

# Sensitivity Analysis of Generic Deep Geologic Repository with Focus on Spatial Heterogeneity Induced by Stochastic Fracture Network Generation

DM Brooks<sup>a,c,\*</sup>, LP Swiler<sup>a</sup>, E Stein<sup>a</sup>, PE Mariner<sup>a</sup>, E Basurto<sup>a</sup>, T Portone<sup>a</sup>, A Eckert<sup>a</sup> and R Leone<sup>a</sup>

<sup>a</sup>Sandia National Laboratories, Albuquerque NM, 87123 USA

## ARTICLE INFO

**Keywords:**  
Nuclear Waste  
Spent Nuclear Fuel  
Performance Assessment  
Uncertainty

## ABSTRACT

Geologic Disposal Safety Assessment Framework is a state-of-the-art simulation software toolkit for probabilistic post-closure performance assessment of systems for deep geologic disposal of nuclear waste developed by the United States Department of Energy. This paper presents a generic reference case and shows how it is being used to develop and demonstrate performance assessment methods within the Geologic Disposal Safety Assessment Framework that mitigate some of the challenges posed by high uncertainty and limited computational resources. Variance-based global sensitivity analysis is applied to assess the effects of spatial heterogeneity using graph-based summary measures for scalar and time-varying quantities of interest. Behavior of the system with respect to spatial heterogeneity is further investigated using ratios of water fluxes. This analysis shows that spatial heterogeneity is a dominant uncertainty in predictions of repository performance which can be identified in global sensitivity analysis using proxy variables derived from graph descriptions of discrete fracture networks. New quantities of interest defined using water fluxes proved useful for better understanding overall system behavior.

## Nomenclature

DFN	discrete fracture network
DOE	U.S. Department of Energy
DRZ	disturbed rock zone
ECPM	equivalent continuous porous medium
GDSA	Geologic Disposal Safety Assessment
MdRT	median residence time
MTHM	metric tons of heavy metal
MTT	mean travel time
PA	performance assessment
PCE	polynomial chaos expansion
QoI	quantity of interest
SFWST	spent fuel and waste science technology
STT	shortest travel time
UA	uncertainty analysis
UQ	uncertainty quantification
WP	waste package

## 1. Introduction

The United States Department of Energy (DOE) is developing a state-of-the-art simulation software toolkit, Geologic Disposal Safety Assessment (GDSA) Framework, for probabilistic post-closure performance assessment (PA) of systems for deep geologic disposal of nuclear waste. Post-closure refers to the period after repository operations have ceased and performance relies on passive features of the site and engineered systems as well as institution controls to pre-

vent intrusion (International Atomic Energy Agency, 2011). The post-closure period can be on the order of hundreds of thousands to millions of years due to the long half lives of some of the hazardous radionuclides present in nuclear waste. This requires that PA simulations predict system behavior over much longer time scales than can be observed. The GDSA Framework couples increasingly higher fidelity models of subsystem processes into total system PA simulations. These models can include multiphase flow, reactive transport, heat conduction and convection, simple geomechanics, and radionuclide decay and ingrowth. Because of the inherent complexity and time scale, predictions from PA models are necessarily uncertain and the characterization, quantification, and analysis of uncertainty are all integral components of a full PA.

The paradigm for treatment of uncertainty in a typical PA can be described in five general steps: 1) characterization of the uncertainty space using probability distributions to describe uncertain model parameters, 2) sampling of the uncertainty space, 3) propagation of the samples through the system model, 4) quantification of uncertainty for the quantity of interest (QoI), and 5) sensitivity analysis. The first four steps could be described as uncertainty analysis (UA), and the fifth step, sensitivity analysis, is fundamental to support correct interpretation of UA results in the context of PA (Saltelli et al., 2019; Helton et al., 2012). Due to domain size and simulation time scale, the PA simulations in the GDSA Framework require high performance computers to run. This limits the number of samples from the uncertainty space that can be propagated through the system model, which can limit the effectiveness of some tools for UA and sensitivity analysis (Swiler et al., 2019, 2020, 2021).

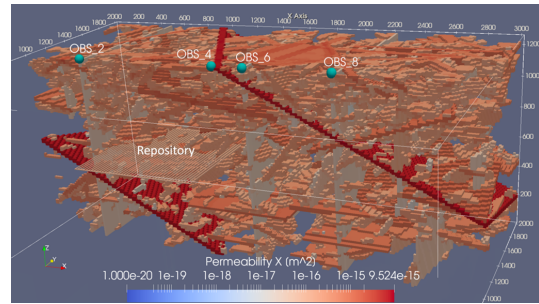
\*Corresponding author. dbrooks@sandia.gov  
ORCID(s): 0000-0002-8709-8738 (D. Brooks)

This paper presents a generic reference case, referred to as the crystalline reference case, which is not representative of a specific site. The case is meant to be illustrative for the purpose of developing and demonstrating PA methods within the GDSA Framework which mitigate some of the challenges posed by high uncertainty and limited computational resources. Limitations on the number of PA simulations highlight the need to consider methods that improve sampling efficiency and enable exploration of the full uncertainty space, even when samples of the PA simulations are relatively sparse. This paper focuses specifically on new QoIs and sensitivity analysis using variance-based sensitivity indices.

The crystalline reference case is a generic repository modeled in a crystalline rock formation. This type of repository would be placed in highly impermeable, fractured rock, where flow is assumed to be dominated by flow through the fractures. The repository model includes nuclear waste packages which breach based on an uncertain normalized general corrosion rate (Mariner et al., 2016). Transport of  $^{129}\text{I}$  is modeled from the waste packages through the repository and surrounding host rock based on additional uncertain permeability, porosity, and radionuclide release parameters. This radionuclide is a fission product in spent nuclear fuel which is of particular interest in repository performance assessment due to its abundance in spent nuclear fuel and its long half life. Concentrations of  $^{129}\text{I}$  are monitored at pre-specified points in the model domain called observation points and the peak concentration is tracked throughout the aquifer. This paper also discusses other potential QoIs that may be more illustrative for understanding specific repository characteristics that affect performance or for understanding how the model behaves. However, because nuclear waste repositories are meant to prevent hazardous radionuclides from reaching the biosphere,  $^{129}\text{I}$  concentrations are an important QoI for PA.

The crystalline reference case has an additional level of complexity that presents challenges for sensitivity analysis. The spatial heterogeneity of potential fracture flow paths through the host rock is a source of uncertainty in PA for a nuclear waste repository located in crystalline rock. Conceptually, a long-lived radionuclide released from a waste package will initially migrate through the buffer material surrounding a waste package 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 and it may eventually reach the biosphere via connected fractures and connected fracture zones. The spatial heterogeneity in the fracture network contributes to uncertainty in the QoIs used to measure system performance. With respect to the uncertainty analysis approach, inclusion of uncertainty in the fracture network can be done using a dual loop sampling structure, which is important for understanding the effects of different types of uncertainty but further complicates the challenges caused by limitations on the number of simulations that can be performed (Swiler et al., 2019).

This paper describes the crystalline reference case, dis-



**Figure 1:** Cut-away of a discrete fracture network realization mapped to porous medium grid colored by permeability, showing the full repository and the far half of the model domain. The three-dimensional structures inside the domain are the repository, five deterministic deformation zones (two of which are colored red due to their high permeability, three colored light white), the fractures of a stochastically generated fracture network (colored peach/orange throughout the model domain), and four teal observation points that were placed in the granite just below the aquifer.

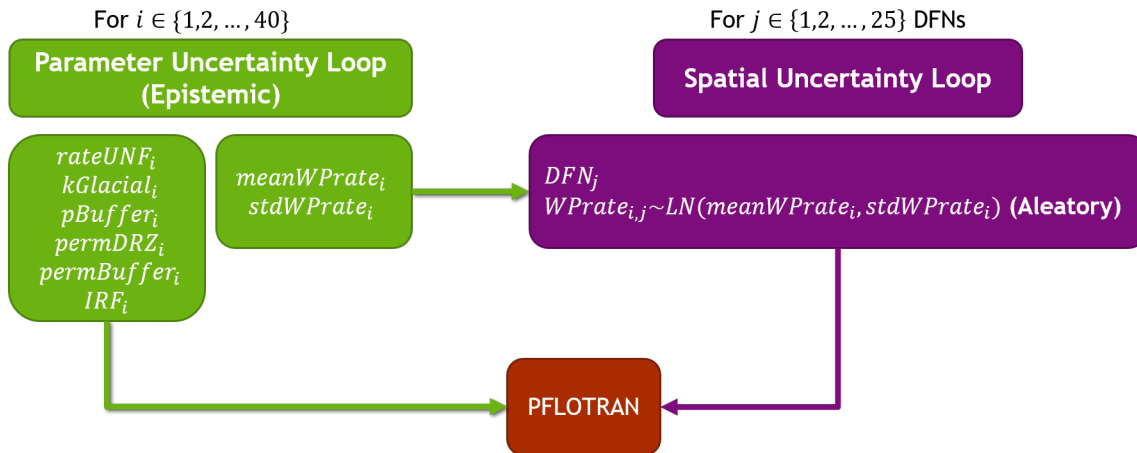
cusses the graph metrics and QoIs that were developed to better understand the effects of spatial heterogeneity, and presents sensitivity analyses that incorporate these developments.

## 2. Model Description

The crystalline reference case simulates a mined repository located at 600 m below land surface in a sparsely fractured crystalline host rock such as granite or metagranite in a stable cratonic terrain. It is assumed that a commercial repository would hold up to 70,000 metric tons of heavy metal (MTHM) of commercial used nuclear fuel, which is the maximum allowed by the Nuclear Waste Policy Act of 1982. This inventory could be accommodated in 168 disposal drifts, each 805 m in length, arranged in facing pairs on either side of a central access hallway (Wang et al., 2015). This layout has drift centers separated by 20 m, and waste packages emplaced lengthwise within the drifts with a spacing of 10 m center-to-center (5-m spacing, end-to-end). Repository access would be via vertical shafts and/or a ramp (Freeze et al., 2013b; Mariner et al., 2011; Wang et al., 2014). This reference case was developed to be a realistic example of a crystalline repository but does not correspond to any specific or proposed site, and models a smaller inventory than the statutory limit of 70,000 MTHM.

Repository performance can be affected by the structural strength of the host rock, depth of burial, host rock permeability, the geometry of fractures in the subsurface, and the chemical environment which affects corrosion and radionuclide solubility and sorption. Engineered barriers for this repository include the characteristics of the constructed repository, the waste form, waste packages, and bentonite buffer (Freeze et al., 2013b; Mariner et al., 2011, 2016).

The model domain for the crystalline reference case (Stein et al., 2017) is 3015-m in length, 2025-m in width, and 1260-m in height. Overlying the host rock is a 15-m



**Figure 2:** Diagram of the dual loop structure used to separate spatial heterogeneity from parameter uncertainty in the analysis. Each sample of the spatial uncertainty loop (purple) corresponds to 40 samples from the parameter uncertainty loop (green). There is interaction between the loops because samples from the parametric loop define the log-normal distribution on the waste package degradation rate that is sampled in the spatial loop. Each PFLOTRAN realization is defined by a parameter sample  $i \in \{1, 2, \dots, 40\}$  and a spatial sample  $j \in \{1, 2, \dots, 25\}$ .

thick overburden of glacial sediments. The repository is located at a depth of 585 m. Forty-two disposal drifts contain 80 waste packages each (3360 waste packages in total). The waste packages have a capacity of 12 pressurized water reactor fuel assemblies (Rechard and Voegelé, 2014). Each waste package has a 5.225 MTHM inventory, so the total inventory for the crystalline reference case is 17,556 MTHM (Freeze et al., 2013a). Drifts are backfilled with bentonite buffer and are surrounded by a 1.67-m thick DRZ. The domain is modeled with an unstructured grid with finer discretization in the repository. Grid cells, which consist of nodes, edges, and boundary faces, are as small as 1.67-m on their shortest edge within the repository and 15-m on their shortest edge elsewhere in the model domain. 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 in (Stein et al., 2017) available for download at <https://pa.sandia.gov>.

## 2.1. Discrete Fracture Networks

As described by Mariner et al. (2016), the representation of fractured crystalline rock in the crystalline reference case is loosely based on the well-characterized, sparsely fractured metagranite at Forsmark, Sweden (Joyce et al., 2014). At Forsmark, large-scale mappable features of concentrated brittle and/or ductile deformation (termed “deformation zones”) bound volumes of relatively undeformed rock. Each volume of relatively undeformed rock (termed a “fracture domain”) is sparsely fractured, and the fractures within each can be described in terms of a number of “fracture sets,” distinguished from each other on the basis of fracture orientation. Each fracture set within a particular fracture domain and depth zone is described using a 3-dimensional Fisher distribution to describe the orientation of fracture

poles in space, a truncated power-law distribution for fracture radii, and a fracture density,  $P_{32}$ , which is defined as the area of fractures per volume of rock ( $m^2/m^3$ ) (Hedin, 2008). For each depth zone within a fracture domain, a relationship is given between fracture radius and fracture transmissivity.

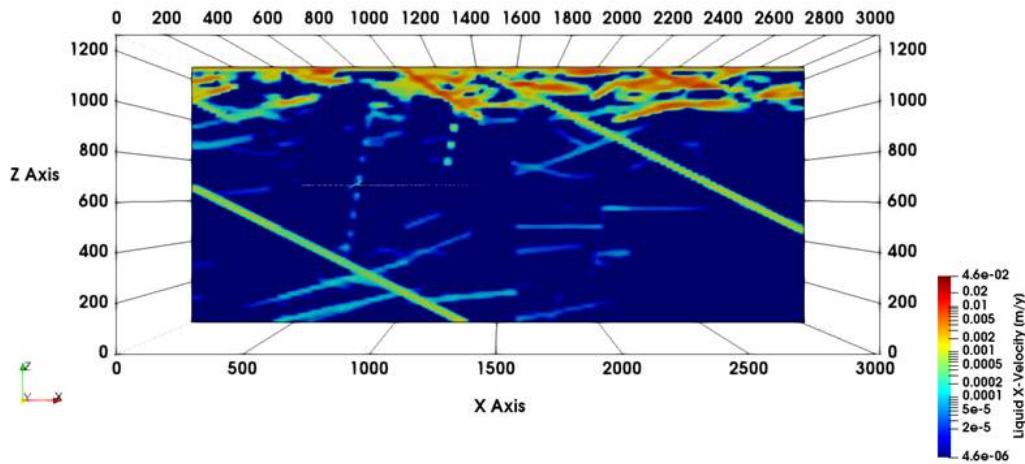
Figure 1 shows the fracture domain for the crystalline reference case. The vertical and offset planes in the figure define deterministic large mappable features (such as faults). The discrete fracture networks (DFNs) are generated stochastically in this domain using probabilistic characteristics. These characteristics are defined based on three fracture sets from Forsmark and three zone depths, resulting in nine total probabilistic fracture set descriptions in the domain. The probabilistic characteristics for each set include fracture density (which decreases over the depth zones so the deepest regions of the domain have the lowest fracture density), radius, orientation, and centroid location. The stochastic fracture set characteristics and deterministic features 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.

Simulating a repository for heat-generating nuclear waste in sparsely fractured crystalline rock requires methods to simulate coupled heat and fluid flow and reactive radionuclide transport in both porous media (e.g., engineered backfill, sedimentary aquifer) and the sparsely fractured, otherwise impermeable, rock (Stein et al., 2017). The DFN must be coupled to the continuum to capture the effects of heat conduction through the rock. For the purposes of probabilistic performance assessment, where thousands of simulations are necessary, a computationally efficient representation of the DFN is needed. To this end, the DFN is up-scaled to an equivalent continuous porous medium (ECPM) using the procedure detailed in Stein et al. (2017). In each grid cell anisotropic permeability, kinematic porosity, and tortuosity are calculated using the outputs of fracture per-

Depth (meters below sea level) / Fracture Set Name	Orientation: Fisher Distribution for Poles			Size: Truncated Power Law for Radii			Fracture Intensity ( $P_{32}$ ) for $r \in [0.038, 564]$ m [m <sup>2</sup> /m <sup>3</sup> ]
	Mean Trend	Mean Plunge	$\kappa$	Power Law $k$	Min Radius $r_0$ [m]	Max Radius $r_{upper}$ [m]	
0-200 / NS	90°	0°	22	2.5	30	500	0.073
0-200 / NE	180°	0°	22	2.7	30	500	0.319
0-200 / HZ	360°	90°	10	2.4	30	500	0.543
200-400 / NS	90°	0°	22	2.5	30	500	0.142
200-400 / NE	180°	0°	22	2.7	30	500	0.345
200-400 / HZ	360°	90°	10	2.4	30	500	0.316
> 400 / NS	90°	0°	22	2.5	30	500	0.094
> 400 / NE	180°	0°	22	2.7	30	500	0.163
> 400 / HZ	360°	90°	10	2.4	30	500	0.141

**Table 1**

Fracture distributions parameters for the crystalline reference case.

**Figure 3:** Cutaway (y-normal) of the streamwise velocity field for one realization, showing that flow throughout the domain is predominantly along the fractures in the DFNs.

meabilities, apertures, orientation (normal vectors), radii, and location from the DFN model. The benefits of this approach include the ability to model heat conduction, computational efficiency for large domains (Stein et al. (2017) reports on average 50x fewer degrees of freedom for ECPM vs DFN meshes, and 6x speedup in simulation times), and trivial coupling to porous medium materials (Stein et al., 2017; Sweeney et al., 2020; Hadgu et al., 2017). Other methods of upscaling a sparse fracture network to a heterogeneous ECPM include those described by Jackson et al. (2000), Svensson (2001), and Sweeney et al. (2020). Upscaling assumptions can lead to differences in behavior between the DFN and ECPM representations (Azizmohammadi and Matthäi, 2017; Bisdorn et al., 2017; Sweeney et al., 2020; Hadgu et al., 2017). Given the DFN properties and upscaling discretization used here, the ECPM representations have slightly lower bulk permeabilities (calculated for a 1-km<sup>3</sup> domain) than the original DFNs and tracer breakthrough times are delayed (Stein et al., 2017). This is likely

due to the increased path length in the ECPM representations (Parashar and Reeves, 2011; Sweeney et al., 2020); a correction following Sweeney et al. (2020) will be added in future studies.

All DFNs are generated according to the same distributions for fracture radii, orientations, and centroid locations (Table 1); however, despite this, they are still significantly different from each other. This heterogeneity between DFNs is consistent with the generic nature of the reference case and our inability to completely describe the fracture network at a hypothetical fixed site. As can be seen in Figure 3, which shows a cutaway of the streamwise velocity field, flow throughout the domain is predominantly along the fractures in the DFNs; this effect on the flow makes the spatial heterogeneity a significant source of uncertainty affecting the QoIs, as discussed in the sensitivity analysis (Section 4).

## 2.2. Waste Package General Corrosion Model

The treatment of uncertainty is discussed in Section 3.1, but in this section we highlight the uncertainty characterization for the waste package general corrosion model which is handled a bit differently than the other parameter uncertainties. Most of the parameters (permeabilities, porosities, sorption parameters) are treated as epistemic uncertainties arising from lack of knowledge about the particular value to use. This is in contrast to aleatory uncertainties which represent inherent, random variability. The waste package general corrosion model is unique in the context of uncertainty and sensitivity analysis because both aleatory and epistemic uncertainty are characterized and propagated through the model. The waste package general corrosion model implemented in PFLOTRAN (Mariner et al., 2016) calculates normalized thickness of the waste package wall at each time step as a function of a base waste package general corrosion rate, a canister material constant, and temperature. Waste package breach occurs when the wall thickness reaches zero. The initial wall thickness is normalized to 1, and is reduced at each time step as a function of the effective waste package general corrosion rate  $R_{eff}$ :

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

where  $R$  ( $\text{yr}^{-1}$ ) is the general corrosion rate at  $60^\circ\text{C}$  (333.15 Kelvin),  $T$  (Kelvin) is the local temperature, and  $C$  is the canister material constant.

This general corrosion rate model and its associated uncertainty are conceptual because additional specificity is not necessary for a generic reference case. The normalized general corrosion rate is unknown (an epistemic uncertainty) but is also assumed to vary spatially within the repository, so it is sampled differently than other uncertain parameters. The spatial variability in  $R$  is assumed to follow a lognormal distribution which is itself uncertain. A lognormal distribution is used because the corrosion rate may vary multiple orders of magnitude; it was expert-defined to be consistent with the original generic reference case conception (Weck et al., 2014). In a performance assessment, however, the distribution would need to be refined to account for the materials and waste form for that specific case. The mean and standard deviation of the lognormal distribution are assigned uncertainty distributions and are sampled once for each simulation to define the lognormal distribution on  $R$  for all waste packages in that simulation. This distribution is then sampled independently for each waste package to define its general corrosion rate, which is applied uniformly to the package surface. Sampling of the distribution parameters implements the epistemic uncertainty in the general corrosion rate. Sampling each waste package independently from that distribution implements spatial variability throughout the repository. Sensitivity analysis can detect the effects of uncertainty in the distribution parameters on QoIs; however, the spatial variability is not captured or described parametrically, so it is more difficult to understand how it affects QoIs.

## 2.3. Software

The repository for the crystalline reference case is modeled using PFLOTRAN in the GDSA Framework. PFLOTRAN is an open source, state-of-the-art, massively parallel subsurface flow and reactive transport simulator (Hammond et al., 2014, 2008; Lichtner et al., 2015). Written in object-oriented Fortran 2003, PFLOTRAN is a porous medium continuum code that can model multiphase flow, reactive transport, heat conduction and convection, simple geomechanics, and radionuclide decay and ingrowth. PFLOTRAN development for GDSA Framework is described by Mariner et al. (2015, 2017, 2016) and Sevougian et al. (2018). Installation instructions are available at <https://www.pfлотran.org>.

Fractured crystalline rock is modeled using stochastic discrete fracture networks, which are two-dimensional planes distributed in the three-dimensional model domain. The fracture networks are generated using dfnWorks (Hyman et al., 2015), 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 (Stein et al., 2017).

The implementation of uncertainty in the crystalline reference case model is handled using Dakota, both to sample uncertain parameters for uncertainty propagation and to perform post-simulation sensitivity analyses. 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) (Adams et al., 2020). The algorithms allow a user to explore a computational simulation identify important parameters, the effects of uncertainties on the system, and optimize designs.

Dakota contains the uncertainty and sensitivity analysis methods commonly used in assessment of computational models, including sampling methods; uncertainty propagation; simple, rank, and partial correlation analysis; surrogate models which can be constructed from a limited number of sample runs and then queried extensively for visualization, sensitivity analysis or UQ; variance-based decomposition methods; and other approaches such as one-at-a-time methods. Dakota also contains methods to perform “nested” sampling involving separate sample loops over epistemic and aleatory uncertain parameters. This capability was used to sample the discrete fracture networks and individual waste package general corrosion rates in a separate loop from the epistemic uncertainties. Finally, Dakota has capabilities to submit runs to high-performance computers and manage concurrent evaluations of a computational model, where each model evaluation may be run in parallel on multiple cores. The Dakota software is publicly available at <https://dakota.sandia.gov>.

## 3. Analysis Methods

The crystalline reference case was studied using an uncertainty analysis framework. The focus, however, was on methods to better characterize the effects of spatial hetero-

**Table 2**  
Uncertainty distributions propagated in the crystalline reference case UA

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

generality, which were tested with global sensitivity analysis.

### 3.1. Analysis Structure

The uncertainty analysis was structured to separate spatial heterogeneity from parametric uncertainty to investigate the relative contributions to total uncertainty from each type. This separation is accomplished with a dual loop sampling structure, as shown in Figure 2. The loops are nested without repeated sampling so that there are 40 unique samples of the variables in the parameter uncertainty loop for each realization of the spatial uncertainty loop. There are 25 DFNs in the spatial uncertainty loop, thus a total of 1000 PFLOTRAN simulations were run. The intention behind sampling the parameter loop 1000 times instead of repeating the same 40 samples for each spatial realization is to obtain better coverage of the epistemic space. With a repeated-sampling scheme, there would be a total of 40 unique values for each epistemic uncertain variable over all 1000 simulations; with the sampling scheme applied for this analysis, there are 1000 unique values for each epistemic uncertain variable. Previous analyses have been performed with repeated samples (Swiler et al., 2020). The sample sizes were chosen based on computational limitations, with the epistemic loop containing more samples because it contains more independent uncertainties. The uncertain inputs to the analysis are described in Table 2.

Uncertain inputs in the parameter loop were sampled using Latin hypercube sampling. The sampled values for *meanWPrate* and *stdWPrate* from the parameter loop define a distribution on general corrosion rate for the waste packages and samples from this distribution were assigned to specific waste packages in the spatial sampling loop. To perform the analysis, samples were propagated through the deterministic repository model in PFLOTRAN, which was executed to predict QoIs up to 10<sup>6</sup> years post-closure of the repository. Simulations took around 30 minutes on average to complete on 512 processors.

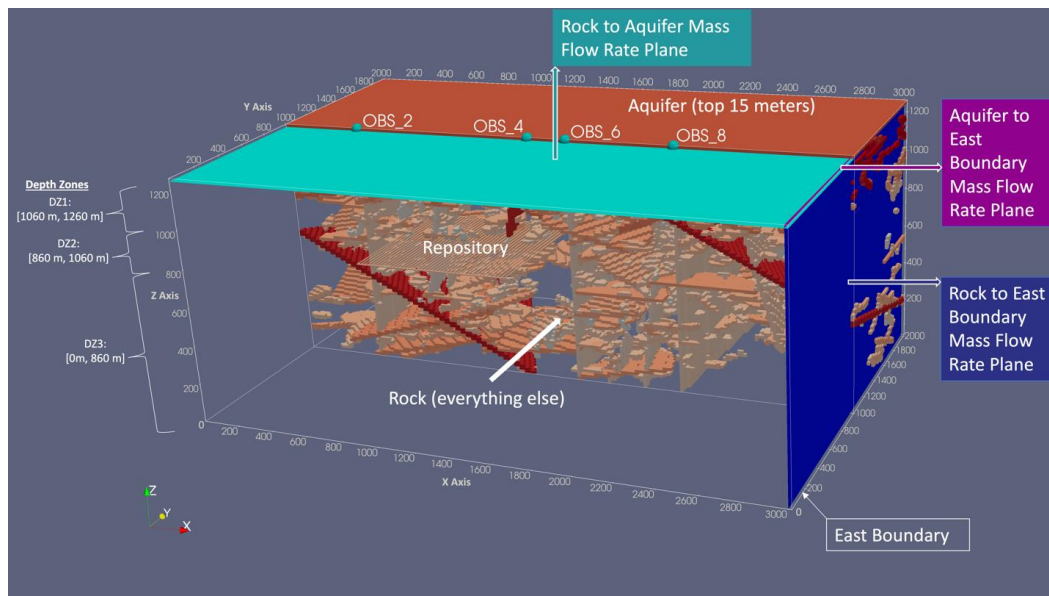
### 3.2. Graph Metrics

The spatial heterogeneity included in this UA presents a particular challenge because it is not described parametri-

cally like the other uncertainties. It is, however, clearly significant. The crystalline rock has low permeability, so connected fractures are the dominant pathways for radionuclide transport from the repository to the aquifer. In previous analysis of the reference case, the same epistemic samples were repeated for each DFN, yet the QoIs still differed by multiple orders of magnitude across DFNs, indicating a clearly important effect (Swiler et al., 2020). Summary characteristics were calculated to describe features of the DFNs in an effort to quantify at least some of their effects parametrically. We refer to these characteristics as graph metrics because of the manner in which they are calculated (from graphs) and because they attempt to measure key characteristics of the DFNs.

Graphs are a powerful tool for representing DFNs because they are able to capture the inherent network topology of the DFN without the high computational cost of meshing thousands of fractures ranging in size from millimeter to kilometer. The difference in computational cost is striking; derived quantities from a graph representation of a DFN can be computed in a matter of seconds or minutes, while a fully-resolved, 3D DFN flow and transport simulation would require several hours to run. Graphs have been used with some success as reduced-order models (with corrections) for predicting flow and transport through the DFN, and as surrogates in UQ propagation studies in this context (Viswanathan et al., 2018; Srinivasan et al., 2018). For this study, however, graphs were not applied within the simulations; they simply summarize important information about the full DFNs used in the simulations.

The graphs were constructed using *dfnWork*'s *dfnGraph* utility (Viswanathan et al., 2018; Srinivasan et al., 2018). Postprocessing was performed using *dfnWorks* and *NetworkX*, a python network analysis package (Hagberg et al., 2008), to calculate the graph metrics. There are two types of graphs that can be constructed from a DFN: a fracture graph, wherein fractures are assigned to nodes and intersections to edges, and an intersection graph, wherein intersections are assigned to nodes and fractures to edges. The type of graph used depends on what needs to be calculated. Calculations



**Figure 4:** Cut-away of one realization mapped to porous medium grid colored by permeability. The features are the same as described in Figure 1. This figure shows three flow planes corresponding to the flux  $Q_oI_s$ , and the location of three depth zones.

are performed on the graphs to estimate characteristics of the DFNs.

The graph metrics used in this analysis are *AveDegree*, *Intersections*, and *STT* (shortest travel time). *AveDegree* is the average number of intersections per fracture. This measures how connected the network is, on average, over the entire domain. *Intersections* is the number of fractures intersecting the repository. This measures the number of potential flow pathways out of the repository region. *STT* is the relative shortest travel time between the repository and the aquifer, which is calculated by scaling the shortest travel time for each DFN by the median shortest travel time over all DFNs. This is a measure of ease of flow between the repository and the aquifer.

Unlike the other two metrics, the procedure for computing the shortest travel time is applied with the intersection graph (with intersections as nodes and fracture as edges) rather than the fracture graph. For each DFN, a shortest travel time was computed using *dfnGraph*'s flow functionality in *dfnWorks*, which assigns weights to the edges of the graph based on the area and permeability of each fracture. These weights are used to estimate a flow time between input and output nodes in the graph, instantiated with inflow and outflow pressures and fluid viscosity specified by the user. The default values were used for these quantities, set to  $2 \times 10^6$  Pa,  $1 \times 10^6$  Pa and  $8.9 \times 10^4$  Pa·s respectively. There is no reason to assign any physical relevance to the absolute shortest travel time derived from this procedure, since the pressures and viscosity were assigned in an arbitrary fashion. Instead, this graph metric should be seen as a relative ranking of the speed with which fluid can move from the repository to the aquifer across DFNs.

These graph metrics were selected for inclusion in sensitivity analysis because they summarize features of the DFNs

that affect flow out of repository and through the host rock, so we expect them to capture some of the influence of the DFNs on  $Q_oI_s$ . However, this list of graph metrics cannot capture all effects from the DFNs, and research into additional graph metrics is ongoing. Though the graph metrics improve our understanding of DFN effects, their use still interferes with sensitivity analyses because of the sampling structure for the UA. The dual loop structure associates 40 realizations of the parametric uncertainty loop with each DFN; this means that there are 1000 unique samples in the parameter loop but only 25 unique DFNs in the spatial loop and thus each graph metric is repeated 40 times in the data set used for sensitivity analysis.

### 3.3. Quantities of Interest

The  $Q_oI$  for performance assessment research with the crystalline reference case has historically been the peak  $^{129}I$  concentration in the aquifer (Swiler et al., 2019, 2020). In a true performance assessment (for specific proposed site), the performance metric would likely be in the form of a dose to a member of the public. However, analysis of the crystalline reference case has not yet focused on the biosphere model, so the dose is not calculated. We treat the peak  $^{129}I$  concentration in the aquifer as a substitute performance metric for the current stage of the model because exposure to the public occurs via the aquifer. The  $^{129}I$  concentration is monitored throughout the entire aquifer at each time step. The peak  $^{129}I$  concentration is then calculated by taking the maximum of this concentration over both time and space. In other words, the peak concentration in a simulation is the highest concentration reached at any location in the aquifer at any time during the simulation.

Additional  $Q_oI_s$  have been developed to gain a better understanding of model behavior, both with respect to the repository and with respect to the host rock. They also con-

tribute to our understanding of the DFN effects. The QoIs 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 (Swiler et al., 2019; Mariner et al., 2020; Swiler et al., 2020).

Fluxes and flux ratios are the most recent additions to this list of potential QoIs. These include a fractional mass flux and ratios of water fluxes. Fluxes and flux ratios are considered at different points in time based on physical phenomena. Immediately following the closure of the repository, heat from the waste packages causes rapid expansion of water. This pressure results in a very early peak (within the first year) in water flux to the aquifer. After this initial peak, however, the surrounding rock absorbs heat from the waste packages more slowly and subsequently heats water entering the domain from the west. This water expands, affecting flow rates and directions. This is referred to as the thermal pulse. QoIs are evaluated at three thousand years to examine behavior during this pulse. Once the waste packages cool, the surrounding rock and water subsequently cool, and flow rates and directions may change again. QoIs are evaluated at one million years to understand behavior during these relatively undisturbed conditions.

Water flux ratios are indicators of the natural barrier system performance. This analysis focused on two such flux ratios. 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 three thousand 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 three thousand years and one million years. The flow planes corresponding to the fluxes in these ratios are illustrated in Figure 4. 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.

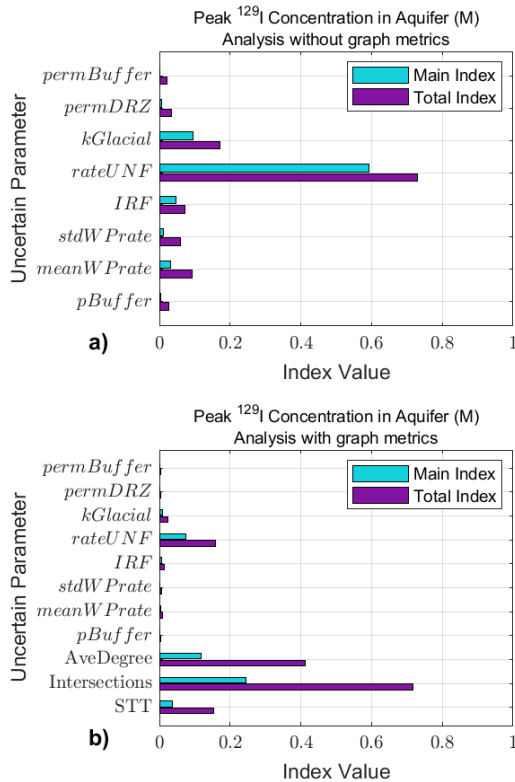
The scalar QoIs discussed in Section 4 are Peak  $^{129}\text{I}$  concentration [M], the ratio between the aquifer to east boundary flux and the rock to east boundary flux at 1 million years, and the ratio between the rock to aquifer flux and rock to east boundary flux at 1 million years. The time-dependent QoIs discussed in 4 are the maximum  $^{129}\text{I}$  concentration in the aquifer, the flux from the rock to the aquifer, the rock to east boundary flux normalized by the flux at 1 million years, and the aquifer to east boundary water normalized by the flux at 1 million years.

### 3.4. Sensitivity Analysis Methods

This analysis focused on exploration of new QoIs and their utility for increasing understanding of the repository and its performance, as well as performance of the model. Sensitivity analysis was performed using various surrogate models to estimate main and total effect Sobol' indices using Dakota (Adams et al., 2020; Swiler et al., 2019, 2020). Sobol' indices are variance-based indices. The main effect index indicates the fraction of the total response variance that can be attributed to a particular parameter alone. For example, if a parameter has a main effect index of 0.3 with respect to a response, that means that 30% of the variance in the response can be attributed to the effect of that parameter's variance. The total effect index indicates the combined or total effect of a variable and its interaction with other variables on the variance of the response. The main effect indices should sum to one if there are no interaction effects; in the presence of interaction effects, their sum will be less than one. The total effect indices are bounded below by one because they include interactions in addition to the main effects. If the main and total effect indices are the same value for a given parameter, this indicates the parameter does not have significant interaction effects on the response with respect to other parameters (Sobol', 1993; Homma and Saltelli, 1996).

The calculation of these sensitivity indices is a costly computation, requiring tens of thousands of model evaluations. Because this is not feasible with this model, the 1000 sample realizations were used to generate two types of surrogate models that were then used to perform the Sobol' variance-based decomposition. The two surrogates used were polynomial regression (a second-order quadratic regression) and polynomial chaos expansion (PCE). PCE presents a computational advantage since Sobol' indices may be calculated directly from the expansion coefficients, whereas sampling is required to estimate Sobol' indices with the quadratic regression (Sudret, 2008; Xiu, 2010; Ghanem et al., 2017; Ghanem and Red-Horse, 2017; Le Gratiet et al., 2017). Sensitivity results from the PCE and quadratic regression models were almost equal, so only the PCE results are presented for simplicity.

One of the underlying assumptions for the surrogate models is independence between the input variables. This is clearly true for the parametric uncertainties since they are independently sampled. However, it is not necessarily true for the graph metrics. Linear correlation coefficients were calculated between the graph metrics and showed that there are no significant correlations between the metrics for these simulations; the highest correlation was 0.09 (p-value 0.66). However, due to the small sample size (25), it is not clear if the lack of correlation could be by random chance. Future analyses should always include an assessment of graph metric correlations. Graph metrics also differ from the parametric uncertainties because the graph metrics are not sampled from a distribution and distributions are required for all input variables in Dakota. Continuous distributions were fit to the graph metrics so Dakota could include them in the PCE



**Figure 5:** Sensitivity analysis results without (a) and with graph (b) metrics for peak  $^{129}\text{I}$  concentration in the aquifer. Larger main and total effect indices identify *rateUNF*, *kGlacial*, and all of the graph metrics as important drivers of variation in the peak  $^{129}\text{I}$  concentration.

surrogates.

The surrogates were applied twice, once without including the graph metrics as inputs and a second time with the graph metrics as inputs. 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 in the DFN. The two analyses are compared to evaluate the influence of spatial uncertainties and the effectiveness of spatial uncertainty summary measures. 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 40 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. When the graph metrics are not included in the analysis, this may overemphasize the parametric uncertainties since they are the only attributable sources of variation. The true relative contributions from the two sources of uncertainty are likely somewhere between the two biased sensitivity results.

The surrogate model methods and the calculation of the Sobol' indices are discussed extensively in (Swiler et al., 2019).

## 4. Sensitivity Analysis Results

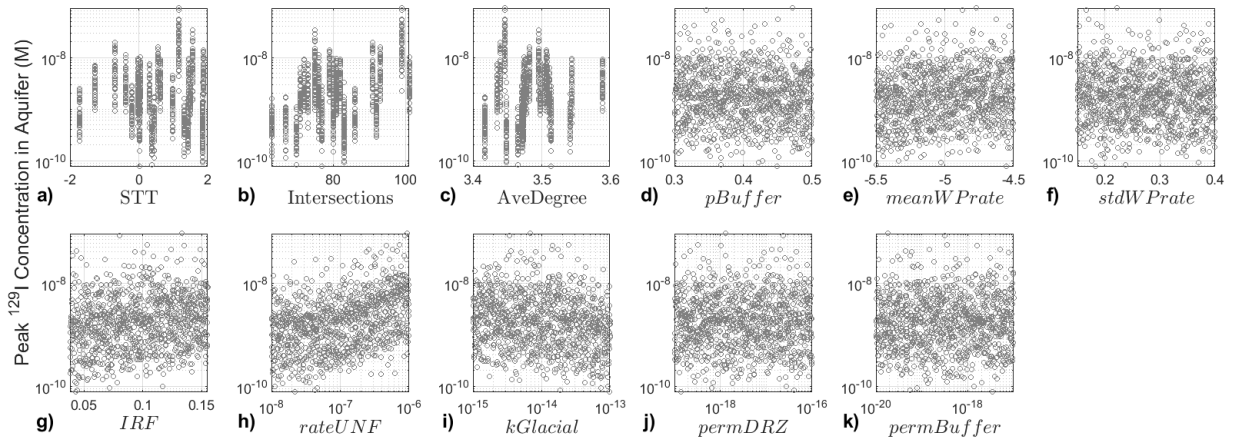
Sensitivity analysis results are presented for second-order PCE surrogate models; Sobol' indices from quadratic regression were nearly equal to those from PCE so they are omitted. Analyses were also performed with multivariate adaptive regression splines and Gaussian process models, but these surrogate models performed poorly for some QoIs. Though it would be valuable to study why some surrogates performed better than others and how to improve surrogate performance, that was not the focus of this study. First, we discuss results for scalar QoIs where the sensitivity analysis was performed either for the QoI at a specified time in the simulation, or for a QoI that is aggregated over time. Then we discuss results for QoIs that are tracked over time, for which the sensitivity analysis was repeated at each time step. Results are presented for particularly interesting QoIs; results for additional QoIs are discussed in (Swiler et al., 2021).

### 4.1. Scalar QoI Results

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 two plots, which contain the main and total Sobol' index values for the sensitivity analyses that were performed with and without the graph metrics. In this analysis, the PCE surrogate is second order, so the total index for a parameter includes its main effect and effects from two-way interactions. When the total index is significantly greater than the main index, this indicates significant interactions.

The sensitivity analysis results for peak  $^{129}\text{I}$  concentration are plotted in Figure 5. Figure 5a shows the results when the sensitivity analysis excludes graph metrics and thus cannot attribute variance to the spatial uncertainty. Figure 5b shows the results when the sensitivity analysis includes graph metrics and thus variance can be attributed to the spatial uncertainty. In the top plot, which does 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 interactions. This does not necessarily mean that *kGlacial* and *rateUNF* have a significant interaction with each other. Rather, it means that all of their interactions with the other parameters, when combined, are significant.

The analysis in Figure 5b includes three graph metrics to represent spatial heterogeneity. With these metrics included, the indices for *kGlacial* and *rateUNF* are substantially reduced. 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. The peak  $^{129}\text{I}$  concentration is plotted against the parameters and graph metrics in Figure 6. The trend between peak  $^{129}\text{I}$  concentration and *rateUNF* (h) is distinct and there are apparent



**Figure 6:** Peak  $^{129}\text{I}$  concentration in the aquifer plotted against parametric uncertainties and graph metrics. The scatterplots support the importance of the graph metrics (a,b,c),  $rateUNF$  (h), and  $kGlacial$  (i) on the peak  $^{129}\text{I}$  concentration with less pronounced effects from  $IRF$  (g) and  $meanWPrate$  (e).

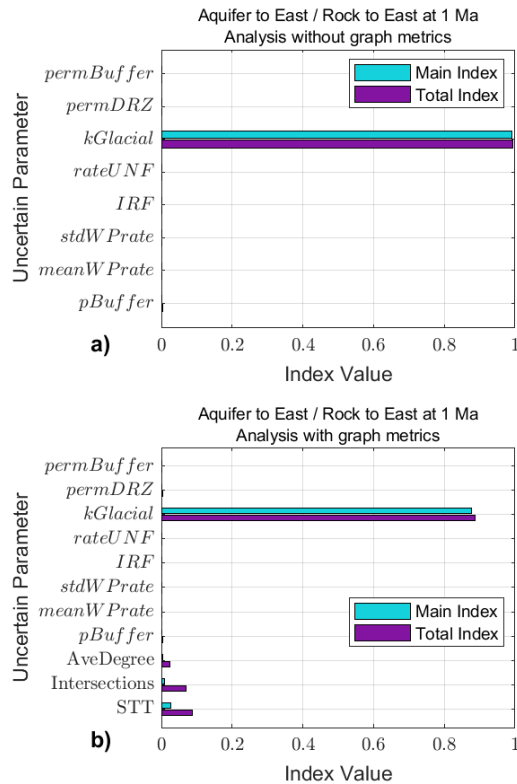
minor trends with respect to  $kGlacial$ (i) and  $IRF$ (g). Additionally, the scatter plots show trends between the graph metrics (a,b,c) and peak  $^{129}\text{I}$  concentration that correspond to the high main indices for the graph metrics. The importance of the graph metrics makes sense because 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.

Both analyses were performed on the same simulations, which included spatial heterogeneity, with the only difference being whether the surrogate models used to estimate the sensitivity indices included a measure of spatial heterogeneity. In the absence of graph metrics, the surrogates estimate  $rateUNF$  contributes to over 70% of the variance in  $^{129}\text{I}$ . Though  $rateUNF$  is clearly important, the analysis including graph metrics shows that this is an overestimate of the parameter's influence. This demonstrates that surrogate models can provide inaccurate results and may misattribute variance due to spatial heterogeneity when measures of the spatial heterogeneity are not included in surrogate model construction.

Results for the ratio between the aquifer-to-east-boundary flux and the rock-to-east-boundary flux are plotted in Figure 7. These results are for the flux ratio at 1 Ma (end of simulation). There is little difference between the analysis that excluded graph metrics (a) and the analysis that included them (b) for this QoI. This makes sense because the trend with  $kGlacial$  in Figure 8 is strong and obvious trends are not apparent for any of the graph metrics. This effect from  $kGlacial$  is expected; increasing permeability enables increased flow in the aquifer, and thus increases flux from the aquifer to the east boundary and from the rock into the aquifer. 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. However, the rock-to-east-boundary flux depends on the bulk properties of the fractured rock domain. The method used to generate DFNs is intended to incorporate spatial heterogeneity while maintaining these bulk properties, so the lack of significant sensitivity to the graph metrics for this QoI may indicate success in maintaining bulk properties across DFNs. This example may motivate exploration of additional graph metrics in the future; the current graph metrics clearly improve understanding for some QoIs but are not sufficient to fully characterize the effects of spatial heterogeneity.

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 9 are an interesting case because almost all of the variance is attributed to interaction effects when the graph metrics are not included in the analysis (a). This is an instance of overfitting, which can be seen by looking at the scatterplots in Figure 10. Plots a,b, and c 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 attributes that variance to the parameters. Since none of those parameters have a significant effect on their own, the variance is attributed to interactions. There are sufficiently many parameters compared to the total number of simulations that the PCE surrogate model can explain much of the variance via two-way interactions, but these results are spurious. The



**Figure 7:** Sensitivity analysis results without (a) and with (b) graph metrics for the ratio between the aquifer-to-east-boundary water flux and the rock-to-east-boundary water flux. High main and total effect indices in both analyses identify *kGlacial* as the most influential parameter on uncertainty in this flux ratio.

conclusion of over-fitting is also intuitive because it does not make sense phenomenologically for *rateUNF*, *IRF*, *stdWPrate*, or *meanWPrate* to affect these fluxes. These parameters only affect the waste packages and waste form, not the water flux in the surrounding rock.

This example demonstrates the necessity of interrogating sensitivity analysis results critically to ensure that conclusions make phenomenological sense and reflect patterns that truly exist in the data. This type of analysis becomes much more challenging when higher order interaction effects are included in surrogate models than can be readily visualized.

#### 4.2. Time-Dependent QoI Results

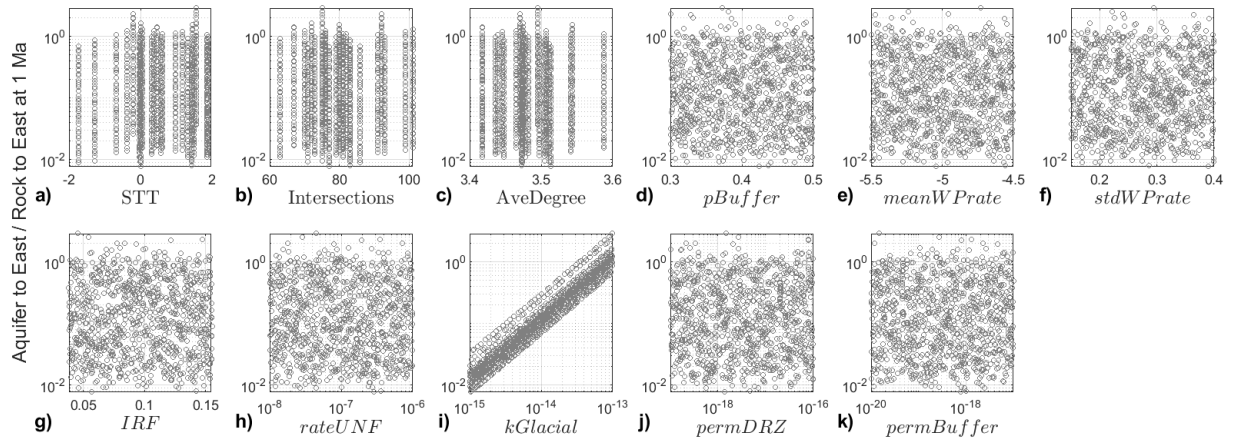
This section presents sensitivity analysis results for QoIs that are tracked over time. The analyses were performed precisely as was done with the scalar QoIs except that the process of fitting the surrogate model and estimating Sobol' indices was repeated at each time step. This means that the QoI value at each timestep was treated as an individual QoI and sensitivity results are plotted over time to demonstrate how sensitivity changes. We note that the time-dependent Sobol' sensitivity indices require significant computation: for each QoI, a surrogate must be constructed at each time point, then

used in the sensitivity index calculations. This is a one-time cost that is part of the postprocessing analysis. The benefit is that it allows one to more clearly see how the importance of various parameters changes over time. With that information, we can better relate the sensitivity analysis results to the physics of the problem. 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.

Sensitivity results over time for maximum  $^{129}\text{I}$  are plotted in Figure 11. Figure 11a and Figure 11b show the main and total index values over time for the surrogates that do not account for spatial heterogeneity. Figure 11c and Figure 11d 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 results. Figure 11e shows the time history of the QoI and Figure 11f shows the variance between the realizations over time. This type of plot can be helpful for understanding seemingly aberrant behavior in 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. Figure 11g shows the temperature time histories for the centermost waste package in the repository, observation point 4 (see Figure 1), and the point in the DRZ that is directly above the centermost waste package in the repository. These temperatures are so similar across realizations that the curves are nearly indistinguishable. 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 1. Note that, due to differences in scale, there are two separate vertical axes in Figure 11g. 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.

When the graph metrics are not included (Figure 11a and Figure 11b), the parameters driving waste package corrosion (*meanWPrate*, *stdWPrate*) dominate, especially in the period of  $10^3$  to  $10^5$  years. 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 from the scalar QoI analysis that the sensitivity results for peak  $^{129}\text{I}$  concentration (see Figure 5) 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 11. 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 (Figure 11g) 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



**Figure 8:** Ratio between the aquifer-to-east-boundary flux and the rock-to-east-boundary flux plotted against parametric uncertainties and graph metrics. The strong positive trend with respect to  $kGlacial$  (i) supports the sensitivity analysis results with none of the other scatterplots demonstrating a significant trend.

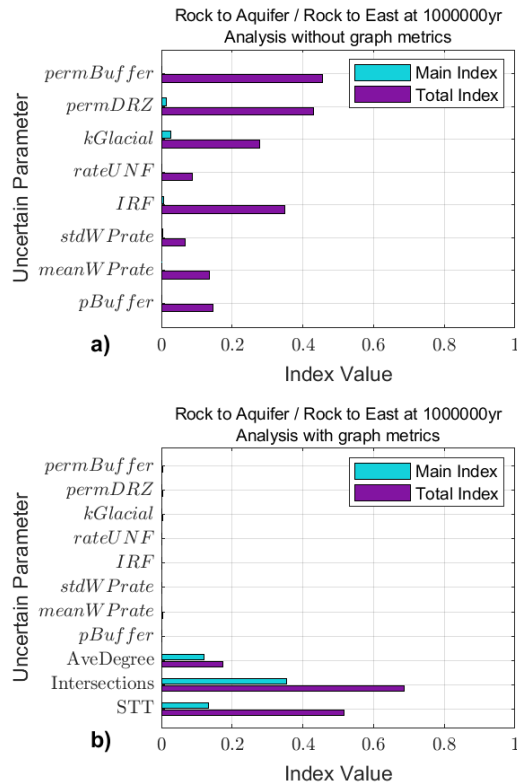
with scatter plots at individual time points. Figure 12 shows the maximum  $^{129}\text{I}$  concentration at  $10^5$  and  $10^6$  years versus  $meanWPrate$  (a,c) and the number of intersections (b,d). The trend with  $meanWPrate$  is strong up at  $10^5$  years, but weaker at the end of the simulation, whereas the trend with the number of intersections is strongest at the end of the simulation. The increase of importance for  $rateUNF$  at the end of the simulation makes sense because by that time the waste packages have failed so waste package degradation does not control radionuclide release; waste form degradation controls it.  $kGlacial$  affects peak  $^{129}\text{I}$  because high permeability causes rapid flow in the aquifer and therefore higher dilution. When average degree and intersections increase this creates more fracture connections between the repository and other fractures which can affect upward flow.

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. This makes sense because radionuclides enter the repository when the waste form degrades and the engineered barrier fails. When the engineered barrier (the waste package) is intact, it prevents radionuclides from the waste form entering the repository. But once the engineered barrier breaches, any degradation of the waste form releases radionuclides into the repository. The time-dependent Sobol' index results are useful for understanding these types of changes over time without having to generate and examine hundreds of scatter plots.

Figure 13 shows the global sensitivity analysis results for the water flux from the rock to the aquifer over time. The total effects for the parameter uncertainties are not significant when the graph metrics are included (d), and only  $kGlacial$  appears to indicate a potentially significant effect when the graph metrics are excluded (b). This, combined with the low main effects for all of the other uncertain parameters in the analysis without graph metrics (a), suggests that the dominant uncertainty affecting water flux from the rock to

the aquifer is spatial heterogeneity. This also explains why this flux (e) and the corresponding sensitivity analysis results vary less over time than some other QoIs; the spatial heterogeneity is fixed with respect to time. 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 (c,d), 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 (d) than main effect Sobol' indices (c), 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 (a). 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 (e), as expected. The relatively high total effects for some parameters in the analysis without graph metrics (b) may be the result of over-fitting; none of the parameter uncertainties can explain the variance, but there are enough parameters for the surrogate to explain some of it via interactions.

Global sensitivity results for the flux from the rock to the east boundary normalized by the flux at 1 Ma are plotted in Figure 14. This is another new QoI and, conceptually, the higher this ratio is, the higher the flux from the rock to east boundary. Only the total effects indices are plotted for this QoI. When the graph metrics are not included in the surrogate model (a), the Sobol' indices attribute most of the variance to  $kGlacial$ , which was expected. With graph metrics included in the surrogate model (b),  $kGlacial$  is still significant, but the shortest travel time and number of intersections



**Figure 9:** Sensitivity analysis results without (a) and with (b) graph metrics for the ratio between the rock-to-aquifer water flux and the rock-to-east-boundary water flux. High total effects indices with low main effects indices in the analysis without graph metrics (a) suggest overfitting in the surrogate model used to estimate the indices. Large main and total effects indices for the graph metrics (b) show the dominance of spatial heterogeneity on this flux ratio.

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 (b,c). 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).

Sensitivity results over time for the flux from the aquifer to the east boundary normalized by the flux at 1 Ma are plotted in Figure 15. Only the total effects indices are plotted for this QoI. The sharp drop in importance for *kGlacial* just before 1000 years (a,b) seems unreasonable. However, the time series plot for this QoI (c) shows that this drop corresponds to when the water flux increases and the variance drops to nearly zero. 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.

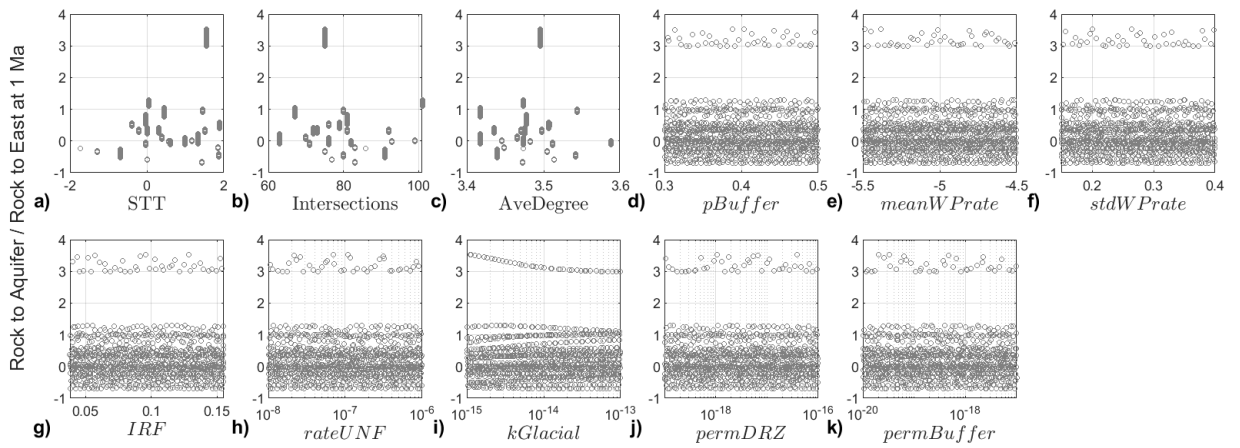
## 5. Conclusions

The results presented in this paper demonstrate significant advancements in ongoing study of the crystalline reference case for development of PA methods (Swiler et al., 2019, 2020, 2021). The addition of graph metrics that quantify realized characteristics of the DFNs clearly improved understanding of system performance and system behavior. 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. However, graph metrics were not significant in all analyses, even when spatial heterogeneity should be a significant contributor to QoI uncertainty. This indicates that the graph metrics selected do not describe all of the significant spatial heterogeneity. Refinement or addition of more graph metrics may be needed to better capture this uncertainty.

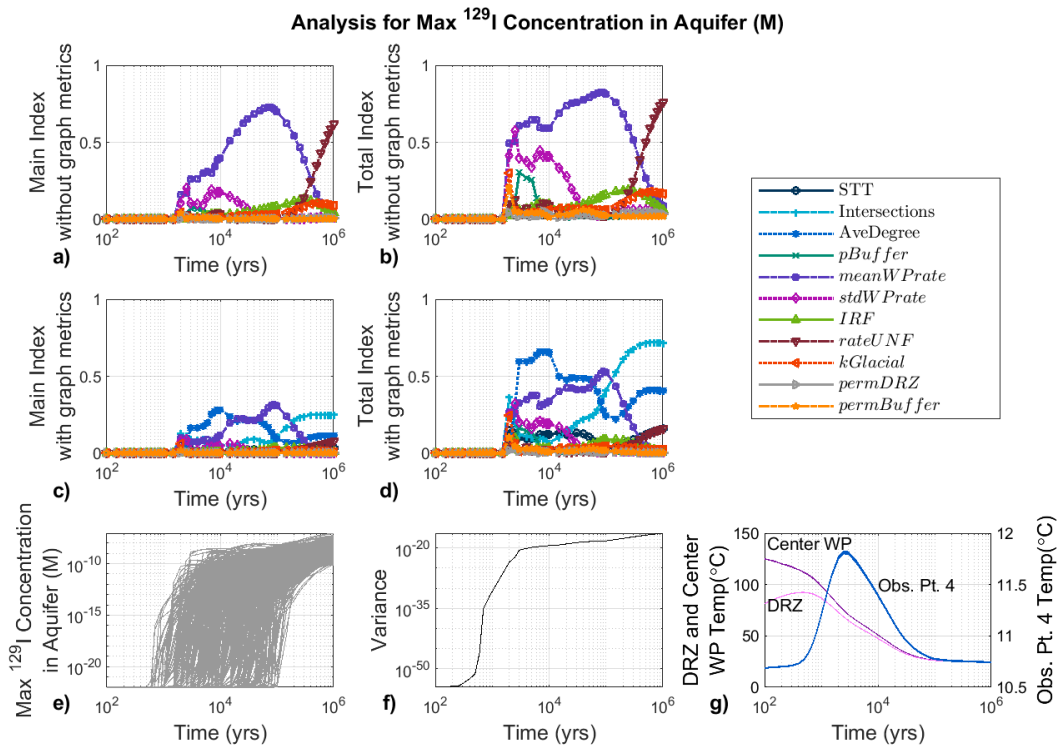
The addition of flux ratios also improved understanding of spatial effects, especially when considered with the graph metrics. Sensitivity analyses confirmed that the dominant uncertainties driving uncertainty in fluxes make phenomenological sense. The number of fracture intersections with the repository 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 with peak  $^{129}\text{I}$ , though the number of intersections has a stronger correlation.

Time-dependent sensitivity analyses were also a new capability applied in this study, which has not previously been performed for the crystalline reference case (Swiler et al., 2019, 2020, 2021). Comparison with scalar analysis and scatter plots confirmed the accuracy of these sensitivity results. Comparison to thermal profiles for waste packages and the DRZ demonstrated that such analyses provide a more nuanced understanding of changes over time.

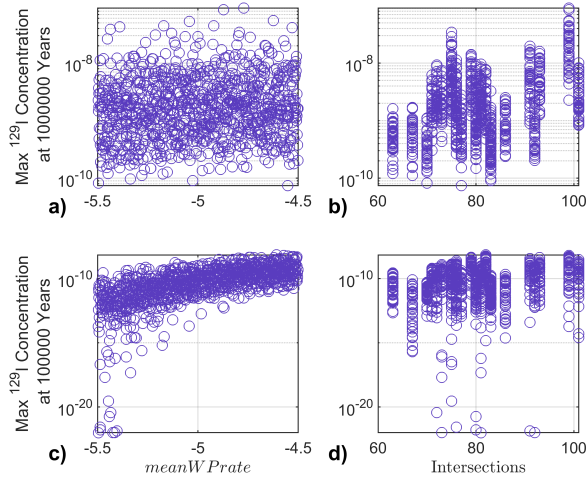
Advancements for the crystalline reference case to support PA method development are ongoing. This analysis supports further development of graph metrics to describe spatial heterogeneity, use of flux ratio QoIs to supplement  $^{129}\text{I}$  analyses, and inclusion of time-dependent sensitivity analyses to understand performance changes over time.



**Figure 10:** Ratio between the rock-to-aquifer flux and the rock-to-east-boundary flux plotted against parametric uncertainties and graph metrics. Trends are apparent for the graph metrics (a,b,c), but the only significant relationship with a parameter uncertainty is  $kGlacial$  (i). There is a clear trend between  $kGlacial$  and this flux ratio within each spatial realization, but the variation in the ratio between spatial realizations is much larger than the variation within a spatial realization, confirming the dominant effects of spatial heterogeneity.

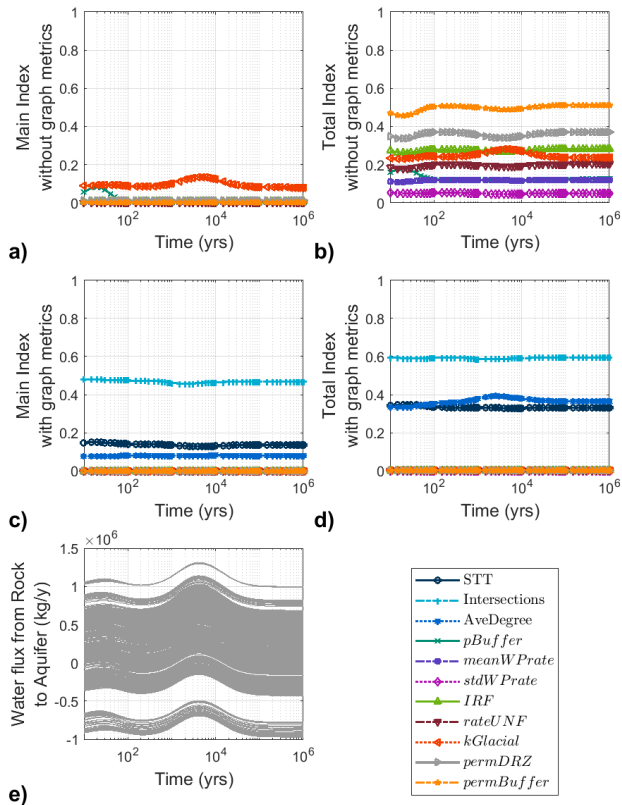


**Figure 11:** Sensitivity analysis results with and without graph metrics for the maximum  $^{129}\text{I}$  concentration in the aquifer tracked over time. Main and total effect indices are shown over time for the analysis without graph metrics (a,b) and with graph metrics (c,d). The time series results for this QoI are shown (e) with the variance between these curves (f). Temperatures for the DRZ, center WP, and observation point 4 (g) support physical interpretation of the timing for waste package parameter significance in the sensitivity results.

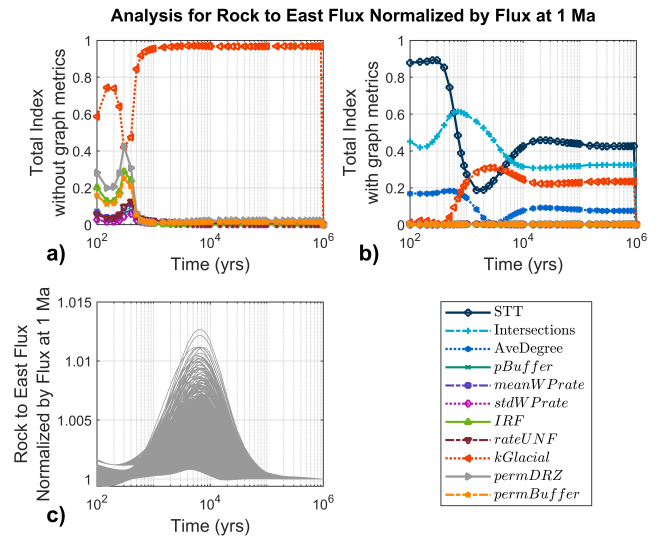


**Figure 12:** Scatter plots of the maximum  $^{129}\text{I}$  concentration in the aquifer tracked over time plotted against the  $\text{meanWPrate}$  (a,c) and the number of intersections (b,d). This sequence of plots shows the higher importance of  $\text{meanWPrate}$  earlier in the simulation (c) and the higher importance of the Intersections graph metric towards the end of the simulation (b).

**Analysis for Water flux from Rock to Aquifer (kg/y)**

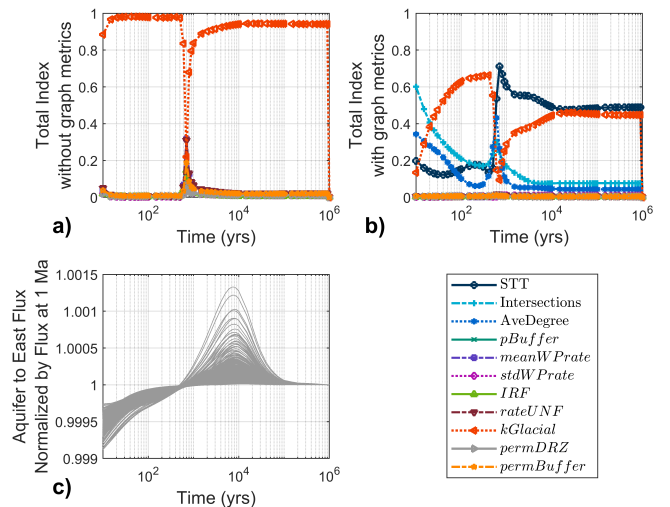


**Figure 13:** Main and Total effects indices from the sensitivity analysis without (a,b) and with (c,d) graph metrics for the water flux from the rock to the aquifer. The time series plot for this QoI (e) shows the relatively low variation in this QoI with respect to time.



**Figure 14:** Total effects indices from the sensitivity analysis without (a) and with (b) graph metrics for the water flux from the rock to the east boundary normalized by the flux at 1 Ma. Differences in the timing of importance between graph metrics and the time series plot (c) supports physical interpretation.

**Analysis for Aquifer to East Flux Normalized by Flux at 1 Ma**



**Figure 15:** Total effects indices from the sensitivity analysis results without (a) and with (b) graph metrics for the water flux from the rock to the aquifer normalized by the flux at 1 Ma. The time series plot (c) explains the sudden change in sensitivity analysis results around  $5 \times 10^2$  years.

## 6. Acknowledgments

This work has been funded by the Geologic Disposal Safety Assessment (GDSA) program under the Spent Fuel and Waste Science and Technology (SFWSST) Campaign of the U.S. Department of Energy Office of Nuclear Energy, Office of Fuel Cycle Technology. The authors thank their colleague, Mariah Smith, for generating Figure 3 of this paper.

The work has been performed at Sandia National Laboratories. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and 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-NA-0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

This article has been authored by employees of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employees own all right, title and interest in and to the article and are solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan <https://www.energy.gov/downloads/doe-public-access-plan>.

## References

- Adams, B.M., Bohnhoff, W.J., Dalbey, K.R., Ebeida, M.S., Eddy, J., 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., Winokur, J.G., 2020. DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: version 6.13 user's manual. Technical Report SAND2020-12495. Sandia National Laboratories. doi:10.2172/1817318.
- Azizmohammadi, S., Matthäi, S.K., 2017. Is the permeability of naturally fractured rocks scale dependent? *Water Resources Research* 53, 8041–8063. doi:10.1002/2016WR019764.
- Bisdom, K., Nick, H., Bertotti, G., 2017. An integrated workflow for stress and flow modelling using outcrop-derived discrete fracture networks. *Computers & Geosciences* 103, 21–35. doi:10.1016/j.cageo.2017.02.019.
- Freeze, G., Gardner, W., Vaughn, P., Sevougian, S., Mariner, P., Mousseau, V., Hammond, G., 2013a. Enhancements to Generic Disposal System Modeling Capabilities. Technical Report SAND2013-10532P, FCRD-UFD-2014-000062. Sandia National Laboratories. Albuquerque, New Mexico.
- Freeze, G., Voegelé, M., Vaughn, P., Prouty, J., Nutt, W., Hardin, E., Sevougian, S., 2013b. Generic Deep Geologic Disposal Safety Case. Technical Report FCRD-UFD-2012-000146 Rev. 1, SAND2013-0974P. Sandia National Laboratories. Albuquerque, New Mexico.
- Ghanem, R., Higdon, D., Owahdi, H., 2017. Handbook of uncertainty quantification. Volume 6. Springer.
- Ghanem, R., Red-Horse, J., 2017. Polynomial chaos: Modeling, estimation, and approximation, in: Ghanem et al. (2017).
- Hadgu, T., Karra, S., Kalinina, E., Makedonska, N., Hyman, J.D., Klise, K., Viswanathan, H.S., Wang, Y., 2017. A comparative study of discrete fracture network and equivalent continuum models for simulating flow and transport in the far field of a hypothetical nuclear waste repository in crystalline host rock. *Journal of Hydrology* 553, 59–70. doi:https://doi.org/10.1016/j.jhydrol.2017.07.046.
- Hagberg, A., Swart, P., Chult, D., 2008. Exploring network structure, dynamics, and function using networkx.
- Hammond, G., Lichtner, P., Mills, R., Lu, C., 2008. Toward petascale computing in geosciences: application to the hanford 300 area, in: *Journal of Physics, SciDAC 2008: Scientific Discovery through Advanced Computing*. doi:10.1088/1742-6596/125/1/012051.
- Hammond, G.E., Lichtner, P.C., Mills, R., 2014. Evaluating the performance of parallel subsurface simulators: An illustrative example with PFLOTRAN. *Water resources research* 50, 208–228. doi:10.1002/2012WR013483.
- Hedin, A., 2008. Semi-analytic stereological analysis of waste package/fracture intersections in a granitic rock nuclear waste repository. *Mathematical Geosciences* 40, 619–637. doi:10.1007/s11004-008-9175-3.
- Helton, J.C., Hansen, C.W., Sallaberry, C.J., 2012. Uncertainty and sensitivity analysis in performance assessment for the proposed high-level radioactive waste repository at yucca mountain, nevada. *Reliability Engineering & System Safety* 107, 44–63. doi:10.1016/j.res.2011.07.002.
- Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety* 52, 1–17. doi:10.1016/0951-8320(96)00002-6.
- Hyman, J.D., Karra, S., Makedonska, N., Gable, C.W., Painter, S.L., Viswanathan, H.S., 2015. dfnworks: A discrete fracture network framework for modeling subsurface flow and transport. *Computers & Geosciences* 84, 10–19. doi:10.1016/j.cageo.2015.08.001.
- International Atomic Energy Agency, 2011. Disposal of radioactive waste, specific safety requirements. IAEA Safety Standards Series No. SSR-5. Vienna.
- Jackson, C.P., Hoch, A.R., Todman, S., 2000. Self consistency of a heterogeneous continuum porous medium representation of a fractured medium. *Water Resources Research* 36, 189–202. doi:10.1029/1999WR900249.
- Joyce, S., Hartley, L., Applegate, D., Hoek, J., Jackson, P., 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, 1233–1249. doi:10.1007/s10040-014-1165-6.
- Le Gratiet, L., Marelli, S., Sudret, B., 2017. Metamodel-based sensitivity analysis: Polynomial chaos expansions and gaussian processes, in: Ghanem et al. (2017).
- Lichtner, P.C., Hammond, G.E., Lu, C., Karra, S., Bisht, G., Andre, B., Mills, R., Kumar, J., 2015. PFLOTRAN user manual: A massively parallel reactive flow and transport model for describing surface and subsurface processes. Technical Report LA-UR-15-20403. Los Alamos National Laboratory (LANL), Los Alamos, NM; Sandia National Laboratories (SNL), Albuquerque, NM. doi:10.2172/1168703.
- Mariner, P., Gardner, W., Hamond, G., Sevougian, S., Stein, E., 2015. Application of Generic Disposal System Models. Technical Report FCRD-UFD-2015-000126, SAND2015-10037 R. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1226515.
- Mariner, P., Hardin, E., Hansen, F., Freeze, G., Lord, A., Goldstein, B., R.H., P., 2011. Granite Disposal of U.S. High-Level Radioactive Waste. Technical Report SAND2011-6203. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1029794.
- Mariner, P., Nole, M., Basurto, E., Berg, T., Chang, K., Debusschere, B., Eckert, A., Ebeida, M., Gross, M., Hammond, G., Harvey, J., Jordan, S., Kultman, K., LaForce, T., Leone, R., III, W.M., Mills, M., Park, H., Pertty, F., III, A.S., Seidl, D., Sevougian, S., Stein, E., Swiler, L., 2020. Advances in GDSA Framework Development and Process Model Integration. Technical Report M2SF-20SNO10304042/SAND2020-

- 10787R. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1671380.
- Mariner, P.E., Stein, E.R., Frederick, J.M., Sevougian, S.D., Hammond, G.E., 2017. Advances in Geologic Disposal System Modeling and Shale Reference Cases. Technical Report SFWD-SFWST-2017-000044, SAND2017-10304 R. Sandia National Laboratories. doi:10.2172/1395750.
- Mariner, P.E., Stein, E.R., Frederick, J.M., Sevougian, S.D., Hammond, G.E., Fascitelli, D., 2016. Advances in Geologic Disposal System Modeling and Application to Crystalline Rock. Technical Report FCRD-UFD-2016-000440, SAND2016-9610 R. Sandia National Laboratories. doi:10.2172/1333708.
- Parashar, R., Reeves, D., 2011. Computation of flow and transport in fracture networks on a continuum grid, in: Proceedings of MODFLOW and MORE2011: Integrated Hydrologic Modeling, Golden, C).
- Rechard, R.P., Voegele, M.D., 2014. Evolution of repository and waste package designs for yucca mountain disposal system for spent nuclear fuel and high-level radioactive waste. Reliability Engineering & System Safety 122, 53–73. doi:10.1016/j.res.2013.06.018.
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environmental Modelling and Software 114, 29–39. doi:10.1016/j.envsoft.2019.01.012.
- Sevougian, S., Hammond, G., Mariner, P., Stein, E., Frederick, J., Mackinnon, R., 2018. GDSA Framework: High-Performance Safety Assessment Software to Support the Safety Case (SAND2018-9975C), in: Proceedings of the IGSC Safety Case Symposium, Rotterdam, The Netherlands.
- Sobol', I., 1993. Sensitivity analysis for nonlinear mathematical models. Mathematical Modeling and Computational Experiment 1, 407–414.
- Srinivasan, G., Hyman, J., Osthus, D., Moore, B., O'Malley, D., Karra, S., Viswanathan, H., 2018. Quantifying topological uncertainty in fractured systems using graph theory and machine learning. Scientific Reports 8, 1–11. doi:10.1038/s41598-018-30117-1.
- Stein, E., Frederick, J.M., Hammond, G.E., Kuhlman, K.L., Mariner, P., Sevougian, S.D., 2017. Modeling Coupled Reactive Flow Processes in Fractured Crystalline Rock. Technical Report SAND2017-0230C. Sandia National Laboratories. Albuquerque, New Mexico.
- Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety 93, 964–979. doi:10.1016/j.res.2007.04.002.
- Svensson, U., 2001. A continuum representation of fracture networks. part i: Method and basic test cases. Journal of Hydrology 250, 170–186. doi:10.1016/S0022-1694(01)00435-8.
- Sweeney, M.R., Gable, C.W., Karra, S., Stauffer, P.H., Pawar, R.J., Hyman, J.D., 2020. Upscaled discrete fracture matrix model (udfm): an octree-refined continuum representation of fractured porous media. Computational Geosciences 24, 293–310. doi:10.1007/s10596-019-09921-9.
- Swiler, L., Basurto, E., Brooks, D., Eckert, A., Leone, R., Mariner, P., Portone, T., Smith, M., Stein, E., 2021. Uncertainty and Sensitivity Analysis Methods and Applications in the GDSA Framework (FY2021). Technical Report M3SF-21SN010304042, SAND2021-9903 R. Sandia National Laboratories. Albuquerque, New Mexico.
- Swiler, L., Helton, J., Basurto, E., Brooks, D., Mariner, P., Moore, L., Mohanty, S., Sevougian, S., Stein, E., 2019. Status Report on Uncertainty Quantification and Sensitivity Analysis Tools in the Geologic Disposal Safety Assessment (GDSA) Framework. Technical Report M2SF-19SN010304031, SAND2019-13835R. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1574263.
- Swiler, L., Helton, J., Basurto, E., Brooks, D., Mariner, P., Portone, T., Stein, E., 2020. Advances in Uncertainty and Sensitivity Analysis Methods and Applications in GDSA Framework. Technical Report M3SF-20SN010304032, SAND2020-10802 R. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1671381.
- Viswanathan, H., Hyman, J., Karra, S., O'Malley, D., Srinivasan, S., Haggberg, A., Srinivasan, G., 2018. Advancing graph-based algorithms for predicting flow and transport in fractured rock. Water Resources Research 54, 6085–6099. doi:10.1029/2017WR022368.
- Wang, Y., Hadgu, T., Matteo, E., Kruichak, J., Mills, M., Tinnacher, R., Davis, J., Viswanathan, H., Chu, S., Dittrich, T., Hyman, F., Karra, S., Makedonska, N., Reimus, P., Zavarin, M., Zhao, P., Joseph, C., Begg, J., Dai, Z., Kersting, A., Jerden, J., Copple, J., Cruse, T., Ebert, W., 2015. Used Fuel Disposal in Crystalline Rocks: FY15 Progress Report. Technical Report FCRD-UFD-2015-000125. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1229704.
- Wang, Y., Matteo, E., Rutqvist, J., Davis, J., Zheng, L., Houseworth, J., Birkholzer, J., Dittrich, T., W. Gable, C., Karra, S., Makedonska, N., Chu, S., Harp, D., Painter, S., Reimus, P., Perry, F., Zhao, P., Begg, J., Zavarin, M., Tumey, S., Dai, Z., Kersting, A., Jerden, J., Frey, K., Copple, J., Ebert, W., 2014. Used Fuel Disposal in Crystalline Rocks: Status and FY14 Progress. Technical Report FCRD-UFD-2014-000060, SAND2014-17992 R. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1160341.
- Weck, P.F., Colón, C.F.J., Sassani, D.H., Zheng, L., Rutqvist, J., Steefel, C.I., Kim, K., Nakagawa, S., Houseworth, J., Birkholzer, J., Caporuscio, F.A., Cheshire, M., Rearick, M.S., McCarney, M.K., Zavarin, M., Benedicto, A., Kerstink, A.B., Sutton, M., Jerden, J., Frey, K.E., Copple, J.M., Ebert, W., 2014. Evaluation of Used Fuel Disposition in Clay-Bearing Rock. Technical Report FCRD-UFD-2014-000056, SAND2014-18303 R. Sandia National Laboratories. Albuquerque, New Mexico. doi:10.2172/1164611.
- Xiu, D., 2010. Numerical methods for stochastic computations. Princeton university press, Princeton, N.J.