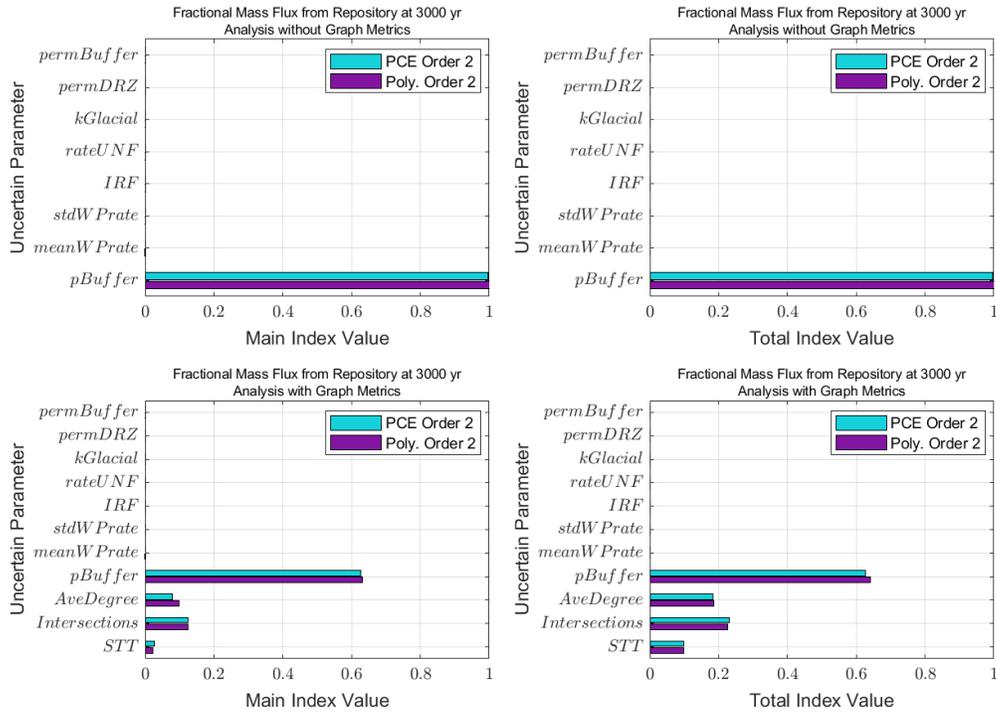
Figure 4-15. Scatterplots of mean travel time at peak  $^{129}\text{I}$  versus uncertain parameters and graph metrics.



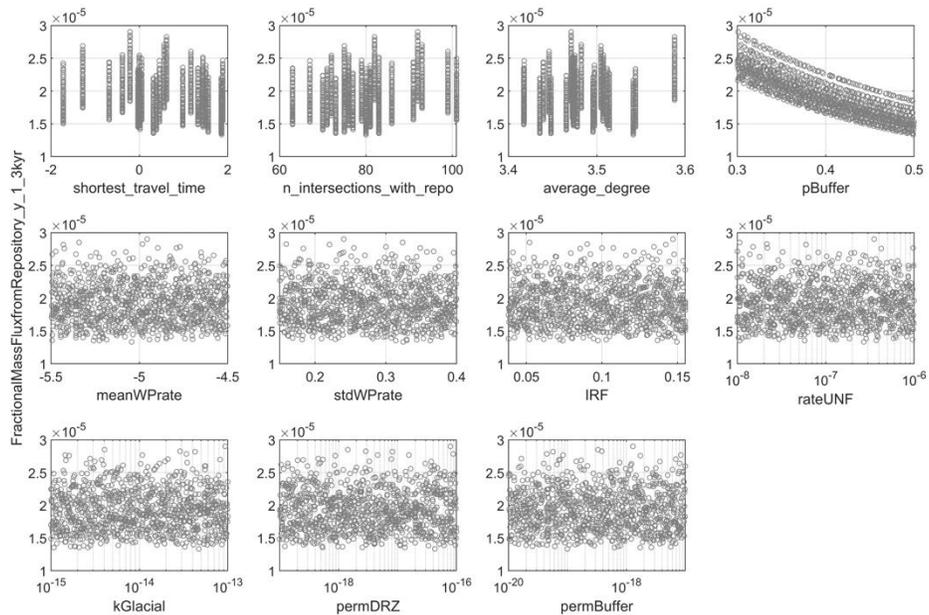
**Figure 4-16. Global sensitivity analysis results for the median residence time of tracer in repository.**



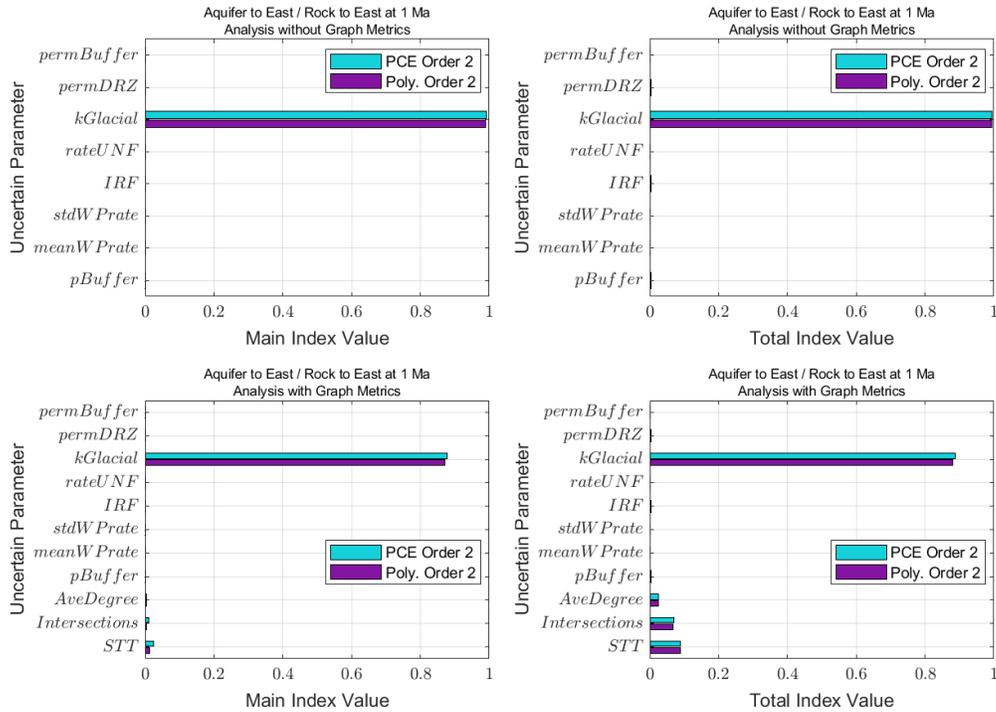
**Figure 4-17. Scatterplots of the median residence time of tracer in repository versus uncertain parameters and graph metrics.**



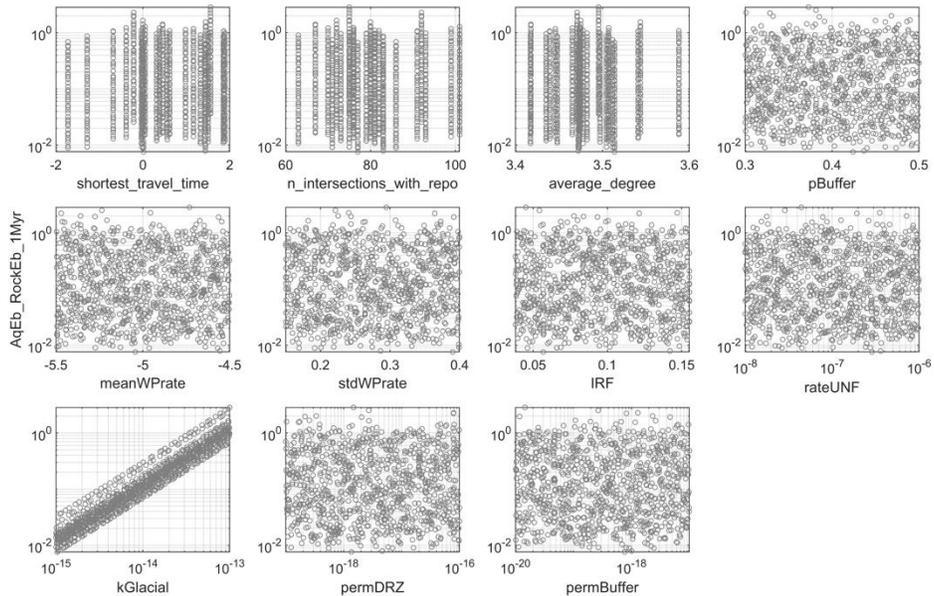
**Figure 4-18. Global sensitivity analysis results for the fractional mass flux from the repository at 3000 years**



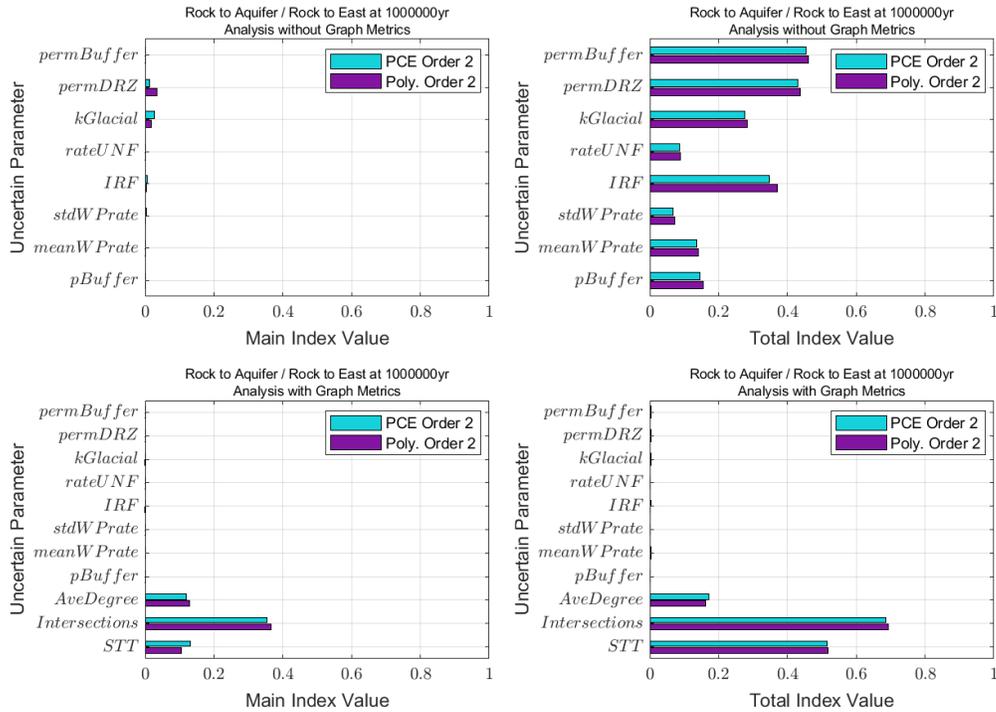
**Figure 4-19. Scatterplots of the fractional mass flux from the repository at 3000 years versus uncertain parameters and graph metrics.**



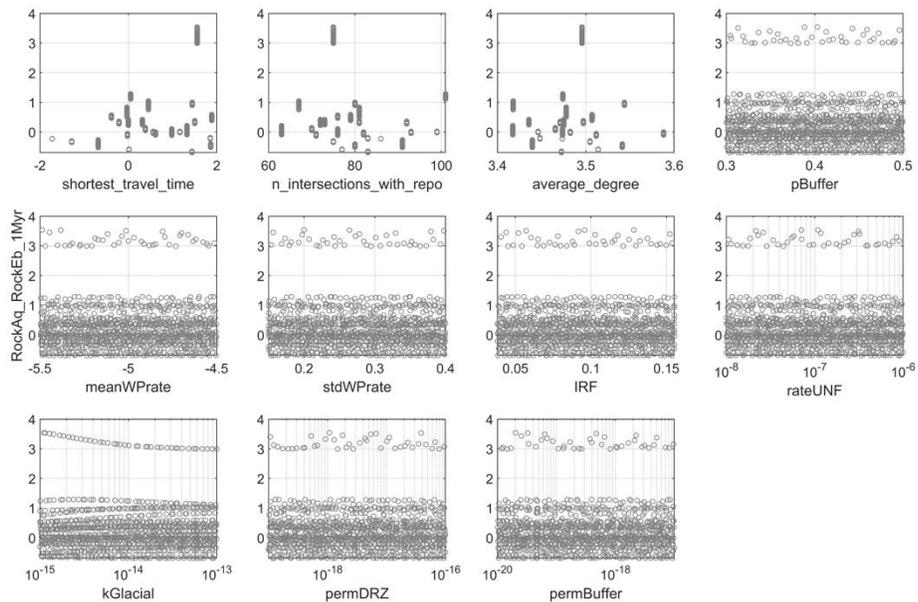
**Figure 4-20. Global sensitivity analysis results for the ratio between the aquifer to east boundary flux and the rock to east boundary flux at 1 Ma.**



**Figure 4-21. Scatterplots of the ratio between the aquifer to east boundary flux and the rock to east boundary flux at 1 Ma versus uncertain parameters and graph metrics.**



**Figure 4-22. Global sensitivity analysis results for the ratio between the rock to aquifer flux and the rock to east boundary flux at 1 Ma.**



**Figure 4-23. Scatterplots of the ratio between the rock to aquifer flux and the rock to east boundary flux at 1 Ma versus uncertain parameters and graph metrics.**

The ratio between the rock-to-aquifer flux and the rock-to-east-boundary flux at 1Ma is the final scalar QoI analyzed in this section. The sensitivity analysis results in Figure 4-22 are an interesting departure from results from the other QoIs because almost all the variance is attributed to interaction effects when the graph metrics are not included in the analysis. This is an instance of over-fitting, which can be seen by looking at the scatterplots in Figure 4-23. The first three plots graph the flux ratio against the graph metrics. At each value of the graph metric (for all graph metrics) the points are very closely clustered. In other words, the variation within the 40 parameter loop samples in a single spatial realization is small. This trend occurs for all spatial realizations, so most of the variance is between spatial realizations (due to spatial heterogeneity) and not due to within-group or parameter uncertainty. This means that when the surrogate model does not have access to any variables that characterize spatial heterogeneity, it tries to attribute that variance to the parameters. Since none of those parameters have a significant effect on their own, its only option for trying to characterize the variance is through interactions. There are sufficiently many parameters compared to the total number of simulations that the PCE and polynomial surrogate models can explain much of the variance via two-way interactions, but these results are spurious.

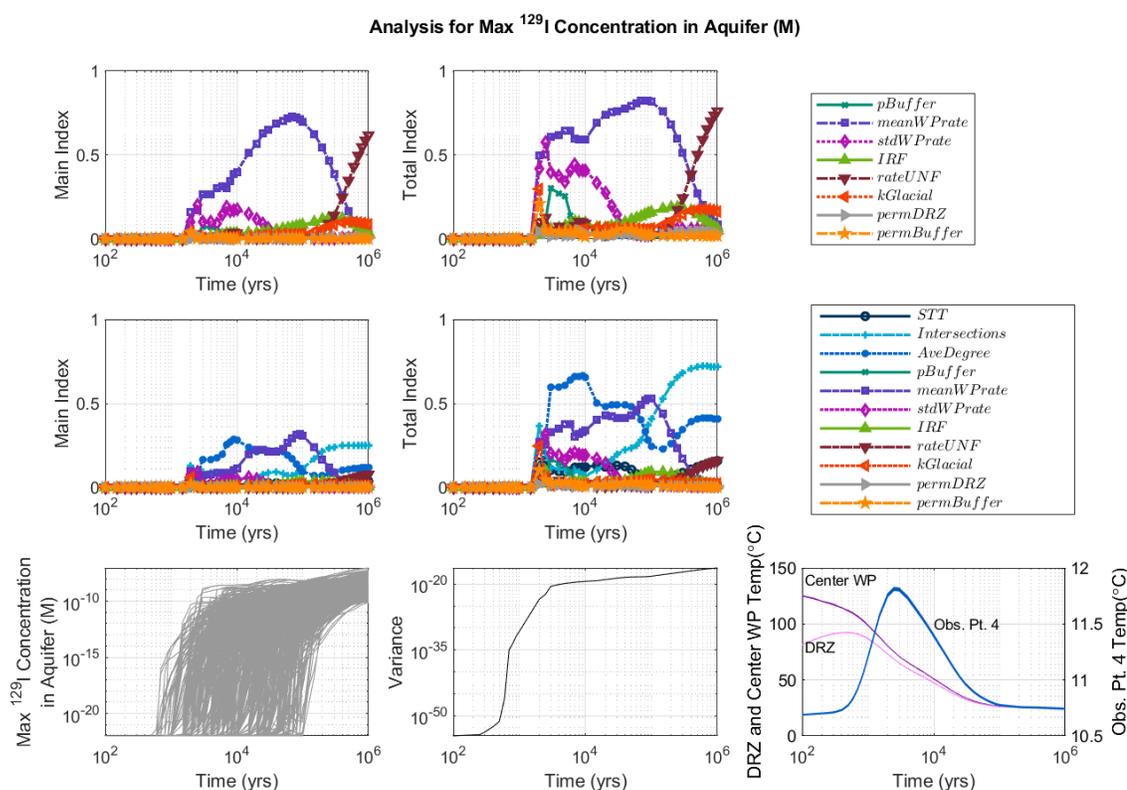
### 4.3.2 Sensitivity Results for Time-Dependent QoIs

This section presents sensitivity analysis results for QoIs that are tracked over time. The analyses were performed precisely as was done with the scalar QoIs except that the process of fitting the surrogate model and estimating Sobol' indices was repeated at each time step. This means that the QoI value at each timestep was treated as an individual QoI and sensitivity results are plotted over time to demonstrate how sensitivity changes. As in the scalar QoI analysis, the Sobol' indices were estimated with surrogates that did not include graph metrics to represent spatial heterogeneity and with surrogates that included these metrics.

In each of the time-dependent sensitivity analysis figures, the top two plots show the main and total index values over time for the surrogates that do not account for spatial heterogeneity. The plots in the middle row of the figures show the main and total index values over time for the surrogates that do account for spatial heterogeneity. The bottom row of plots is meant to facilitate more in-depth understanding of the sensitivity analysis result. The leftmost plot shows the time history of the QoI. The middle plot shows the variance between the realizations over time. This is helpful for understanding seemingly aberrant behavior in the sensitivity indices since this can happen when the variance is close to zero. The indices are a variance-decomposition tool for sensitivity analysis, so they make sense only if there is enough variance to decompose. The final plot shows the temperature time histories for the centermost waste package in the repository, observation point 4 (see Figure 4-5), and the point in the DRZ that is directly above the centermost waste package in the repository. Observation point 4 is in the rock just below the aquifer, in line with the center of the repository and on the first deterministic fracture connected to the aquifer to the east of the repository as shown in Figure 4-10. Note that, due to differences in scale, there are two separate vertical axes in the bottom right figure in the time-dependent results. The temperature at observation point 4 is substantially cooler; it is plotted on a different scale so the time at which it reaches its peak is discernable.

Sensitivity results over time for maximum  $^{129}\text{I}$  are plotted in Figure 4-24. When the graph metrics are not included, the parameters driving waste package corrosion (*meanWPrate*, *stdWPrate*) dominate. When the graph metrics are included, these effects are diminished and the

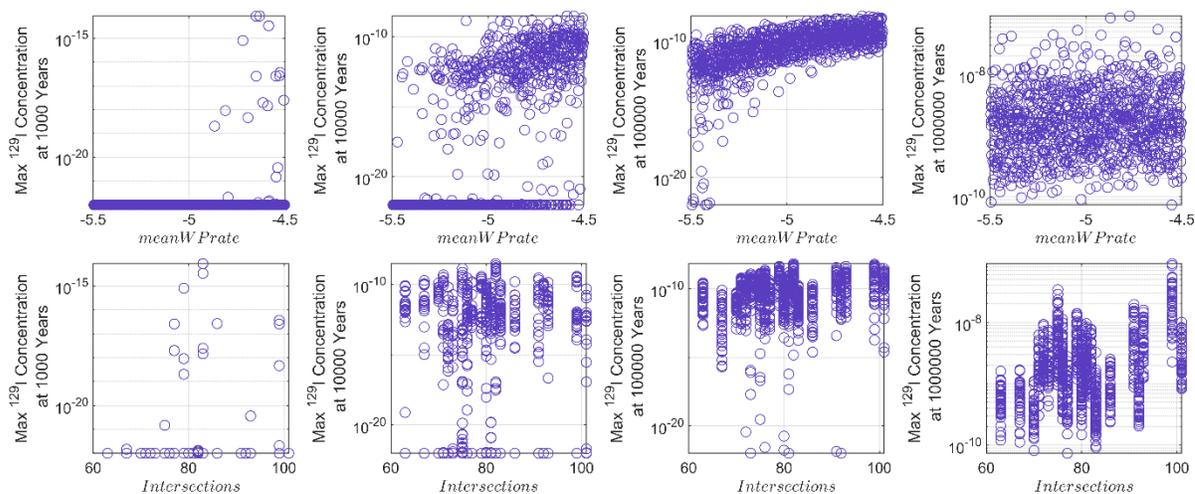
average degree rises in importance, with the number of intersections becoming important towards the end of the simulation. Recall that the sensitivity analysis results for peak  $^{129}\text{I}$  concentration (see Figure 4-12) identified *rateUNF* as the most important parameter uncertainty and *kGlacial* as somewhat important, which is consistent with the results near the end of the simulation in Figure 4-24. This means that the two analyses are in agreement, since the peak  $^{129}\text{I}$  concentration occurs near the end of the simulation. Comparison of the sensitivity results over time to the temperature plots for the centermost WP, the DRZ, and observation point 4, shows that the rising importance of *rateUNF*, *kGlacial*, and graph metrics corresponds with the leveling off of these temperatures. This change in the order of importance for parameters can also be verified with scatter plots at individual time points. Figure 4-25 shows the maximum  $^{129}\text{I}$  concentration at  $10^3$ ,  $10^4$ ,  $10^5$ , and  $10^6$  years versus *meanWPrate* (top) and the number of intersections (bottom). The trend with *meanWPrate* is strong up to  $10^5$  years, but weaker at the end of the simulation, whereas the trend with the number of intersections is strongest at the end of the simulation. This suggests that waste package corrosion drives uncertainty in the maximum  $^{129}\text{I}$  concentration up to around  $10^5$  years and waste form degradation drives uncertainty later. The time-dependent Sobol' index results are useful in these types of changes over time without having to generate and examine hundreds of scatter plots.



**Figure 4-24 Global sensitivity analysis results over time for the maximum  $^{129}\text{I}$  concentration in the aquifer**

Figure 4-26 shows the time-dependent global sensitivity analysis results for the log<sub>10</sub>-transformed maximum  $^{129}\text{I}$  concentration. This analysis was performed to test the capacity for

transformations to improve surrogate model quality for sensitivity analysis. The maximum  $^{129}\text{I}$  concentrations are extremely low, especially early in the simulation. Hence the variance is also low at some time points even though there are differences in the concentrations. Though maximum  $^{129}\text{I}$  concentrations this low may not be relevant with respect to repository performance, understanding which input uncertainties cause this variation can be helpful for understanding model behavior and performing the analysis using a transformation may inform future strategies for improving surrogate modeling. Transforming to a linear scale adjusts the variance to a more reasonable scale so it can be more easily characterized by the surrogate models.



**Figure 4-25. Scatterplots of the maximum  $^{129}\text{I}$  concentration versus *meanWPrate* and the number of intersections with the repository travel time at four time points.**

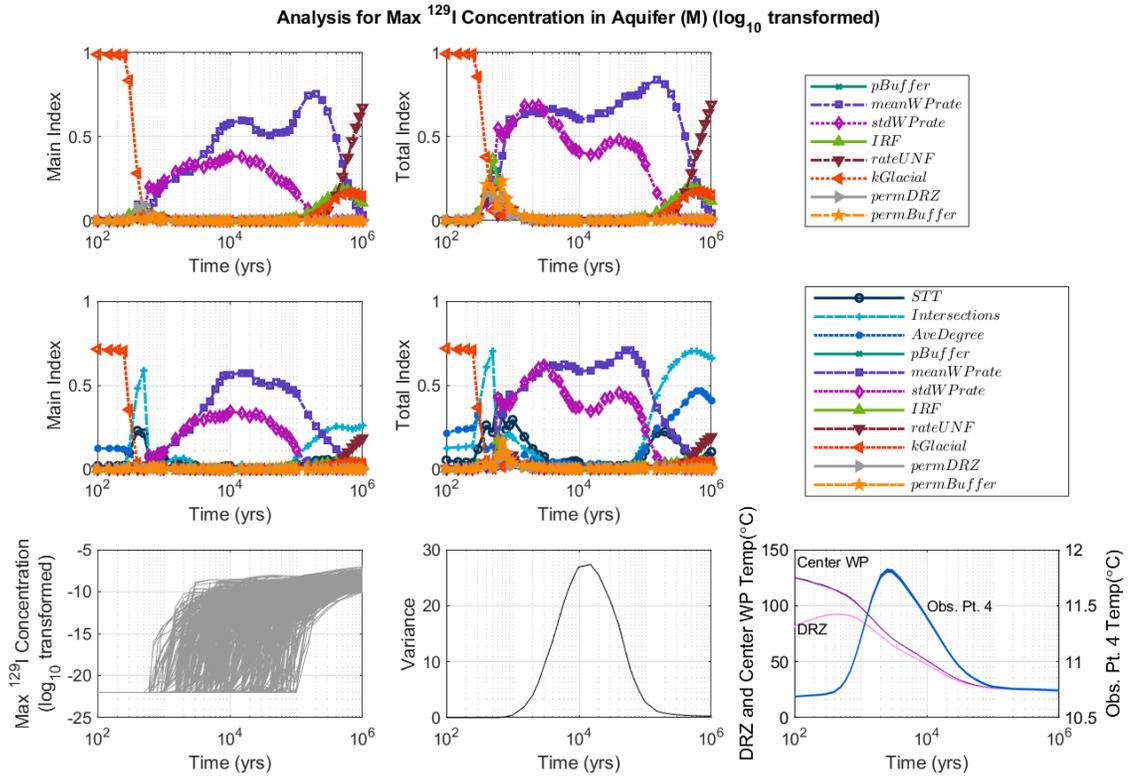
The results in Figure 4-26 are similar to the results for untransformed  $^{129}\text{I}$  concentrations in Figure 4-24, except that the waste package corrosion rate parameters do not decrease as much in significance when the graph metrics are included in the analysis in the middle of the simulation. This is consistent with the scatter plots in the top of Figure 4-25, which show that there is a demonstrable trend between *meanWPrate* and  $^{129}\text{I}$  concentration, at 10,000 years. The analysis on untransformed  $^{129}\text{I}$  concentrations did not identify a large main effect from *meanWPrate* at that time in the analysis that included graph metrics. The transformation leads to more clearly significant main effects for the waste package degradation parameters. Sensitivity results near 100 years should not be considered – these results are spurious due to the lack of any meaningful variance this early in the simulation. Transformations may not be appropriate for all QoIs, but in this case the transformation is appropriate because of the scale of the QoI. Similarly, the sensitivity analysis results added to understanding in this case rather than contradicting information from the analysis on the untransformed QoI.

Time-dependent global sensitivity analysis results for the fraction of tracer remaining in the repository are shown in Figure 4-27. These results are interesting because of the different behavior of the sensitivity indices over time between the analysis graph with graph metrics and the analysis without graph metrics. Without graph metrics, nearly all the variance is attributed to the porosity

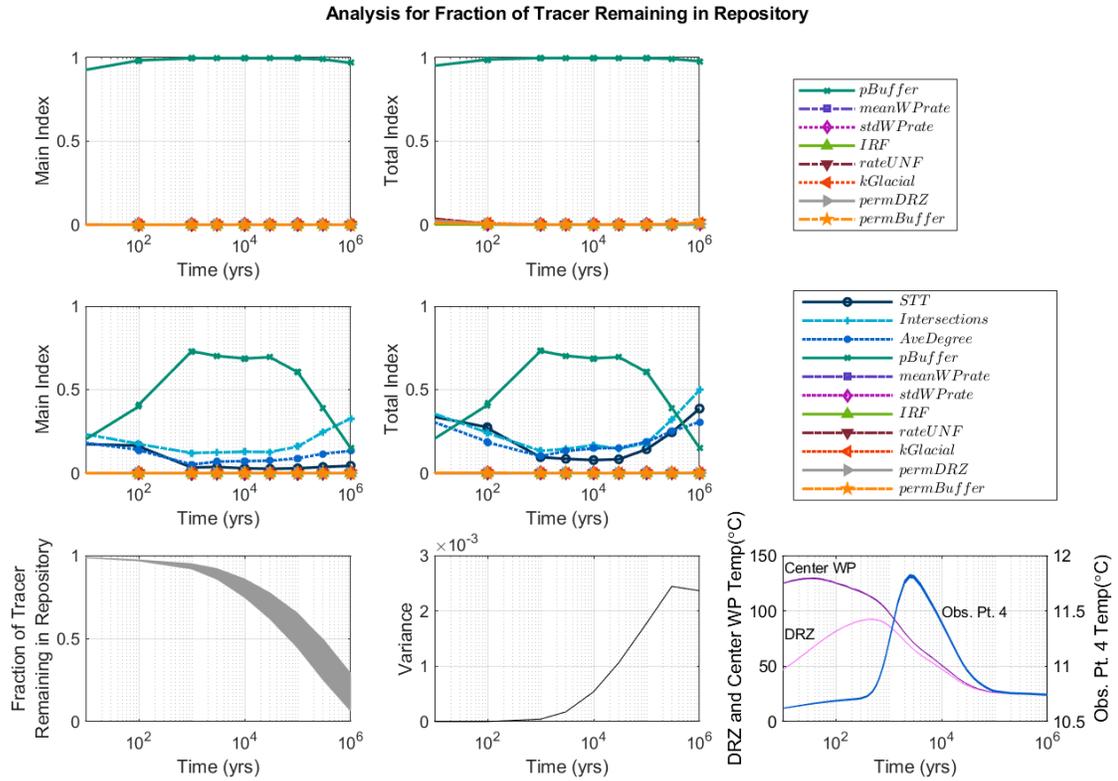
of buffer. This makes sense because high porosity in the buffer will lead to lower fractional mass flux (Figure 4-19) out of the repository, so more of the tracer will remain. When graph metrics are included in the analysis, they have some importance, though *pBuffer* still dominates for most of the simulation.

Figure 4-28 shows the global sensitivity analysis results for the water flux from the rock to the aquifer over time. The main and total effects for the parameter uncertainties are not significant when the graph metrics are included, and only *kGlacial* appears to have a potentially significant effect when the graph metrics are excluded. This, combined with the low main effects for all of the other uncertain parameters in the analysis without graph metrics, suggests that the dominant uncertainty affecting water flux from the rock to the aquifer is spatial heterogeneity. Of the different measures of spatial heterogeneity, the number of intersections with the repository has the most significant main effect.

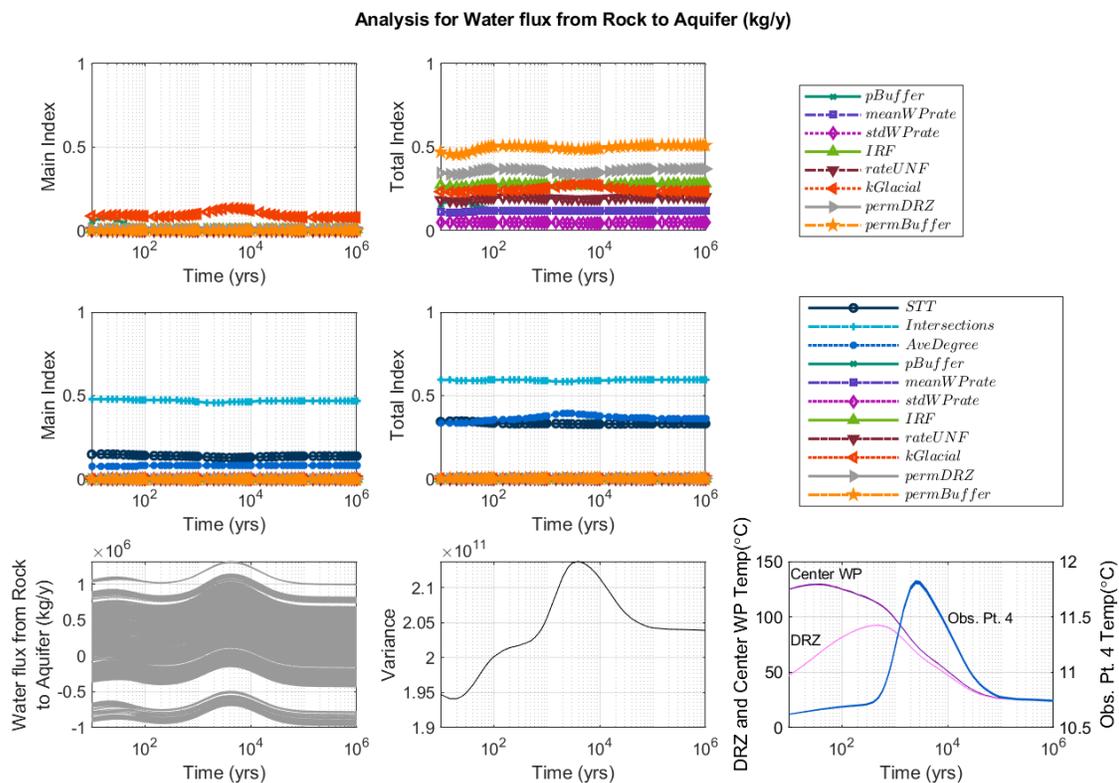
The results for this QoI show that the flux of water from the rock to the aquifer is primarily affected by differences in stochastically placed fractures. Though the analysis including graph metrics shows that spatial heterogeneity is the dominant uncertainty, it is interesting and unexpected that uncertainty in number of intersections with the repository is more important than uncertainty in the average degree. All of the graph metrics have higher total effect Sobol' indices than main effect Sobol' indices, suggesting significant interactions. With respect to the parameter uncertainties, it makes sense that *kGlacial* appears to have a small potentially significant main effect when the graph metrics are not included in the analysis. As *kGlacial* increases, the aquifer becomes less resistant to inflow. The slight peak in the importance of *kGlacial* corresponds to the increase in upward flow during the thermal period, as expected. Much like the analysis in Figure 4-22, the relatively high total effects for some parameters in the analysis without graph metrics may be the result of overfitting; none of the parameter uncertainties can explain the variance, but there are enough parameters for the surrogate to explain some of it via interactions.



**Figure 4-26. Global sensitivity analysis results over time for the log (base 10) transformed maximum  $^{129}\text{I}$  concentration in the aquifer.**

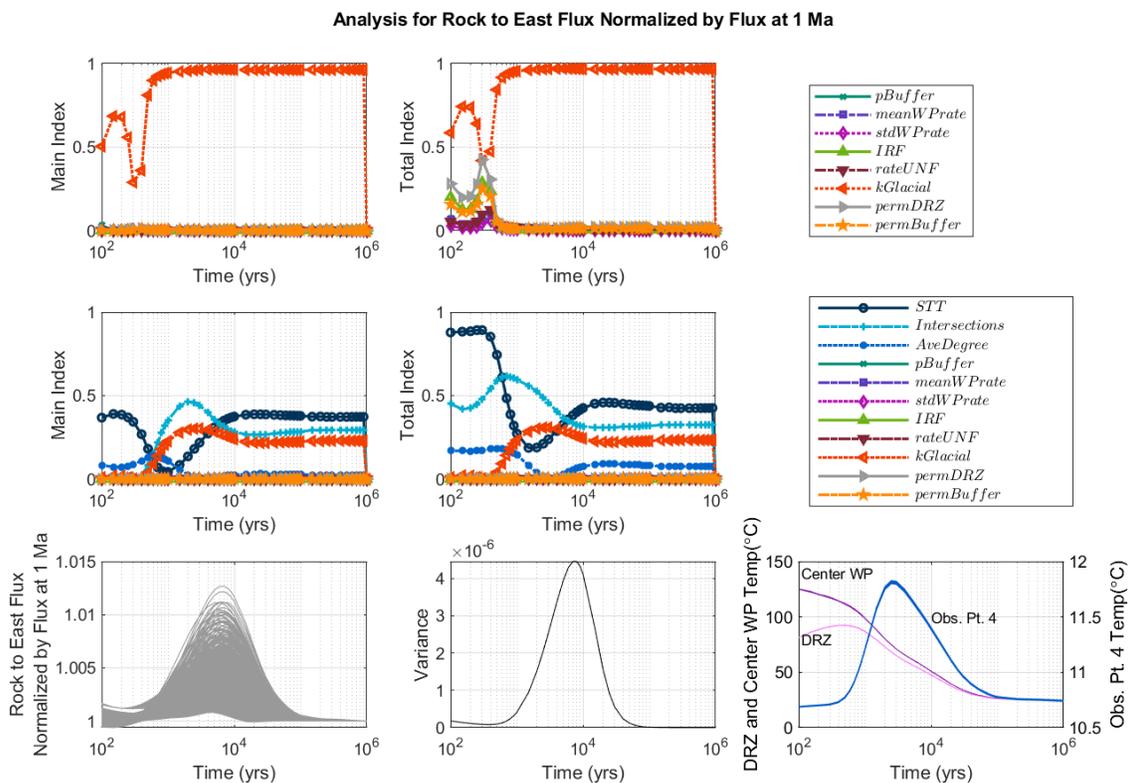


**Figure 4-27. Global sensitivity analysis results over time for the fraction of tracer remaining in the repository.**



**Figure 4-28. Global sensitivity analysis results over time for the water flux from the rock to the aquifer.**

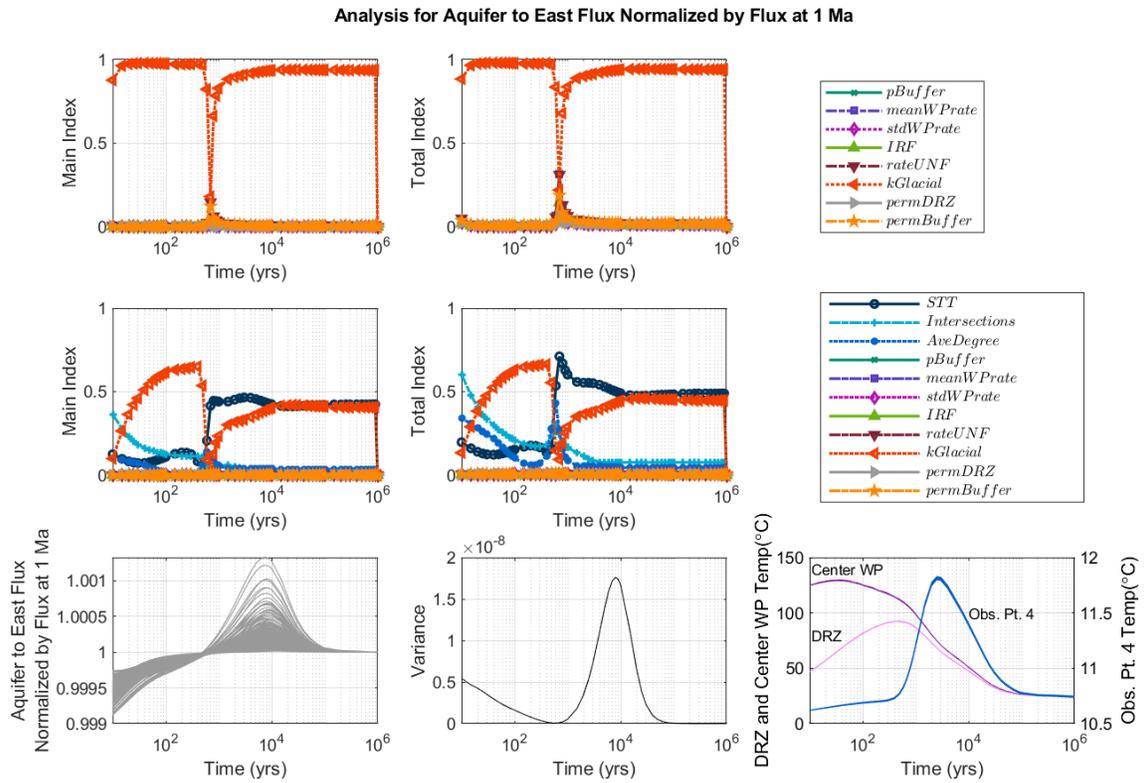
Global sensitivity results for the flux from the rock to the east boundary normalized by the flux at 1 Ma are plotted in Figure 4-29. This is another new QoI and, conceptually, the higher this ratio is, the higher the flux from the rock to east boundary. When the graph metrics are not included in the surrogate model, the Sobol' indices attribute most of the variance to *kGlacial*, which was expected. With graph metrics included in the surrogate model, *kGlacial* is still significant, but the shortest travel time and number of intersections gain importance. At around 1000 years, the number of intersections dominates, but the shortest travel time becomes more significant later on as the flux begins to decrease. This seems reasonable because the number of intersections with the repository affects only flow out of the repository (early), but the shortest travel time characterizes travel into the aquifer (later).



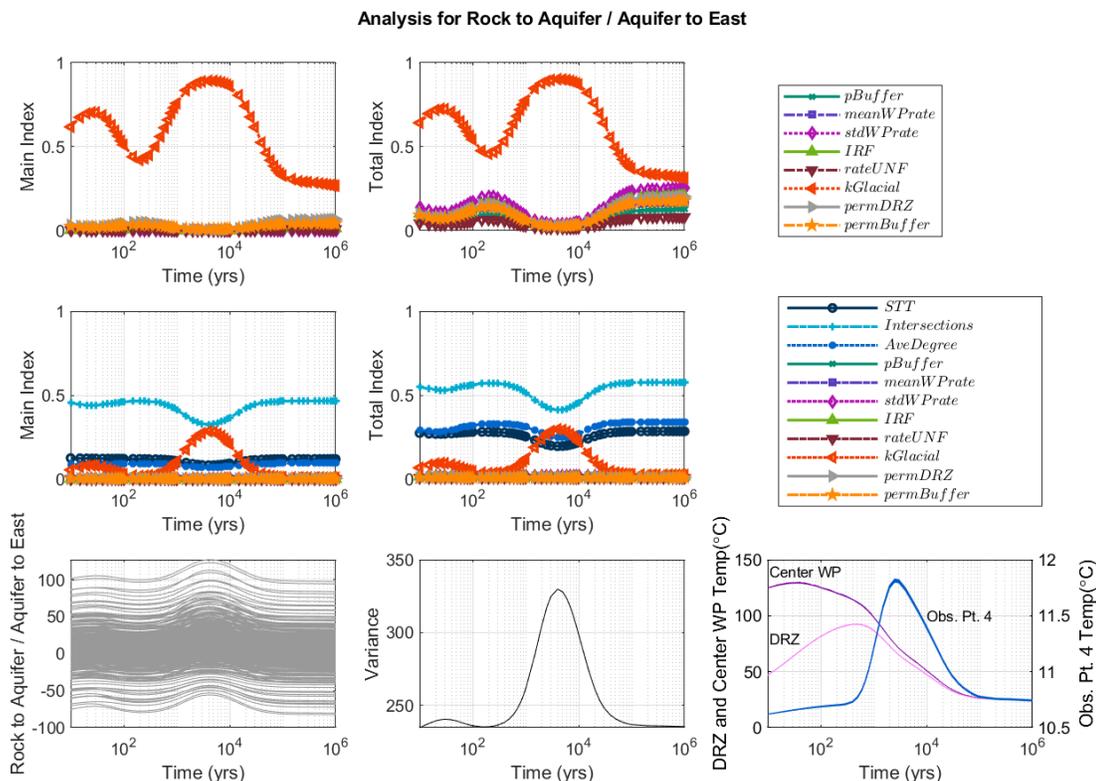
**Figure 4-29. Global sensitivity analysis results over time for the rock to east boundary flux normalized by the flux at 1 Ma.**

Sensitivity results over time for the flux from the aquifer to the east boundary are plotted in Figure 4-30. These results demonstrate the value in adding a plot of the variance over time to the analysis. The sharp drop in importance for *kGlacial* just before 1000 years seems unreasonable. However, the variance plot and the time series plot for the QoI show that this drop corresponds to when the variance decreases to zero, when the water flux increases. As in the other analyses, inclusion of the graph metrics decreases the importance of *kGlacial*, though it still explains more than half of the variance early in the simulation. For the graph metrics, the number of intersections with the repository has some importance early (when flux is low) but the shortest travel time is more important later in the simulation (when flux is higher), as was the case with the rock to east boundary flux.

Figure 4-31 shows the global sensitivity analysis results for the ratio between the rock to aquifer flux and the aquifer to east boundary flux over time. This QoI is meant to reflect upward flow in the rock. Inclusion of graph metrics dramatically reduces the importance of *kGlacial* throughout the simulation, though there is still a peak between 1000 and 10,000 years which also corresponds to a peak in the value of the flux ratio.



**Figure 4-30. Global sensitivity analysis results over time for the aquifer to east boundary flux normalized by the flux at 1 Ma.**



**Figure 4-31. Global sensitivity analysis results over time for the ratio between the rock to aquifer flux and the aquifer to east boundary flux.**

#### 4.4 Summary

In this chapter, we presented many sensitivity analysis results for a large study performed on 1000 PFLOTRAN runs performed on an updated crystalline reference case. We outlined the case study updates, listed the inputs and quantities of interest, and presented the Sobol’ variance decomposition results (main and total effect indices) for scalar QoIs. In addition, we performed time-dependent sensitivity analysis, where the Sobol’ sensitivity indices are calculated at many time points throughout the simulation. We emphasize that this is a new SA capability (time-dependent SA) that has not previously been performed on the crystalline reference case. We also note that this capability involves a significant amount of computation and software infrastructure. For each time point in Figures 4-24 to 4-31, a surrogate must be constructed over all the simulation results for the QoI at that time point, and the SA calculations to determine the sensitivity indices must be performed. This additional computational infrastructure has the benefit of more clearly seeing how various parameters change their importance over time. With that information, we can better relate the sensitivity analysis results to the physics of the problem.

Specific conclusions from the results presented above are:

- The log-transformation improved results for  $^{129}\text{I}$ . This transformation was justified by comparing the GSA results to scatter plots at different points in time to ensure that the relationships identified in the GSA are apparent in the data at corresponding times. This also made sense to do because of the very low variance in  $^{129}\text{I}$  due to variation between low

orders of magnitude. Comparing the time-dependent variance plot to QoI time series plot helped us see where we should be getting sensitivity results but weren't due to negligible variance.

- The uncertainty in *IRF* had less of an impact relative to *rateUNF* on the peak  $^{129}\text{I}$  concentration. This is likely because the distribution on *IRF* was much narrower than for *rateUNF*. Based on the correlation between *IRF* and mean peak  $^{129}\text{I}$  in the simulations, an *IRF* value of 0.1 generates nearly half of the  $^{129}\text{I}$  in the mean peak  $^{129}\text{I}$  concentration.
- We developed an important capability (identified last year as a potential area for development) to generate and plot time-dependent sensitivity analyses.
- Graph metrics tended to dominate sensitivity analysis results when included in the surrogate model construction, suggesting that spatial heterogeneity may drive uncertainty in many of the QoIs. However, it was not clear from our analysis to what extent this dominance may be influenced by the repeated values of the graph metrics in the sampling structure (e.g. one DFN run with 40 epistemic samples means those 40 samples all have the same set of graph metrics). We were able to identify when parameters or spatial heterogeneity are important, but it is not always clear which type of uncertainty dominates due to this potential effect.
- Sensitivity analyses for fluxes, flux ratios, and repository retention QoI (MdRT) made phenomenological sense. The addition of these QoIs is an important advancement for the crystalline reference case.
- The number of fracture intersections with the repository correlates strongly with the direction of flow through the rock; there is more upward flow in the rock relative to horizontal flow as the number of intersections increases. The number of intersections also correlates with higher peak  $^{129}\text{I}$ , meaning fracture intersection avoidance has significant performance implications. The average degree and STT also correlate well, though the number of intersections has a stronger correlation.

Additionally, the analysis and results suggest future opportunities for development:

- One area for future study would be to further interrogate our ability to measure the relative importance of spatial heterogeneity and epistemic parameter uncertainties using surrogate models in the context of repeated sampling. Statistically, it could make sense to sample a unique spatial loop for each epistemic loop (so remove all repeated samples in our scheme) to better understand the extent to which repeated samples may interfere with surrogate construction. However, the burden of generating that many ECPMs may make this option impractical and alternative methods for investigation may be necessary.
- Similarly, this analysis differed from last year in that we generated new samples within each epistemic loop under the assumption that removing repeated sampling in the epistemic loop would provide better coverage of the sample space and, consequently, better surrogate models. Though this qualitatively appeared to be the case because sensitivity analysis results were straightforward to interpret, direct comparisons between the analyses with the

different sampling strategies could inform a better understanding of the magnitude of this effect. This change in sampling strategy came at the cost of easily distinguishing between the effects of spatial heterogeneity and epistemic parameter uncertainty. We can assess whether this change was worth that cost by understanding the extent to which it improved sensitivity analysis results.

- The next step for flux, flux ratio, and repository retention QoIs is to investigate the relationship between these QoIs and peak  $^{129}\text{I}$  further. Essentially, we want to tie the QoIs we are using to better understand the behavior of our model to the QoIs we are using to understand repository performance.

## 4.5 References: Chapter 4

1. Stein, E. R., J. M. Frederick, G. E. Hammond, K. L. Kuhlman, P. E. Mariner and S. D. Sevougian (2017). "Modeling Coupled Reactive Flow Processes in Fractured Crystalline Rock," *Proceedings of the 16th International High-Level Radioactive Waste Management (IHLRWM 2017) Conference*, Charlotte, North Carolina, April 9-13, American Nuclear Society.
2. Joyce, S., L. Hartley, D. Applegate, J. Hoek and P. Jackson (2014). "Multi-scale groundwater flow modeling during temperate climate conditions for the safety assessment of the proposed high-level nuclear waste repository site at Forsmark, Sweden," *Hydrogeology Journal*, **22**(6):1233-1249
3. Hyman, J. D., S. Karra, N. Makedonska, C. W. Gable, S. L. Painter and H. S. Viswanathan (2015). "dfnWorks: A discrete fracture network framework for modeling subsurface flow and transport," *Computers & Geoscience*, **84**:10-19.
4. Mariner, P.E., E.R. Stein, J.M. Frederick, S.D. Sevougian, G.E. Hammond, and D.G. Fascitelli (2016), *Advances in Geologic Disposal System Modeling and Application to Crystalline Rock*, FCRD-UFD-2016-000440, SAND2016-9610 R, Sandia National Laboratories, Albuquerque, NM.
5. SKB 2006. *Long-term safety for KBS-3 repositories at Forsmark and Laxemar – a first evaluation*. SKB TR-06-09. Svensk Kärnbränslehantering AB, Stockholm, Sweden.
6. Hammond, G.E., P.C. Lichtner, C. Lu and R.T. Mills (2011). "PFLOTRAN: Reactive Flow and Transport Code for Use on Laptops to Leadership-Class Supercomputers." *Groundwater Reactive Transport Models*. F. Zhang, G. T. Yeh and J. Parker. Bentham Science Publishers.
7. Lichtner, P.C. and G.E. Hammond (2012). *Quick Reference Guide: PFLOTRAN 2.0 (LA-CC-09-047) Multiphase-Multicomponent-Multiscale Massively Parallel Reactive Transport Code*. LA-UR-06-7048. December 8, 2012. Los Alamos National Laboratory, Los Alamos, New Mexico.
8. Adams, B.M., Bauman, L.E., Bohnhoff, W.J., Dalbey, K.R., Ebeida, M.S., Eddy, J.P., Eldred, M.S., Hough, P.D., Hu, K.T., Jakeman, J.D., Stephens, J.A., Swiler, L.P., Vigil, D.M., and Wildey, T.M., [\*Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity\*](#)

- Analysis: Version 6.0 User's Manual*, SAND2014-4633, July 2014. Other versions released December 2009 ([Version 5.0](#)), December 2010 ([Version 5.1](#)), November 2011 ([Version 5.2](#)), February 2013 ([Version 5.3](#)), May 2013 ([Version 5.3.1](#)), November 2013 ([Version 5.4](#)), November 2014 ([Version 6.1](#)), May 2015 ([Version 6.2](#)), November 2015 ([Version 6.3](#)), May 2016 ([Version 6.4](#)), November 2016 ([Version 6.5](#)), May 2017 ([Version 6.6](#)), November 2017 ([Version 6.7](#)), May 2018 ([Version 6.8](#)), November 2018 ([Version 6.9](#)), May 2019 ([Version 6.10](#)).
9. Sevougian, S. D., E. R. Stein, G. E. Hammond, P. E. Mariner, J. M. Frederick, and E. Basurto 2018. "Simulating the Effect of Fracture Connectivity on Repository Performance with GDSA Framework – 18589," in *Proceedings of WM2018 Conference*, March 18-22, 2018, Phoenix, AZ, USA, SAND2017-12677C.
  10. Mariner, P.E., Stein, E., Frederick, J.M, Sevougian, S.D., and Hammond, G.E. *Advances in Geologic Disposal System Modeling and Shale Reference Cases*. SFWD-SFWST-2017-000044 / SAND2017-10304R. Sandia National Laboratories, Albuquerque, NM (2017).
  11. Viswanathan, H. S., Hyman, J. D., Karra, S., O'Malley, D., Srinivasan, S., Hagberg, A., & Srinivasan, G. (2018). "Advancing graph-based algorithms for predicting flow and transport in fractured rock." *Water Resources Research*, 54(9), 6085-6099.
  12. Srinivasan, G., Hyman, J. D., Osthus, D. A., Moore, B. A., O'Malley, D., Karra, S., ... & Viswanathan, H. S. (2018). "Quantifying topological uncertainty in fractured systems using graph theory and machine learning." *Scientific reports*, 8(1), 1-11.
  13. Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart, "Exploring network structure, dynamics, and function using NetworkX", in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008.
  14. L.P. Swiler, J.C. Helton, E. Basurto, D.M. Brooks P.E. Mariner, L.M. Moore, S. Mohanty, S.D. Sevougian, and E.R. Stein. "Status Report on Uncertainty Quantification and Sensitivity Analysis Tools in the Geologic Disposal Safety Assessment (GDSA) Framework." SAND2019-13835R. Sandia National Laboratories, Albuquerque, NM (2019).
  15. Mariner, P. E., M.A. Nole, E. Basurto, T.M. Berg, K.W. Chang, B.J. Debusschere, A.C. Eckert, M.S. Ebeida, M. Gross, G.E. Hammond, J. Harvey, S.H. Jordan, K.L. Kuhlman, T.C. LaForce, R.C. Leone, W.C. McLendon III, M.M. Mills, H.D. Park, F.V. Perry, A. Salazar III, D.T. Seidl, S.D. Sevougian, E.R. Stein and L.P. Swiler. "Advances in GDSA Framework Development and Process Model Integration." Sandia National Laboratories, Albuquerque, NM (2020).
  16. Mariner, P. E., E. Basurto, D. M. Brooks, E. R. Stein and L. P. Swiler. "A Study of Epistemic and Spatial Aleatory Uncertainties on Crystalline Repository Performance." SAND2019-13399 PE. Albuquerque, New Mexico, Sandia National Laboratories.
  17. LaVenue, A.M. and RamaRao, B.S. "A Modeling approach to address spatial variability within the Culebra Dolomite transmissivity field." SAND92-7306.

18. Rechar, R.P, B. W. Arnold, B. A. Robinson, J. E. Houseworth, “Transport modeling in performance assessments for the Yucca Mountain disposal system for spent nuclear fuel and high-level radioactive waste,” *Reliability Engineering & System Safety*, Volume 122, 2014.
19. Helton J.C. 2011. Quantification of Margins and Uncertainties: Conceptual and Computational Basis. *Reliability Engineering and System Safety* 96:976-1013.
20. Helton J.C. and C.J. Sallaberry. 2012. Uncertainty and Sensitivity Analysis: From Regulatory Requirements to Conceptual Structure and Computational Implementation. *IFIP Advances in Information and Communication Technology* 377 AICT:60-76.
21. Hacking I. 2006. The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference. 2<sup>nd</sup> edition. London; New York: Cambridge University Press.
22. Swiler, L.P., E. Basurto, D.M. Brooks, A. C. Eckert, P. E. Mariner, T. Portone and E.R. Stein. “Advances in Uncertainty Quantification and Sensitivity Analysis Methods and Applications in GDSA Framework.” SAND2020-10802 R. Sandia National Laboratories, Albuquerque, NM.
23. Mariner, P. E., W. P. Gardner, G. E. Hammond, S. D. Sevougian, and E. R. Stein (2015). Application of Generic Disposal System Models. FCRD-UFD-2015-000126 / SAND2015-10037R. Sandia National Laboratories, Albuquerque, NM.
24. Anton Leijnse. 1992. Three-dimensional modeling of coupled flow and transport in porous media. Ph.D. Dissertation. University of Notre Dame, USA. Order Number: UMI Order No. GAX93-06340.
25. Carter, J. T., A. J. Luptak, J. Gastelum, C. Stockman, and A. Miller 2013. *Fuel Cycle Potential Waste Inventory for Disposition*. FCRD-USED-2010-000031 Rev 6. Savannah River National Laboratory, Aiken, SC.
26. Espriu-Gascon, A., A. Martínez-Torrents, D. Serrano-Purroy, J. Giménez, J. de Pablo, I. Casas, “Contribution of phases segregated from the UO<sub>2</sub> matrix to the release of radionuclides from spent nuclear fuel and duration of the Instant Release Fraction (IRF)”, *Journal of Nuclear Materials*, Vol. 532, 2020.<https://doi.org/10.1016/j.jnucmat.2020.152066>.
27. Johnson, L, C, Ferry, C. Poinssot, P. Lovera, “Spent fuel radionuclide source-term model for assessing spent fuel performance in geological disposal. Part I: Assessment of the instant release fraction.” *Journal of Nuclear Materials*, Vol. 346. 2005. <https://doi.org/10.1016/j.jnucmat.2005.04.071>.
28. Lemmens, K., E. González-Robles, B. Kienzler, E. Curti, D. Serrano-Purroy, R. Sureda, A. Martínez-Torrents, O. Roth, E. Slonszki, T. Mennecart, I. Günther-Leopold, Z. Hózer, “Instant release of fission products in leaching experiments with high burn-up nuclear fuels in the framework of the Euratom project FIRST- Nuclides.” *Journal of Nuclear Materials*, Vol. 484, 2017, <https://doi.org/10.1016/j.jnucmat.2016.10.048>.

29. Sassani, D. C., C. F. Jove-Colon, P. Weck, J. L. Jerden Jr., K. E. Frey, T. Cruse, W. L. Ebert, E. C. Buck, R. S. Wittman, F. N. Skomurski, K. J. Cantrell, B. K. McNamara, C. Z. Soderquist. "Integration of EBS Models with Generic Disposal System Models." SAND2012-7762 P and FCRD-UFD-2012-000277.
30. Stevenson, D.R., E.T. Kozak, C.C Davison, M. Gascoyne, R.A. Broadfoot. "Hydrogeologic Characteristics of Domains of Sparsely Fractured Rock in the Granitic Lac du Bonnet Batholith, Southeastern Manitoba, Canada", AECL-11558, COG-96-117, June 1996.
31. Nagra. Project Opalinus clay, demonstration of disposal feasibility for spent fuel, vitrified high-level waste and long-lived intermediate-level waste. Safety Report, Technical report 02-05 (2002).

## 5. CRYSTALLINE REFERENCE CASE UA NESTED WORKFLOW

The GDSA computational framework is largely comprised of two primary computational capabilities: PFLOTRAN and Dakota, as described in Section 1.2. An analysis supporting the performance assessment of a geologic repository necessarily requires the development and use of many additional connective computational components beyond these two primary components, including the development of input files, scripts that connect the capabilities of PFLOTRAN and Dakota, scripts to submit calculations to computational resources, capabilities to gather and post-process results, and much more. The collection of these computational components that comprise the complete body of work required to produce results of interest from a computational simulation capability, such as the GDSA framework, is hereafter referred to as an *analysis workflow*.

The development of analysis workflows to support GDSA analyses often occurs on an analysis-by-analysis basis and requires the expertise of highly experienced modelers/analysts who are familiar with the PFLOTRAN code base. Often, these complex analysis workflows involve many manual steps and continuous monitoring of simulations. The reproducibility of these analyses depends on how well the modeler/analyst organized and documented what was done, which can make it difficult to replicate previous analyses, to hand-off analyses between analysts, and/or to train new analysts to produce analysis workflows.

To improve analysis workflow automation, development, reproducibility, and traceability for repository PA simulations, the GDSA team began developing automated analysis workflows using the Next-Generation Workflow (NGW) capability in FY20. The NGW capability is an open-source engine that was developed at Sandia National Laboratories to provide analysts with a capability to construct, execute, and communicate end-to-end computational simulation analysis workflows [1]. This capability is a graphical, node-based interface that includes many pre-programmed support functions which are often utilized within computational simulation analysis workflows. NGW is available within the Dakota Graphical User Interface (GUI) [2] and is thus available to the GDSA analysis community.

Development of analysis workflows utilizing NGW for the GDSA Framework focuses on providing an automated workflow capability with the following objectives:

- Reduce the learning curve for new users to set up and run simulations and analysis workflows
- Speed up analysis workflow execution time
- Eliminate/reduce the need for manual intervention and allow for automated monitoring
- Reduce the potential for the introduction of human errors
- Increase traceability and reproducibility

A series of analysis workflows utilizing NGW for GDSA analysis exemplars, including a short course exercise and the Crystalline Reference Case, were developed in FY20 to pilot the use of this capability for the GDSA program [3]. Section 3.2.1 of the FY20 GDSA Framework report [3]

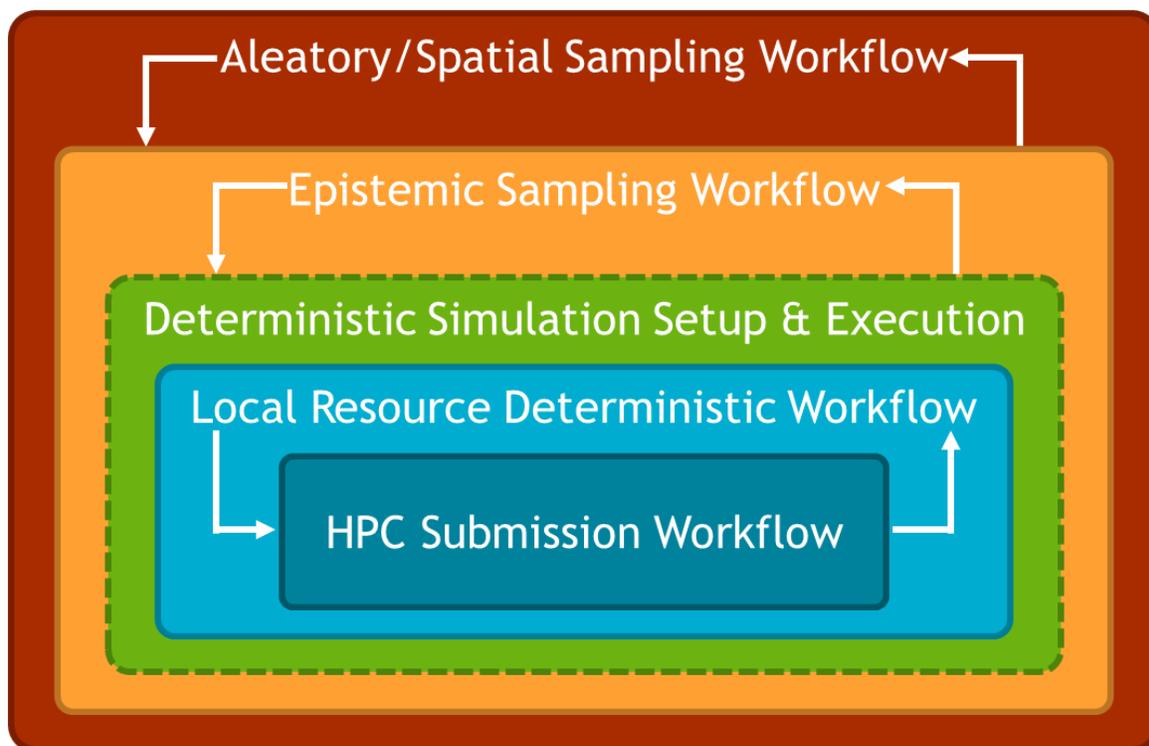
details the development of these workflows, provides examples of their use, and proposes a path forward for additional development. Appendix C of [3] provides a complete graphical workflow, demonstrating an uncertainty quantification study. This workflow sets parameters of interest, inserts these parameters into the PFLOTRAN input file, collects additional input files, runs PFLOTRAN, and develops and runs a post-processing script. The eventual goal of this work is to develop an analysis workflow library that can be made available to the GDSA analysis community.

The development and application of the analysis workflows for the Crystalline Reference Case described above continued in FY21 in support of the Uncertainty Analysis (UA) detailed in Chapter 4. The workflows developed for the UA are organized in a nested structure which is hereafter referred to as the *Crystalline Reference Case UA Nested Workflow*. Crystalline Reference Case UA Nested Workflow development, including the extension of the workflow to new QoIs and hardening the workflow to be more robust to changes in systems and supporting software versions, took place throughout the year to support the final production runs whose results are reported in Section 4.3. This development is detailed in Section 5.1. Crystalline Reference Case UA Nested Workflow application, described in Section 5.2, took place throughout the year as the workflow development progressed and included testing as well as two full UQ studies for the Crystalline Reference Case. In addition to the development and application of the Crystalline Reference Case UA Nested Workflow, a training course was developed to introduce PFLOTRAN developers and analysts to the NGW capability and its application for the GDSA framework. This course development and delivery is described in Section 5.3. Finally, takeaways and outcomes of the use of the analysis workflows are provided in Section 5.3, along with a discussion of future work.

## **5.1 Crystalline Reference Case UA Nested Workflow Development**

### **5.1.1 Crystalline Reference Case UA Nested Workflow Structure Summary**

The Crystalline Reference Case UA Nested Workflow is a series of nested analysis workflows that was first developed in FY20 to support the FY20 Crystalline Reference Case UA. These workflows are detailed in Appendix C of [3]. At a high level, the workflow structure is divided into three components including a deterministic engine to set up, run, and post-process a single PFLOTRAN simulation that is nested in the dual-loop sampling structure that is applied for the Crystalline Reference Case UA with an inner epistemic sampling loop representing parameter uncertainty and an outer aleatory/spatial sampling loop that represents uncertainty in the subsurface. This high-level structure is depicted visually in Figure 5-1. A more detailed view of the workflow structure is shown in Figure 5-2, which depicts the major files, actions, and codes that are employed at each level of the nested workflow structure.



**Figure 5-1. High-level view of the Crystalline Reference Case UA Nested Workflow structure.**

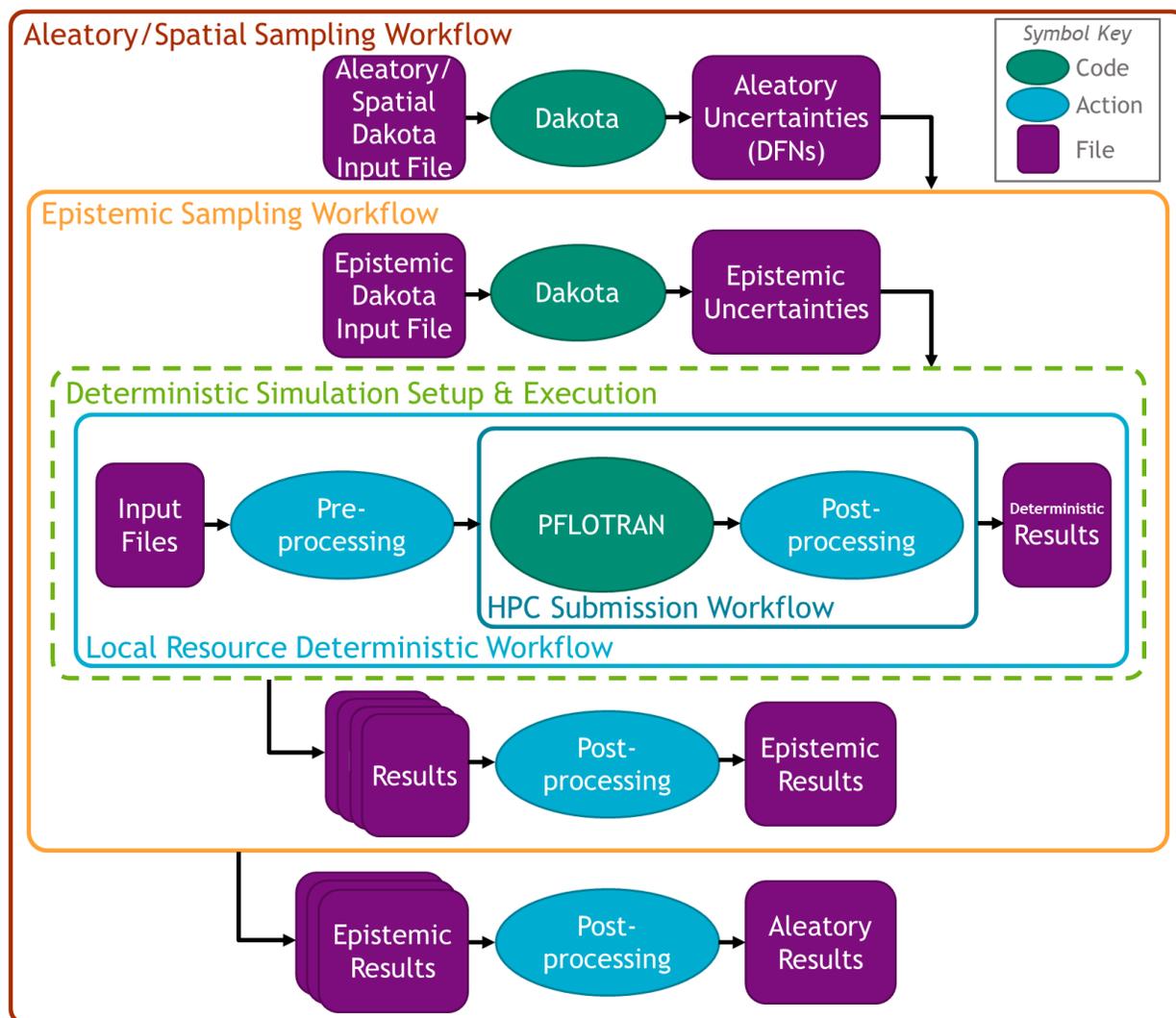
In this figure, each workflow element, distinguished by a solid border, represents a single workflow “.iwf” file that contains specific actions delineating each workflow. These actions include the transfer of information and data across “.iwf” files in the nested structure shown above. Each workflow that is contained within a workflow is implemented as a nested workflow node, which allows this transfer of data to occur.

At the center of the Crystalline Reference Case UA Nested Workflow structure are two workflows, shown in blue in Figure 5-1, that together comprise a single deterministic simulation run (including setup and execution), as shown in green. The first of these workflows, the Local Resource Deterministic Workflow, sets up the run on a local computational resource by collecting required input files that will be called during the calculations, setting the values for input parameters, templating the PFLOTRAN input file with these input values, preparing high-performance computing (HPC) submission scripts and post-processing scripts, and sending all required information and files to the HPC file system. The second workflow, the HPC Submission Workflow, which runs on the remote HPC system, then collects these files on the selected HPC, submits the PFLOTRAN job to the run queue, monitors the progress of the job, and post-processes calculations before returning final results to the first workflow on the local computational resource. These steps are shown in the detailed view of the workflow structure given in Figure 5-2.

The next workflow that is nested around this set of workflows comprising the deterministic simulation is the Epistemic Sampling Workflow, shown in orange in Figure 5-1. This workflow, which runs on the same local computational resource as the Local Resource Deterministic Workflow, includes a Dakota node that samples the uncertain parameters selected for epistemic

propagation and creates directories to execute  $n$  deterministic simulations via the coupled workflows that were described above, where  $n$  is the number of epistemic samples. Within these directories, input files for each deterministic simulation are saved with the values of sampled parameters replaced with the values selected by Dakota. Following the completion of all the epistemic realizations, the Epistemic Sampling Workflow aggregates the post-processed results, calculates summary statistics, and saves and plots QoI time histories at the epistemic level.

The Aleatory/Spatial Sampling Workflow, shown in red in Figure 5-1, is nested around the Epistemic Sampling Workflow. Similar to the Epistemic Sampling Workflow, the Aleatory/Spatial Sampling Workflow includes a Dakota node and also runs on the same local computational resource as the Local Resource Deterministic Workflow. At this level, the aleatory/spatial Dakota input file runs a list parameter study, a study type available in Dakota, over the selected number of DFNs. For each DFN, the Aleatory/Spatial Sampling Workflow initiates the Epistemic Sampling Workflow, which in turn initiates the selected number of epistemic realizations, as described previously. For each of these epistemic realizations, the DFN files for the DFN selected by Dakota are saved for use in the execution of each deterministic simulation. Following the completed execution of all simulations for all DFNs, the Aleatory/Spatial Sampling Workflow again aggregates the post-processed results at the aleatory/spatial level, calculates summary statistics, and saves and plots QoI time histories at the aleatory/spatial level. Finally, the Aleatory/Spatial Sampling Workflow saves the results file in the format that is needed for subsequent sensitivity analyses, which occur outside of the workflow execution. The detailed view of these elements of the Crystalline Reference Case UA Nested Workflow structure, including major files, actions, and codes, is shown in Figure 5-2.



**Figure 5-2. Detailed view of the Crystalline Reference Case UA Nested Workflow structure showing the four workflows along with important files, actions, and codes that are employed for each.**

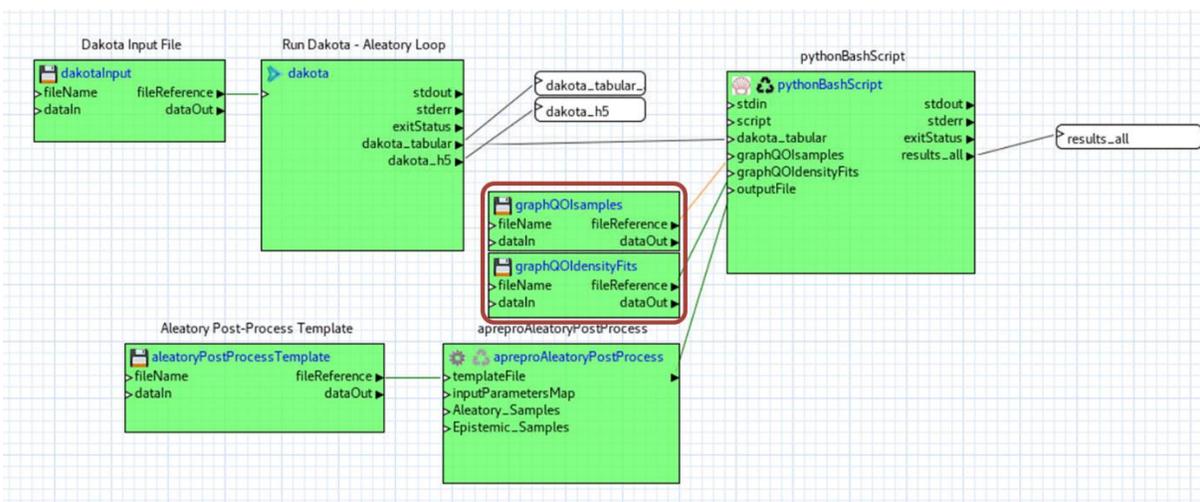
Most of the main features of this workflow structure were completed in FY20, at which point the system of workflows could be used to complete a large UQ study with many nested samples for the FY20 Crystalline Reference Case analysis. However, as work on the Crystalline Reference Case in FY21 progressed and as use of the Crystalline Reference Case UA Nested Workflow extended beyond the original developer, the need for continued development and updates arose. These updates involved changes to specific input files, workflow nodes, and settings, and thus respected the structure presented above. Major updates are described in Section 5.1.2.

### 5.1.2 Crystalline Reference Case UA Nested Workflow Updates

The PFLOTRAN input file for the Crystalline Reference Case was updated this year to incorporate new features for this analysis such as soil compressibility and a different waste package spacing as described in Section 4.2. These changes to the PFLOTRAN input file did not

have a significant impact on the workflow as the only changes needed were to point to the updated input file template and to add a new uncertain parameter, IRF (see Section 4.1.3), as a templated variable, meaning a variable for which the text could be replaced by a value selected by Dakota, in both the input file and the Local Resource Deterministic Workflow and as an uncertain parameter in the epistemic Dakota input file. Although large changes had been made to the PFLOTRAN input deck to update the analysis model, only minor changes were required to allow this updated model to be run through a full UQ study. This means that the Crystalline Reference Case UA Nested Workflow is general enough to accommodate large changes to input files for features that are not sampled.

The second change to the Crystalline Reference Case UA Nested Workflow structure was the automatic inclusion of graph metrics (see Section 3.3) for the DFNs in the workflow results. Updates were made to the DFN generation process to automatically compute and aggregate the graph metrics. These DFN generation changes were made externally to the workflow structure, but subsequent minor changes in the workflow enabled their incorporation in the final results output. The path to the file containing aggregated graph metrics was added to the workflow, and they were pulled in and post-processed along with the simulation results at the aleatory/spatial sampling level which includes them in the final output file that is generated by the workflow. Additionally, the graph metric density fits discussed in Section 3.3 are pulled from their original path into the location of the final results. These additions are outlined in red in Figure 5-3.



**Figure 5-3. Addition of graph metrics to the Aleatory/Spatial Sampling Workflow, outlined in red.**

The remaining changes made to the Crystalline Reference Case UA Nested Workflow involved updates to the Python post-processing scripts to accommodate changes to the analysis QoIs, as described in Section 4.2.7. The majority of these changes were made to the main post-processing script that is executed following each single deterministic PFLOTRAN simulation in the HPC Submission Workflow. Following the selection of new QoIs, the PFLOTRAN output files read by this post-processing script were updated accordingly and additional calculations were made to output the final QoIs. Additional minor changes were made to propagate this update through the nested workflow structure. Beyond the modification of QoIs, post-processing script changes were

made to harden the analysis workflows across Python versions and user settings. These updates are described in Section 5.1.3.

### **5.1.3 Updates to Improve Crystalline Reference Case UA Nested Workflow Reliability**

Following the initial development and testing of the Crystalline Reference Case UA Nested Workflow structure, issues pertaining to software version differences across systems and users began to surface. Although the workflow delivery software remained the same across users, differences across Python modules including versions, available supporting packages and their versions, and user settings made running the workflow to completion by different analysts difficult. Changes were made to the workflow elements and to the testing of post-processing scripts to mitigate these issues. First, user environment variable settings were examined to ensure that these were not able to impact workflow runs by inadvertently loading incompatible versions of Python or Python packages. Second, changes were made to the post-processing scripts to remove functionalities that were not backward compatible. This included modifying the file types utilized in the transfer of data from the HPC Submission Workflow to the Local Resource Deterministic Workflow and through the Epistemic and Aleatory/Spatial Sampling Workflows to ensure backward compatibility. Finally, testing guidelines were implemented such that changes made to a script were tested on the same system (local or HPC), and with the same version of Python, as the workflow elements that would call them. With these changes in place, Python nodes within all workflow elements began to function reliably across users and systems.

The Crystalline Reference Case UA Nested Workflow structure development occurs in parallel to active model development. Because of this, the workflow must be able to automatically accommodate changes such as updated QoIs. Without this flexibility, the workflow would quickly become outdated, barring continuous updates to reflect the needs of analysts interpreting simulation results. The original iteration of the workflow included many hard-coded QoIs, which were not of interest in the FY21 Crystalline Reference Case UA. To minimize the overhead of such changes in future, the hard-coded QoIs were removed, leaving a single response of interest, based on the total  $^{129}\text{I}$ , which is returned across workflow elements to Dakota as an indication that they have successfully run. QoIs are now only modified in one place in the workflow: in the main post-processing script. This allows all workflow elements to be flexible to changes in the selected QoIs and mitigates potential bugs from needing to implement model changes in multiple places in the workflow.

## **5.2 Crystalline Reference Case UA Nested Workflow Application**

The Crystalline Reference Case UA Nested Workflow structure was utilized to run two full sets of UQ studies in FY21. This use was a significant stretch beyond the initial testing that the workflow elements had undergone during the development process and allowed the initial identification of minor fixes that were applied to allow the workflow to complete without analyst intervention. The first UQ study utilizing the analysis workflow structure took place in Q1 of FY21 and studied the same 20 DFNs at the aleatory/spatial sampling level as were studied in the FY20 report, with the same number of 40 epistemic samples. This study included a minor update to the PFLOTTRAN input file from the FY20 model to add integral flux calculations to the simulations.

Comparisons across this study and the FY20 study allowed the analysis team to assess the impact of including this aspect of the model in the UQ study. The workflow was able to complete successfully but encountered several issues. Seemingly random failures of sample realizations or HPC outages caused the progress of all runs to stop, requiring a restart that needed some analyst intervention to succeed. To mitigate this issue in subsequent studies, the script in the workflow that submitted HPC jobs to the queue and periodically checked their progress was updated to include checks for random failures and HPC outages. Next, the concurrency scheme for the nested sampling structure was updated to recommend beginning all samples from the Aleatory/Spatial Sampling Workflow (each representing a different DFN) at the same time with two concurrent epistemic runs submitted to the queue at a time to ensure each Aleatory/Spatial sample was able to run to completion for all epistemic samples in parallel even if another Aleatory/Spatial sample experienced a failure. Previously, the Aleatory/Spatial samples ran in series, meaning that a failure in one Aleatory/Spatial sample's epistemic realizations stopped progress for all remaining Aleatory/Spatial samples. The updated concurrency scheme allows the study to efficiently utilize a feature in the NGW software which skips over runs that have completed if all expected output files are available for transfer between workflow nodes, making restarts much easier. Finally, the handoff of results from the workflow to the subsequent sensitivity analysis execution was refined to ensure that variable naming, file structure, and data organization were set up to work across both the workflow and the scripts used to execute the sensitivity analyses.

The second UQ study completed in FY21 using the Crystalline Reference Case UA Nested Workflow structure involved more extensive changes to the PFLOTRAN model and workflow elements, as described in Sections 5.1.2 and 5.1.3. This study was completed collaboratively across four analysts with a new analyst executing the workflow, where previously the workflow developer alone had completed the workflow execution end-to-end. The transfer of workflows across users posed challenges that were mitigated by workflow element changes, as described in Section 5.1.3. In preparation for workflow execution, tests were run across systems and users to ensure workflow stability. As issues were found, changes were made to scripts and other workflow elements by the team of analysts. These changes were saved and transferred across users using git. This allowed for changes to be made efficiently under version control and represented an important benefit of using the workflow along with git to complete the analysis as setting up the workflow could happen both collaboratively and efficiently. The final execution of the workflow utilized the updated concurrency structure described above. Screenshots of the aleatory user inputs and the HPC user inputs used in the production run are shown in Figure 5-4 and Figure 5-5, respectively. The inputs entered here were then used as variables in the HPC template file. The number of nodes, account number, and job name replace the text placeholders in the final submission script according to user-selected parameters. Although some restarts were required due to HPC outages, the workflow was able to handle these issues with ease and the analysis completed successfully without manual intervention beyond the user starting the workflow again. This is a major improvement from the previous manual workflow employed for the Crystalline Reference Case, which required significant user intervention to restart runs interrupted by HPC outages. The final results for this study are presented in Section 4.3.

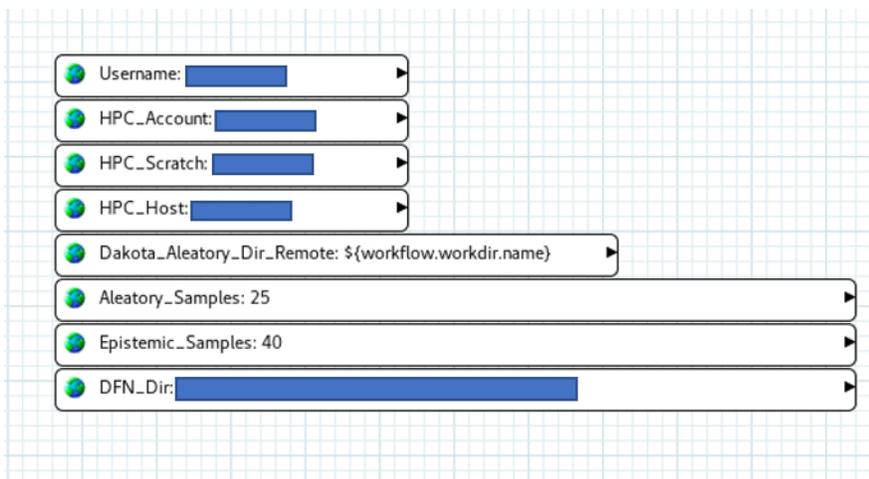


Figure 5-4. Screenshot of aleatory input parameters

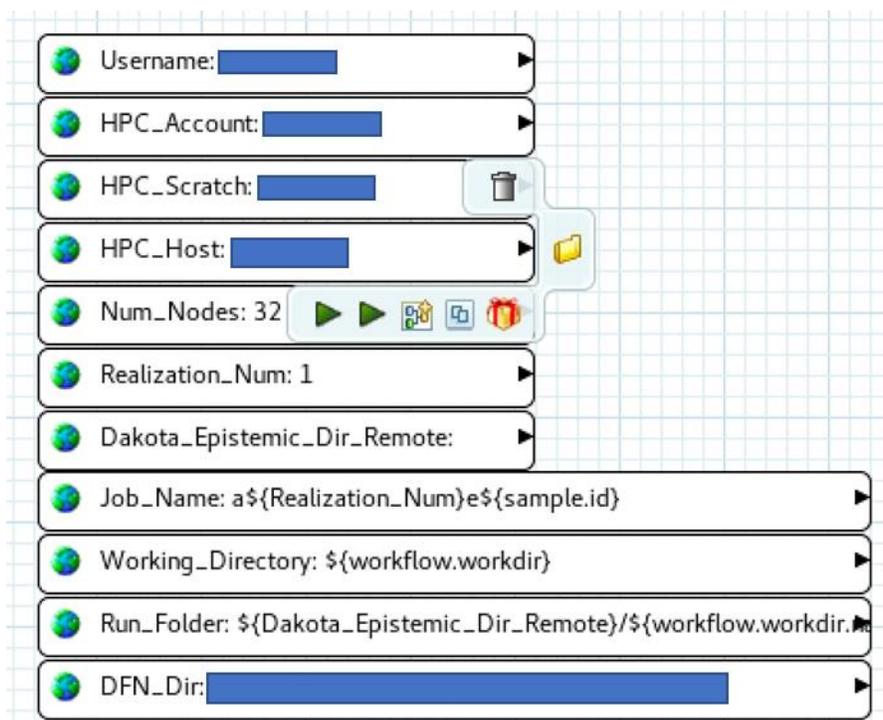


Figure 5-5. Screenshot of HPC user inputs

### 5.3 Outcomes of Crystalline Reference Case UA Nested Workflow Use

The development and application of the Crystalline Reference Case UA Nested Workflow progressed significantly in FY21. The work performed allowed the workflows to move from a proof-of-concept paradigm to a production-supporting capability over the course of the year. The team was able to perform two, rather than one, full UQ studies over the course of the fiscal year

and the final production run results were delivered much earlier to facilitate a longer period for analysis and results documentation. In addition to this, the workflow capability fully documented the end-to-end analysis process in a package that will be included in the GDSA calculation archive. This archive would allow another team to pick up and reproduce the analysis at a later date if needed, which could not have been achieved without the workflow structure in place.

The Crystalline Reference Case UA Nested Workflow structure, coupled to the use of git for version control and file sharing, enabled a significant increase in the collaborative teamwork supporting the final production runs, as mentioned above. In real time, team members were able to make changes to workflow elements, track these changes in git, and share them with others. This allowed the team to finalize preparations for the production runs much more quickly than the previous manual analysis methodology, which was completed by a single analyst. Additionally, testing of changes and additions to the workflow across users and systems ensured that the workflow would successfully execute the full production run study with 1000 simulations.

The Crystalline Reference Case UA Nested Workflow structure proved to be robust against HPC outages during the final production runs. The previous manual analysis methodology required significant intervention to overcome unplanned HPC outages that occurred during runs, setting back the timeline for results delivery. Instead, the workflow was able to run through and skip simulations that had completed successfully and restart simulations that had not once the HPC resource became available again. This represented a huge improvement in the efficiency of completing the final production runs.

An introductory training course for the NGW software and analysis workflows developed using this capability to support GDSA was developed in FY21 and was presented to GDSA code developers, analysts, and program management on March 8, 2021. This training covered the basic elements of the NGW software and included hand-on demonstrations to introduce attendees to using the software through the Dakota GUI. Finally, the training included a walk-through demonstration of the Crystalline Reference Case UA Nested Workflow structure described in Section 5.1.1. This training increased the awareness of the NGW capability across the GDSA team and promoted the use of this capability for future analyses beyond the Crystalline Reference Case UA. The training materials for this course are available upon request.

The workflow capabilities demonstrated for the Crystalline Reference Case will be expanded to additional studies and applications under the GDSA program in future work due to the successful application of these capabilities for the uncertainty studies presented in this report. In addition to extending this capability to new applications, documentation including instructions on setting up and running workflows will be developed to support the adoption of the workflow by new analysts. This will include recommendations for UQ studies and could leverage faster-running examples as study problems.

## **5.4 References: Chapter 5**

1. Orient, G., Clay, R., Friedman-Hill, E., & Hoffman, E., 2020. Next Generation Workflow – an Open Source Software System for Computational Modeling to Support Agile VVUQ. Presented at the ASME V&V Symposium, May 20-22, 2020.

2. Ridgway, E., 2020. Dakota. GUI Version 6.12 User Manual – Next-Gen Workflow. <https://dakota.sandia.gov/content/next-gen-workflow>. Accessed August 5, 2020.
3. Mariner, P. E., M.A. Nole, E. Basurto, T.M. Berg, K.W. Chang, B.J. Debusschere, A.C. Eckert, M.S. Ebeida, M. Gross, G.E. Hammond, J. Harvey, S.H. Jordan, K.L. Kuhlman, T.C. LaForce, R.C. Leone, W.C. McLendon III, M.M. Mills, H.D. Park, F.V. Perry, A. Salazar III, D.T. Seidl, S.D. Sevougian, E.R. Stein and L.P. Swiler. “Advances in GDSA Framework Development and Process Model Integration.” Sandia National Laboratories, Albuquerque, NM (2020).

## 6. CONCLUSIONS

This report covers the main topics addressed in the GDSA UQ/SA work performed in 2021.

The report provides an overview of an exciting new uncertainty quantification method involving models at multiple levels of fidelity. We presented results demonstrating the use of these multifidelity (MF) UQ approaches in calculating Sobol' sensitivity indices. The demonstration was done on a simplified crystalline reference case that retained many of the key qualities of the full reference case. We examined the use of Polynomial Chaos Expansion (PCE) surrogates within the MF framework. We found that the multifidelity framework for sensitivity analysis performed better for some QoIs than others, depending on whether the QoIs exhibited discrete/discontinuous behavior and also whether spatial uncertainty was included. For the case with no spatial uncertainty, MF PCE was able to generate the global sensitivity indices for some performance QoIs (e.g. peak  $^{129}\text{I}$ ) at a fraction (2.5%) of the cost of calculating the sensitivity indices from a full set of high fidelity runs. Our study showed that a MF PCE which exploits a hierarchy of meshes has the potential to improve sensitivity analyses when evaluations at the finest level are limited. More work remains in the validation of the PCE models and the extension of the MF framework to the full production crystalline case.

This year, we used a recently added feature of dfnWorks when generating the discrete fracture networks used in this analysis: the correlated depth-dependent transmissivity. We examined the resulting averaged permeability fields in the ECPM and compared the impact of using correlated constant vs. correlated depth-dependent transmissivity. The transmissivity approach used had a large, statistically significant impact on the geometric mean of the permeability fields at each depth level in the repository. The two upper depth zones had much higher averaged permeability fields when using the depth-dependent transmissivity. However, the choice of transmissivity relationship did not have such a significant effect on most the quantities of interest we examined. This may be due to the small differences in the geometric mean permeability in depth zone 3 with the depth-dependent transmissivity compared with the constant correlated transmissivity: we hypothesize the QoIs relating to repository performance did not show much change given the small change in depth zone 3 permeability, where the repository is located. This is a preliminary study but points to the significance of the DFN parameterization within the full repository performance assessment calculations.

This year, we extended the sensitivity analysis of the crystalline reference case performed in FY19 and FY20. This analysis, in FY21, includes improvements to the fracture network implementation to incorporate depth-dependent transmissivity, modifications to the model domain including soil compressibility and a wider spacing of waste packages resulting in a cooler repository, and new quantities of interest, some of which better reflect the retention capabilities of the repository. These additional QoIs include median residence time of a conservative tracer in the repository, fraction of a "spike" tracer remaining in the repository over time, ratios of various water fluxes, and normalized fluxes. This year, we performed the sensitivity analyses with and without the graph metrics included as parameters, where the graph metrics are proxy variables representing the spatial variability introduced by the DFNs. We present detailed SA results both for scalar QoIs as well as time-dependent QoIs. The time-dependent sensitivity analysis, where the Sobol' sensitivity indices are calculated at many time points throughout the simulation, is a new capability added this year which involves a significant amount of computation and software

infrastructure. For each time point, a surrogate must be constructed over all the simulation results for the QoI at that time point, and the SA calculations to determine the sensitivity indices must be performed. This additional computational infrastructure to support time-dependent SA has the benefit of more clearly seeing how various parameters change their importance over time. With that information, we can better relate the sensitivity analysis results to the physics of the problem. Detailed conclusions about the effects of epistemic parameter uncertainty vs. spatial uncertainty (as measured through graph metrics extracted from the DFNs) on key QoIs such as peak  $^{129}\text{I}$  are outlined in the Chapter 4 summary, 4.4.

Another activity continued this year was the *GDSA Workflow*. This workflow couples Dakota, PFLOTRAN, and NGW (the Next Generation Workflow) to provide analysts with a capability to construct, execute, and communicate end-to-end computational simulation analysis workflows. The *GDSA Workflow* also allows greater reproducibility and traceability of the actual files and scripts used for a particular study. The *GDSA Workflow* was extended this year to work with job submission scripts on the high-performance computing platforms and to support nested sampling for the crystalline reference production case, including the automation of postprocessing that was previously done manually and with scripts. We used the *GDSA Workflow* in our 1000 PFLOTRAN production runs and continue to extend it, with an emphasis on reproducibility and usability.

Finally, we continue to engage with the international community on sensitivity analysis topics. This year, we completed a report investigating sensitivity analysis for multiple case studies with an international working group. Each participating organization separately performed sensitivity analyses on the case studies and results were compared and contrasted to evaluate the performance of different sensitivity analysis methods. This next phase of this collaboration focuses on applying the same process for more complex reference cases, including the crystalline reference case, and developing consensus recommendations on some best practices for sensitivity analysis in the context of repository performance assessment.

There is a rich legacy of UQ/SA being performed in repository assessment. Development, implementation, and demonstration of new tools and methods for uncertainty and sensitivity analysis in *GDSA Framework* will maintain leadership of the repository science community in UQ/SA methods, while also maintaining an infrastructure of proven tools. Geologic repository performance assessment in the U.S. involves coupled, multi-physics modeling at high resolution, large parameter spaces, and greater use of random (stochastic) field modeling. UQ/SA methods discussed in this report, including surrogate modeling to reduce computational expense, variance-based sensitivity analysis to quantify importance of parameter interactions in a multi-physics system, and new multi-fidelity methods, will enable analysis methods to keep pace with the increasing sophistication of the physics models. We seek to improve existing UQ/SA methods, employ new methods, and maintain an infrastructure of proven tools that can be extended to support computationally expensive analyses.

There are several possible avenues of future UQ/SA development, including better understanding and quantification of the effect of spatial heterogeneity, the use of optimization in determining uncertainty bounds, further work with multifidelity UQ approaches, more detailed examination of surrogates and machine learning methods, model form uncertainty, and density-based sensitivity analysis. Future work will involve further development of these potential topics as they align with the *GDSA Framework* and the repository performance assessment needed.