SANDIA REPORT

SAND2004-3440
Unlimited Release
Printed August 2004

An Example Uncertainty and Sensitivity Analysis at the Horonobe Site for Performance Assessment Calculations

Scott C. James

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online order: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

Acknowledgement: This project was funded by Japan Nuclear Cycle Development Institute (JNC). This work was conducted jointly between Sandia National Laboratories (SNL) and JNC under a joint JNC/U.S. Department of Energy (DOE) work agreement. Performance assessment calculations were conducted and analyzed at SNL based on a preliminary model by Kashima, Quintessa, and JNC and include significant input from JNC to make sure the results are relevant for the Japanese nuclear waste program.

Table of Contents

List of Figures

List of Tables

Abstract

Given pre-existing Groundwater Modeling System (GMS) models of the Horonobe Underground Research Laboratory (URL) at both the regional and site scales, this work performs an example uncertainty analysis for performance assessment (PA) applications. After a general overview of uncertainty and sensitivity analysis techniques, the existing GMS site-scale model is converted to a PA model of the steady-state conditions expected after URL closure. This is done to examine the impact of uncertainty in site-specific data in conjunction with conceptual model uncertainty regarding the location of the Oomagari Fault. In addition, a quantitative analysis of the ratio of dispersive to advective forces, the F-ratio, is performed for stochastic realizations of each conceptual model. All analyses indicate that accurate characterization of the Oomagari Fault with respect to both location and hydraulic conductivity is critical to PA calculations. This work defines and outlines typical uncertainty and sensitivity analysis procedures and demonstrates them with example PA calculations relevant to the Horonobe URL.

1. Introduction

Characterization of uncertainty associated with flow and transport models is critically important when environmental models are used for performance and risk assessment. A systematic uncertainty analysis provides insight into the level of confidence in model estimates and aids in assessing how various model estimates should be weighed. Further, it can identify the key sources of uncertainty (such as data gaps) that merit additional site characterization as well as the sources of uncertainty that are not important for performance assessment. Uncertainty analyses use available information to quantify the degree of confidence in existing data and models. They cannot reduce uncertainty – that can only come with the collection of additional data. Instead, the goal is to understand judgments associated with modeling even though the applicability of a model is inherently limited by uncertainties in data. A quantitative uncertainty analysis can ascertain the robustness of model conclusions particularly when data are limited and when simplifying assumptions have been used. Ultimately, models should help target data gathering efforts. In addition, alternate conceptual models can play a large role in process analysis. All valid conceptual models of a site should be included in an uncertainty and sensitivity analysis because they may yield different results and insight into the processes modeled.

In this work, an existing model provided by Quintessa for the dewatering of the Horonobe Underground Research Laboratory (URL) is modified to make it amenable to an example performance assessment model. Actual site data are incorporated into its design and stochastic particle tracking realizations are presented. In addition, a synthetic parameter measuring the ratio of diffusive to advective forces, the F-ratio, is calculated along the particle tracks and analyzed.

2. Background for the Uncertainty Analysis

The following stages are commonly involved in an uncertainty analysis:

Characterization of input uncertainties – estimation of uncertainties in model input parameters (e.g. site data);

Uncertainty propagation – estimation of the uncertainty in model outputs resulting from uncertainty in model inputs (e.g. heterogeneous hydraulic conductivity fields); and,

Characterization of model uncertainty – characterization of output uncertainties associated with different model structures and model formulations (e.g. difference in conceptual model by variation of fault location).

2.1 Reducible and Irreducible Uncertainty

Uncertainty associated with model formulation and application can be classified as 'reducible' or 'irreducible.' Natural uncertainty is inherent or irreducible, whereas data and model uncertainty contain both reducible and irreducible components. The irreducible uncertainty in data and models is generally a result of the presence of natural uncertainty. An example is found in the natural heterogeneity of subsurface geology, for instance in the estimation of a hydraulic conductivity field. Reducible uncertainty can be decreased through improved instrumentation, improvements in model formulation, or increasing the number of data points. An example of reducible uncertainty is the determination of the exact location of a fault, which may be determined through careful data collection. Nevertheless, the distinction between reducible and irreducible model and data uncertainties is largely a matter of convention because it may not be feasible to eliminate the presence of an error (reducible uncertainty). Furthermore, the averaging processes involved in model formulation unavoidably 'lump' together natural and modeling uncertainty often making distinctions difficult. Fortunately, what is perceived as irreducible natural uncertainty may be quantified in a statistical sense with Monte Carlo and other stochastic techniques.

2.2 Sensitivity Analysis

The aim of a sensitivity analysis is to estimate the rate of change in the output of a model with respect to changes in inputs. Such knowledge is important for: 1) evaluating the applicability of the model; 2) determining parameters that require more accuracy; and, 3) understanding the behavior of the system being modeled. The choice of a sensitivity analysis depends on the sensitivity measure employed, the desired accuracy in the estimates of the sensitivity measure, and the computational cost involved. Based on the choice of sensitivity metric and the variation in the model parameters, sensitivity analyses can be broadly classified into the following categories:

Variation of parameters or model formulation – In this approach, the model is run at a set of sample points (different combinations of parameters of concern) or with straightforward changes in model structure (e.g., grid resolution). Sensitivity measures appropriate for this analysis are: the model response from arbitrary parameter variation, normalized model response, and extrema. Of these measures, the extreme values (worst case scenarios) are often critically important to environmental applications. That is, low probability high consequence events must be carefully considered.

Regional sensitivity analysis – Sensitivity is estimated by studying model behavior over the entire range of parameter variation, taking uncertainty in the parameter estimates into account.

Local sensitivity analysis – Here, the focus is on estimates of model sensitivity to input and parameter variation in the vicinity of a sample point. This sensitivity is often characterized through gradients or partial derivatives at the sample point (e.g., radionuclide concentrations at a receptor point). These might include first- and second-order moment techniques.

Sampling based methods involve running the model with various sets of input variables and establishing the relationship between inputs and outputs. These methods typically fall into the category of variation of parameters or model formulation, although certainly they may be used in both local and regional sensitivity analyses.. The most widely used sampling based sensitivity/uncertainty analysis methods are Monte Carlo and Latin Hypercube Sampling.

2.3 Monte Carlo and Latin Hypercube Sampling Methods

Monte Carlo methods are the most widely used means for uncertainty analyses with applications ranging from aerospace engineering to zoology. These methods involve uniform sampling from distributions of inputs and successive model runs until statistically significant distributions of outputs are achieved. They can be used to solve problems with physical probabilistic structures, such as uncertainty propagation in stochastic models, or can be used to solve non-probabilistic problems, such as finding the area under a curve. Monte Carlo (random walk) methods are also used in the solution of problems that can be modeled by the sequence of a set of random steps that eventually converge to a desired solution. Because these methods require a large number of samples (or model runs), their applicability is sometimes limited to simple models. In case of computationally intensive models, the time and resources required by these methods may be prohibitive. A degree of computational efficiency is accomplished through Latin Hypercube Sampling (LHS) that selects from the input distribution in an efficient manner, so that the number of necessary solutions compared to the Monte Carlo method is significantly reduced. In this method, the range of probable values for each uncertain input parameter is divided into ordered segments of equal probability. Thus, the whole parameter space, comprising all uncertain parameters, is partitioned into bins having equal probability, and they are sampled such that each parameter is sampled once from each of its possible segments. The advantage of LHS is that the random samples are generated from ranges of all possible values, thus giving insight into the extremes of the probability distributions of the outputs with considerably less sampling than Monte Carlo methods.

3. Model Data

3.1 Background

Environmental systems are inherently stochastic due to unavoidable unpredictability (randomness). Some quantities are random even in principle, while others that are precisely measurable (e.g., porosity) are modeled as ‘random’ quantities as a practical matter (due to the cost and effort required for precise local measurements at the site scale). For example, hydraulic conductivity has ‘natural uncertainty’ because a precise estimation is impossible. Rather, it is characterized through ensemble averages (measured over some representative volume and/or through random statistics based on site data). On the other hand, some parameters are modeled via a mean value superimposed on a ‘random error’ simply because they are impossible to monitor (e.g., radionuclide retardation parameters).

When quantities vary over time and space or across individuals in a population it is termed ‘variability.’ Variability is the heterogeneity between individual members of a population of measurements and is typically characterized through a frequency distribution. In addition, variability may be interpreted as uncertainty because both can be addressed in terms of frequency distributions. Nevertheless, the implications of the differences in uncertainty and variability are relevant in decision-making. For example, knowledge of the frequency distribution for variability can guide the identification of significant subpopulations that merit

further study. In contrast, the knowledge of uncertainty can aid in determining areas where additional site characterization is needed to reduce uncertainty.

Uncertainties in model parameter estimates stem from a variety of sources. While in principle, many parameters could be measurable up to any desired precision, often there are significant uncertainties associated with field and laboratory measurements. Some uncertainties arise from measurement errors, which involve: 1) random errors in analytic devices (e.g., the imprecision of monitors that measure temperature); 2) systematic biases due to imprecise calibration; or, 3) inaccuracies in the assumptions used to infer the actual quantity of interest from the readings of a 'surrogate' variable (e.g., inferring *K* through water level measurements). Other potential sources of uncertainties in estimates of parameters include misclassification, estimation of parameters through a small sample, and estimation of parameters through non-representative samples. Ultimately, uncertainty in model application arises from uncertainties associated with measurement data used for the model evaluation.

3.2 Horonobe Site Data

Four exploratory boreholes, HDB-1, -3, -4, and -5, were drilled within the domain of the site-scale model. Data collected from these boreholes include among others: fracture frequency, bulk density, temperature, porosity, and saturation. Fracture frequency is chosen as the most important parameter for this study and stochastic realizations of the model domain are constrained by their distributions in each borehole. (This parameter was selected because of the extensive data set available and the because these data fit into the GMS modeling framework well.) GMS allows for definition of up to five material types in a model. Because observed fracture frequency ranges from 0 to 20 fractures/m, borehole data with no fractures/m are defined as Material 1, 1–5 fractures/m as Material 2, 6–10 as Material 3, 11–15 as Material 4, and 16–20 as Material 5 as shown in Figure 1. Because hydraulic conductivity is assumed to be related to fracture frequency (i.e., more highly fractured portions of the media are more conductive), each material is assigned a hydraulic conductivity multiplier, K_f, a parameter used to modify the original hydraulic conductivity to obtain a heterogeneous field (see Table 1). Using the T-PROGS (Carle, 1999) module in GMS, 200 equally probable heterogeneous material fields conditioned on borehole data are generated based on transition probability statistics (EMRL, 2002). Using the defined material types (i.e., fracture frequency) and their associated depths, T-PROGS calculates the probability of transition between materials and assigns material properties to each finite difference cell in the model based on the resulting geostatistics.

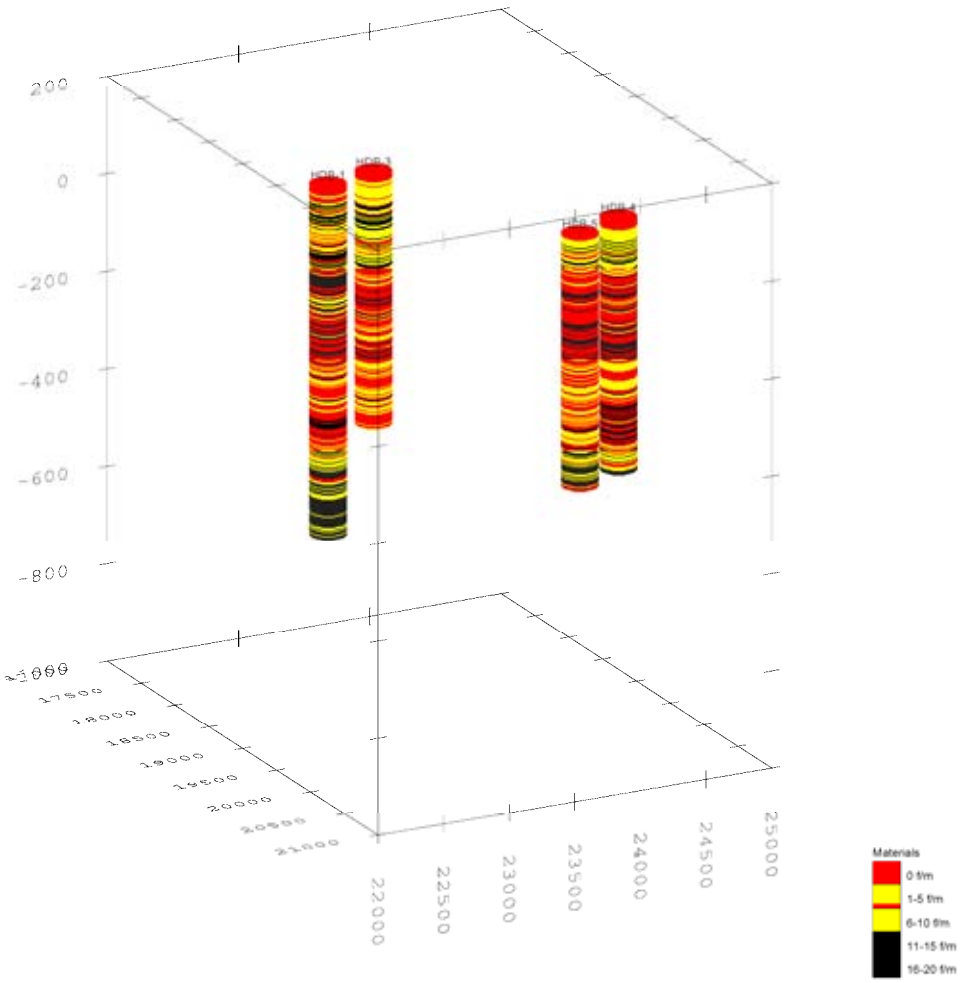

Figure 1: Boreholes HDB-1, -3, -4, and -5 are shown with color indicating the fracture frequency.

Table 1: Fracture frequency related to material type and hydraulic conductivity multiplier. The corresponding proportion of each material from the raw data set is also presented.

Fractures/m	Material	K_f	Proportion
0	1	0.1	33.2%
1–5	2	1.0	52.2%
6–10	3	10.0	12.4%
11–15	4	100.0	1.9%
15–20	5	1,000.0	0.3%

Recall that this is an example PA model and therefore some of the values used in the process are contrived (e.g., K_f in Table 1). An actual PA model would require site-specific parameters (i.e., the hydraulic conductivity as a function of fracture frequency should be measured). Also, it is assumed that the borehole data are adequately representative of the entire site when T-PROGS extrapolates these data to the entire model domain. In a formal PA model, this assumption would need to be substantiated.

4. Model Uncertainty

4.1 Background

Mathematical models are necessarily simplified representations of the phenomena being studied and a key aspect of the modeling process is the judicious choice of model assumptions. The optimal mechanistic model will provide the greatest simplifications while providing a sufficiently accurate representation of the processes affecting the phenomena of interest. Hence, the structure of mathematical models employed to represent flow and transport systems is often a key source of uncertainty. In addition to the significant approximations often inherent in modeling, competing models may also be available (e.g., single- or dual-porosity). Furthermore, the limited spatial or temporal resolution (e.g., numerical grid cell size) of many models is also an approximation that introduces uncertainty into model results. The following are definitions of the different sources of model uncertainties:

Model Structure: Uncertainty arises when there are alternative sets of scientific or technical assumptions for developing a model (e.g., modeling fractured media as a network of fractures or as an equivalent porous medium). In such cases, if the results from competing models yield similar conclusions, then one can be confident that the decision is robust in the face of uncertainty. If, however, alternative model formulations lead to different conclusions, further model evaluation is required.

Model Detail: Models are necessarily simplified to make them tractable. For example, it is often necessary to convert a complex nonlinear equilibrium to a simpler linear model when investigating the transport of reactive radionuclides. Uncertainty in the predictions of simplified models can sometimes be characterized by comparison of their predictions to those of more detailed, inclusive models.

Extrapolation: Models that are validated for one portion of input space may be completely inappropriate for making predictions in other regions of the parameter space. A groundwater flow model based upon current climatic conditions may be inaccurate when applied to a significantly drier or wetter climate. Similarly, in groundwater contaminant modeling, models that are evaluated only for base case contaminant levels, may introduce uncertainties when they are employed in the study of future scenarios that involve significantly different contaminant levels.

Model Resolution: In the application of numerical models, selection of a spatial and/or temporal grid size often involves negotiation between computational and project resources. For example, there is a tradeoff between the computation time (hence cost) and prediction accuracy. In addition, there may be a tradeoff between resolution and the validity of the governing equations of the model at such scales. Typically, a coarse grid resolution introduces approximations and uncertainties into model results. However, a finer grid resolution may not necessarily result in more accurate predictions. For example, a coarse-grid model ignores significant ‘sub-grid’ detail, and a fine-grid model requires more computational resources, and may even produce incorrect results because the macro-scale governing equations may not be appropriate at the micro-scale. This type of uncertainty may be managed with an appropriate selection of model domain parameter values or by comparing results based on different grid sizes.

Model Boundaries: All models have boundaries in terms of time, space, number of chemical species, types of pathways, and so on. Even the selection of a model boundary is a type of simplification. Within the boundary, the model may be an accurate representation, however it must be maintained that overlooked phenomenon may play an unaccounted role in the model scenario.

4.2 The Quintessa GMS Model of Horonobe

Quintessa has developed a Groundwater Modeling System (GMS) model of the Horonobe Underground Research Laboratory (URL) site. They began with a regional-scale model where the average surface elevation was interpolated from a detailed topographic map. The model covers 49 km (north-south), 37.5 km (east-west), and is 5 km deep with a 100-column×100-row×80-layer grid (800,000 cells) using natural boundary conditions (e.g., the seacoast) when possible. The geologic properties of this model (e.g., hydraulic conductivity) are based on the known geology of the area. Within the regional-scale model domain, Quintessa developed both a steady state and transient site-scale GMS model of the Horonobe site to study the effects of dewatering during URL construction. This model is 4.02 km (north-south), 3 km (east-west), and 1,250 m deep with an 83-column×72-row×124-layer grid (741,024 cells) with local refinement in the URL region and near the water table as shown in Figure 2. No recharge is applied to either of the regional- or site-scale models. The boundary conditions for the site-scale model were established through interpolation of the results (because of mesh inconsistencies) from the regional-scale model.

The Quintessa GMS model employed several conceptual models of the Oomagari Fault that runs through the Horonobe site-scale model domain. The Oomagari Fault is distinct from fractured media in that it comprises a fairly wide zone of altered geologic properties due to a major earthquake long ago. Because site characterization at Horonobe is just beginning, the exact location of the fault as well as its hydraulic conductivity (low or high) is unknown. Using Quintessa's steady-state site-scale model, Sandia National Laboratories has developed a stochastic model based on site data to demonstrate the processes and procedures typical of performance assessment (PA) and decision analysis models.

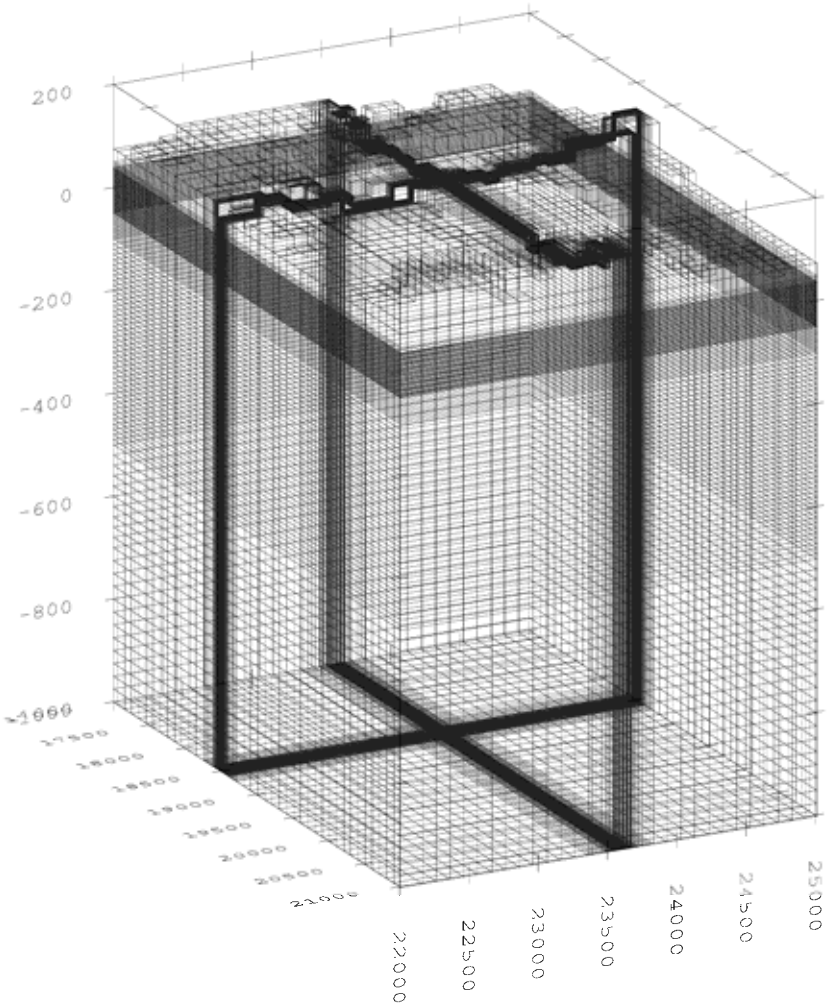

Figure 2: Finite difference mesh used in the Quintessa site-scale model of the Horonobe URL. The figure is stretched by a factor of 3.5 in the vertical to more clearly illustrate the mesh. Axes coordinates are given in meters.

Quintessa generated four fundamental conceptual models of the Horonobe URL each with the Oomagari Fault in an alternate position [center (CF), no fault (NF), east (EF), and west (WF)]. As subsets of each conceptual model, the fault, if present is assigned either high or low hydraulic conductivity with the remaining model domain also independently assigned either high or low hydraulic conductivity (a total of 16 alternate conceptual models). In this work, stochastic heterogeneous hydraulic conductivity fields based on the borehole data for each of Quintessa's four fundamental conceptual models were generated. Specifically, each fault location (Cases CF, EF, and, WF) with high values of hydraulic conductivity for the fault (where the fault acts as a conduit for flow – worst case scenario) and one case with no fault were examined (Case NF). All four fundamental conceptual models evaluated here assume base hydraulic conductivities of $K = 10^{-7}$ m/s for the non-fault upper layers (in certain areas this extends down 430 m into the model), $K = 10^{-10}$ m/s for the non-fault lower layers, and $K = 10^{-5}$ m/s for the fault, which are modified by stochastic random material fields as illustrated in Figure 3.

A total of 200 random material fields were generated with T-PROGS (50 for each of the four conceptual models) and each was used to create a corresponding random hydraulic conductivity field. GMS was used to solve the resulting flow field. Table 2 lists the average (and standard deviation) of the number of grid blocks composed of Materials 1 through 5. Across all 200 material fields from T-PROGS, on average the model is composed of 33.5% Material 1, 52.2% Material 2, 12.0% Material 3, 1.9% Material 4, and 0.3% Material 5, which is consistent with the fracture frequency measurements taken in the boreholes and listed in Table 1.

Table 2: Average (and standard deviation) of the number of gridblocks of each material type for the four conceptual models.

Conceptual model	Material 1	Material 2	Material 3	Material 4	Material 5
CF	249354 (3137)	386737 (3264)	88383 (1786)	14087 (852)	2464 (327)
NF	248180 (2439)	387274 (2129)	89152 (2242)	13793 (697)	2625 (435)
EF	247753 (2916)	386235 (2489)	90538 (2137)	14112 (827)	2386 (397)
WF	247317 (1902)	387758 (2592)	88921 (2271)	14519 (721)	2510 (280)

Once each realization has been completed, the data are exported to an ASCII text file where a FORTRAN pre-processor reads in the material data and original GMS hydraulic conductivity field and then generates a new hydraulic conductivity field based on both inputs. Table 1 defines material properties and their associated hydraulic conductivity multiplier, K_f. Figure 3 shows: (a) original log hydraulic conductivity field (note the high hydraulic conductivity fault); (b) one realization of the material field generated with T-PROGS; and, (c) the resulting log hydraulic conductivity field equal to the original hydraulic conductivity field with each cell's K value multiplied by K_f from the corresponding cell of the material field. The high hydraulic conductivity fault remains, but the overall hydraulic conductivity field for the system is now heterogeneous. It should be noted that the average hydraulic conductivities of Figures 3(a) and 3(c) are unequal.

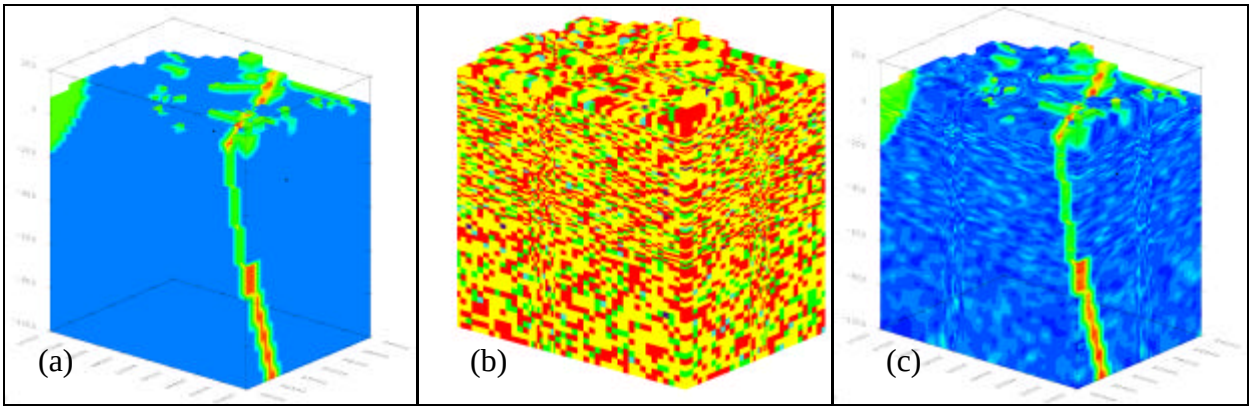

Figure 3: (a) The original log hydraulic conductivity field (Case CF); (b) the stochastic material multiplier field; and (c) the resulting log hydraulic conductivity field.

4.3 Particle Tracking

Particles were released at 250 and 500 m depth in each of the boreholes of Figure 1 (a total of eight particles). They are tracked until they exited the model boundaries. Within a conceptual model, differences in particle pathlengths and velocities (travel times) are due to the heterogeneity of the hydraulic conductivity field. Between conceptual models, the placement of the fault also yields variability in particle pathlengths and velocities. Descriptive statistics for pathlengths and travel times among the 50 stochastic hydraulic conductivity fields are listed in Table 3. Results are also compared across conceptual models to identify differences due to fault location. Figure 4 illustrates an example of the GMS particle tracking results for the hydraulic conductivity field of Figure 3a (CF). The red crosses are particle release locations and the red circles are their locations upon crossing the model boundary. The colors on the plots indicate the hydraulic head calculated with GMS where values range from about 50 m (blue) to about 150 m (red). Clearly, particles are moving downward and exit the system through the northern (y = 25,000 m) boundary in this conceptual model.

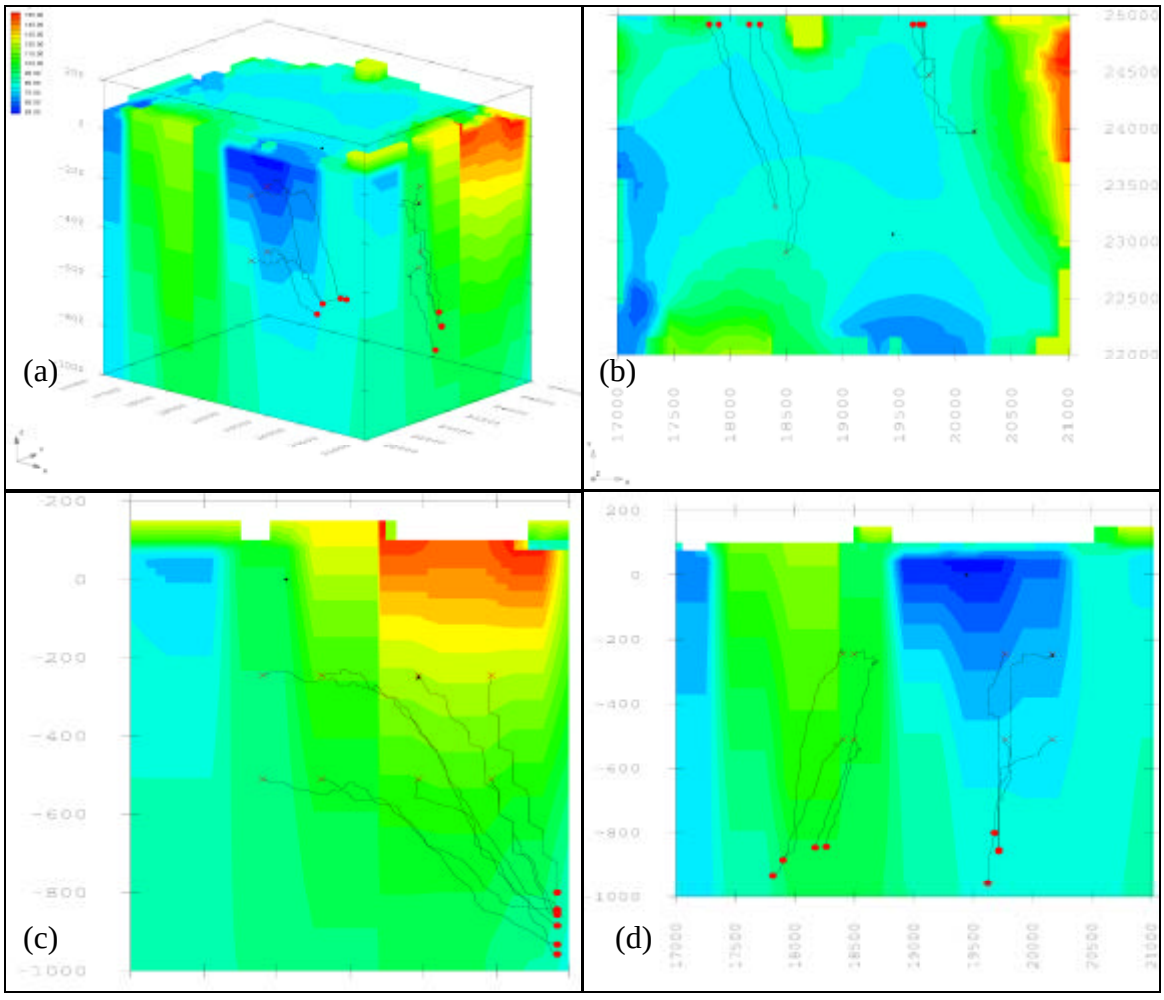

Figure 4: Particle tracks for the flow field corresponding to Figure 3c (Case CF) for the (a) oblique view, (b) top (*x*-*y*) view, (c) side (*z*-*x*) view, and (d) front (*z*-*y*) view.

Plan and map views of the 50 (x, y) particle tracks for each conceptual model for particles released –250 m down each borehole (Particles 1–4) are shown in Figure 5 and for particles

released at –500 m down each borehole (Particles 5–8) in Figure 6. The thick red line indicates the location of the top of the Oomagari Fault. Note that the particle tracks for conceptual models EF and NF are similar because in the EF case, the fault is far enough away from the particle tracks to have no significant impact on their paths. Differences between individual pathways are primarily due to variations in the random stochastic material field yielding local differences in the flow field. Furthermore, Particles 3 and 4 for Case WF are little changed from those in Cases EF and NF because the fault in Case WF is also far enough away from their particle paths to minimally impact their trajectories. The change in location of the fault between Cases CF and WF has a dramatic impact on the trajectories of Particle 1 where in Case CF particles exit across the north boundary and in Case WF they primarily exit across the south boundary. Clearly, the exact location of the fault is critical to determining flowpaths. In comparing Figure 5 and Figure 6, it is noted that the flow field changes noticeably with depth. Specifically, Particles 1 and 5 show different general flow directions even though their starting locations only differ by 250 m in the vertical.

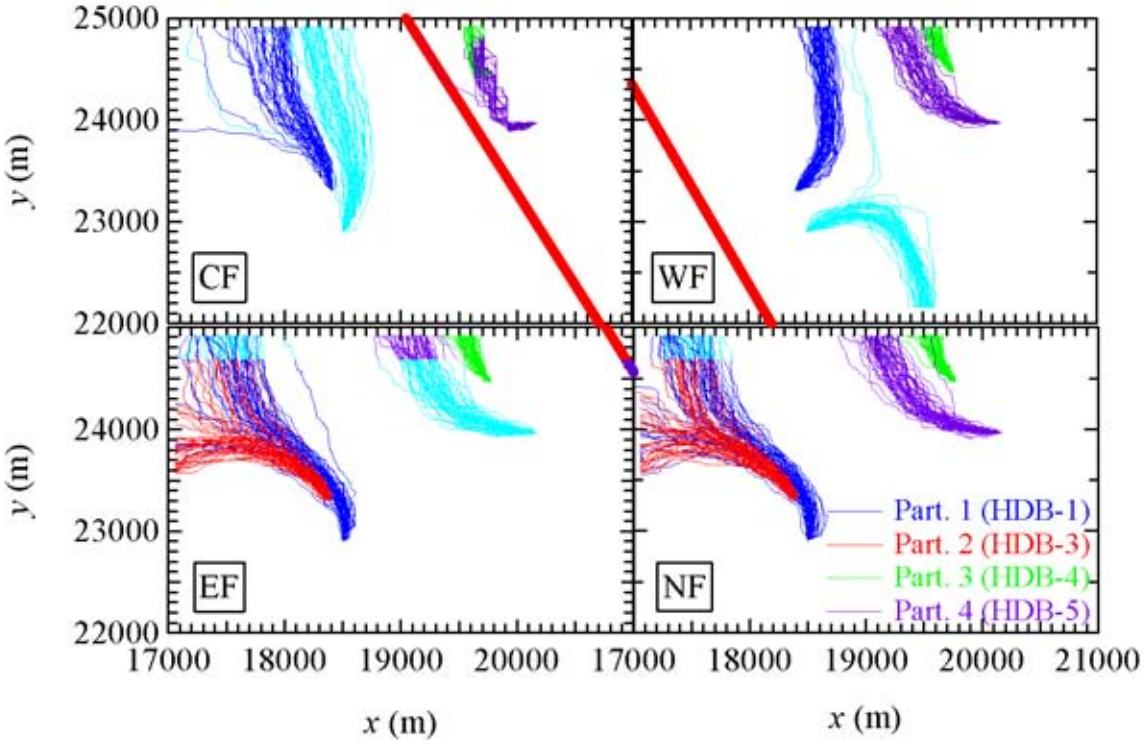

Figure 5: Particle tracks for the fifty realizations of each conceptual model (particles released –250 m down each borehole). The thick red line indicates the location of the Oomagari Fault.

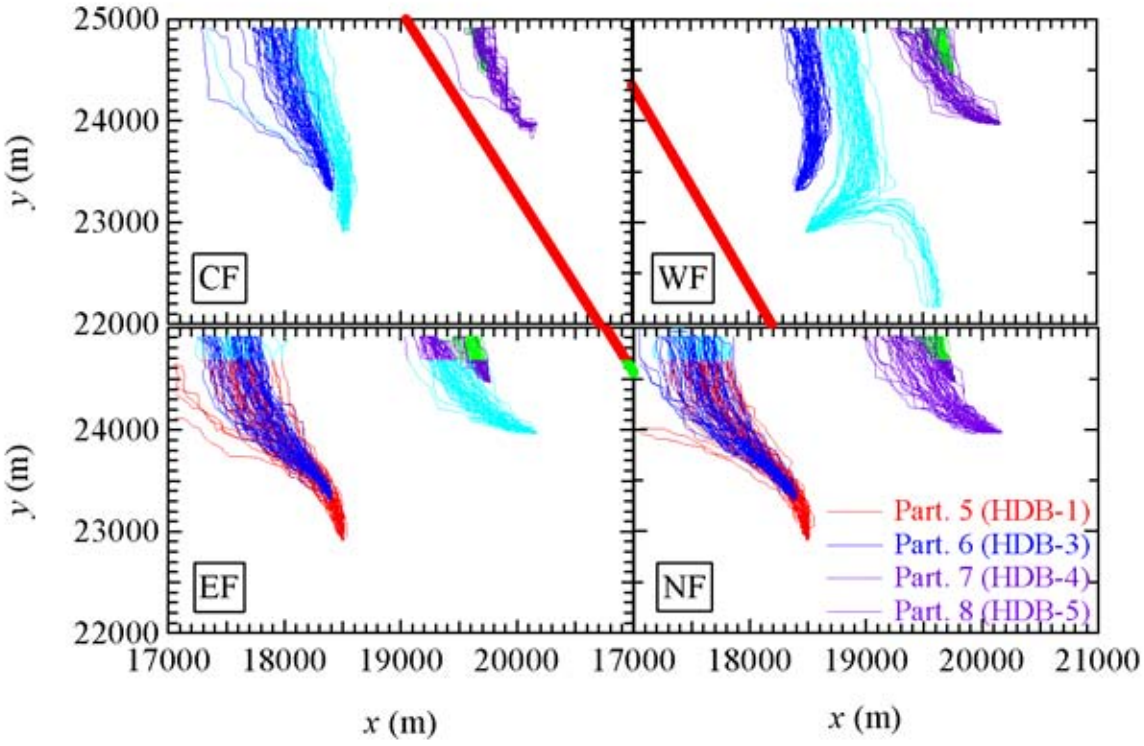

Figure 6: Particle tracks for the fifty realizations of each conceptual model (particles released –500 m down each borehole). The thick red line indicates the location of the Oomagari Fault.

The mean pathlength and travel times for each particle are listed in Table 3 and shown graphically in Figure 7 (recall that porosity is constant across the entire model domain). Although the conceptual models all yield comparable results, it is clear that in Case WF the particle tracks are shortest. This is because of the direction change for Particle 1 and the low variability in pathlines for Particle 2. Across realizations for one particle and one conceptual model, correlations between pathlength and average particle velocity can be as low as –0.48 (Case CF, Particle 8) and as high as 0.45 (Case CF, Particle 3). Nevertheless, there is little correlation (–0.13) between these parameters on the whole and no consistent conclusions can be drawn about the relationship between pathlength and velocity. As can be inferred from Figure 5, pathlengths and travel times for all particles in Cases EF and NF are similar as are those for Particles 3 and 4 of Cases WF, EF, and NF. Because they are located closest to the fault, Particle 3, 4, 7, and 8 for Case CF are the fastest particles.

Table 3: Summary statistics of the eight particle tracks through the four conceptual models (l is pathlength and *v* is velocity).

	l_1 (m)	v_1 (m/s)	l_2 (m)	v_2 (m/s)	l_3 (m)	v_3 (m/s)	l_4 (m)	v_4 (m/s)	l_5 (m)	v_5 (m/s)	l_6 (m)	v_6 (m/s)	l_7 (m)	v_7 (m/s)	l_8 (m)	v_8 (m/s)
CF Avg.	4326	1.1×10^{-4}	1771	7.6×10^{-5}	618	1.1×10^{-4}	1683	3.8×10^{-4}	2550	8.8×10^{-5}	2279	8.0×10^{-5}	495	1.1×10^{-3}	1507	5.4×10^{-4}
CF Stdev.	547	2.0×10^{-5}	71	9.9×10^{-6}	90	6.4×10^{-5}	23	5.5×10^{-4}	178	5.1×10^{-6}	203	7.6×10^{-6}	56	2.4×10^{-3}	64	1.1×10^{-3}
WF Avg.	3144	1.7×10^{-4}	1885	1.3×10^{-4}	538	1.4×10^{-4}	1573	1.3×10^{-4}	2369	6.9×10^{-5}	1796	1.2×10^{-4}	487	1.7×10^{-4}	1382	1.3×10^{-4}
WF Stdev.	509	4.5×10^{-5}	41	3.7×10^{-5}	56	4.0×10^{-5}	78	2.4×10^{-5}	162	2.2×10^{-5}	32	4.7×10^{-5}	67	5.0×10^{-5}	105	2.9×10^{-5}
EF Avg.	3948	1.2×10^{-4}	1824	7.3×10^{-5}	551	1.2×10^{-4}	1747	1.2×10^{-4}	2538	7.9×10^{-5}	2068	8.3×10^{-5}	541	1.3×10^{-4}	1488	1.3×10^{-4}
EF Stdev.	486	2.9×10^{-5}	263	1.0×10^{-5}	55	2.3×10^{-5}	145	3.5×10^{-5}	131	1.9×10^{-5}	109	2.0×10^{-5}	97	6.4×10^{-5}	77	1.9×10^{-5}
NF Avg.	3993	1.1×10^{-4}	2228	7.4×10^{-5}	634	9.7×10^{-5}	1747	1.2×10^{-4}	2615	8.8×10^{-5}	2162	8.0×10^{-5}	542	1.4×10^{-4}	1491	1.1×10^{-4}
NF Stdev.	553	2.2×10^{-5}	277	2.0×10^{-5}	115	2.4×10^{-5}	43	1.6×10^{-5}	167	2.3×10^{-5}	196	1.3×10^{-5}	41	6.6×10^{-5}	82	3.2×10^{-5}

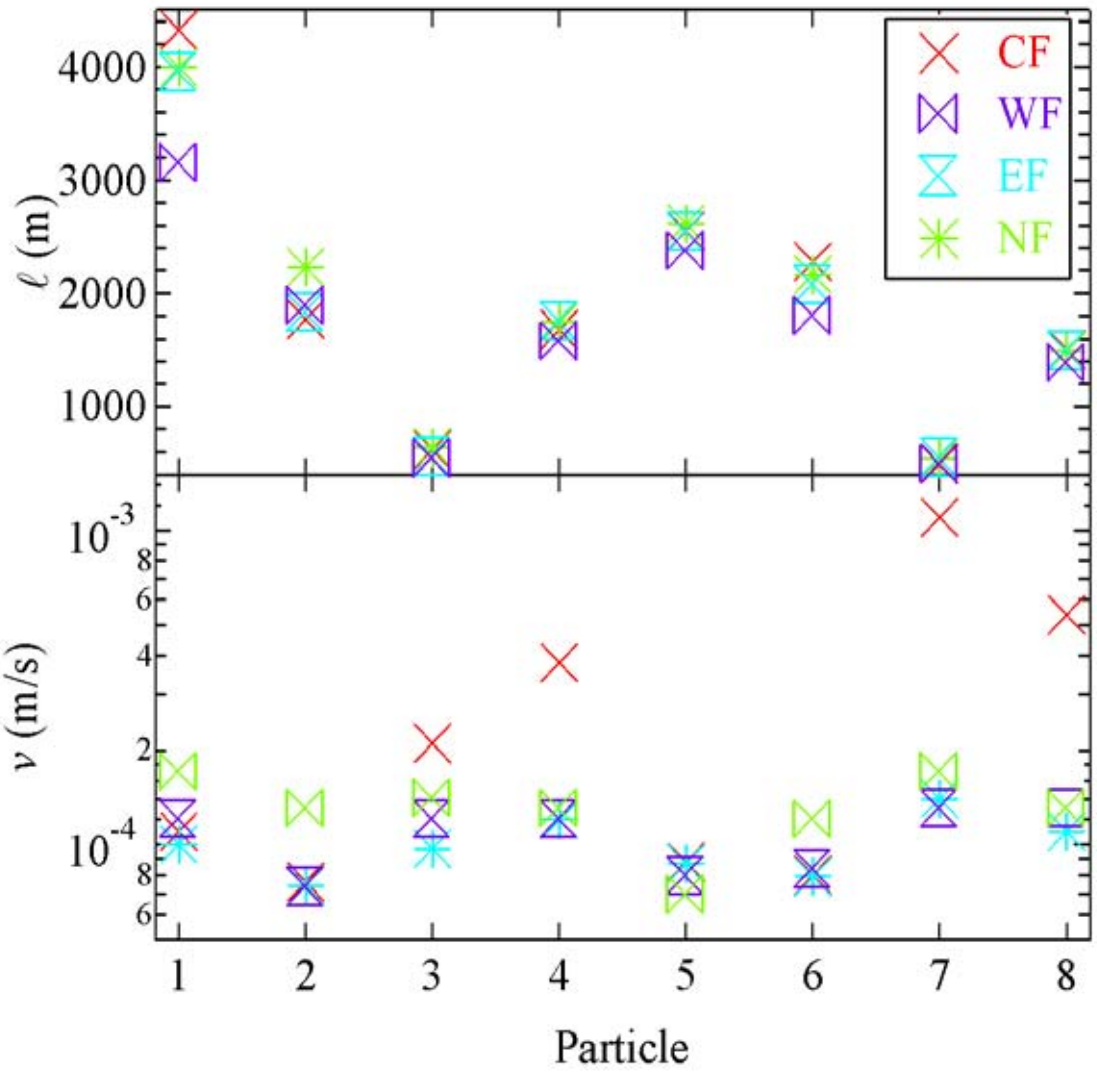

Figure 7: Average pathlength and velocity for the eight particles released from each of the four conceptual models.

The average and standard deviation of changes in depth for each particle from injection until it exits the model domain are presented in Table 4. For Cases CF, EF, and NF, particles released at –250 m have greater displacement than those released at –500 m. This is because all particles travel downward and those released at –500 m more quickly encounter the effects of the bottom no-flow boundary thus limiting their downward displacement. It is not surprising that Case WF is the exception to this trend because of the dramatic shift in paths for Particles 1 and 2.

Table 4: Average and standard deviation of changes in depth for each particle from injection until it exits the model domain for each of the conceptual models.

Case	CF		WF		EF		NF	
Particle	Δz_{avg} (m)	Δz_{stdev} (m)	Δz_{avg} (m)	Δz_{stdev} (m)	Δz_{avg} (m)	Δz_{stdev} (m)	Δz_{avg} (m)	Δz_{stdev} (m)
1	–565	45	–159	80	–602	86	–563	130
2	–552	101	–433	48	–441	205	–507	204
3	–429	115	–170	37	–195	44	–186	40
4	–511	80	–365	52	–400	45	–393	42
5	–359	49	–318	67	–396	33	–393	51
6	–381	49	–314	45	–389	36	–398	40
7	–253	80	–103	34	–121	35	–128	41
8	–343	47	–262	53	–254	59	–264	45

4.4 The F-Ratio

Conceptually, the F-ratio can be thought of as the ratio of factors resisting transport (i.e., path length and surface area available for diffusion) to those factors driving transport (i.e. velocity). The F-ratio is defined as (SKI, 1997)

$$F = \frac{\ell \times S_A}{v_D} = \frac{t \times S_A}{q}, \qquad (1)$$

where l is the MODPATH pathlength through a gridblock, S_A is the surface area of fractures in a grid block, v_D is the Darcy velocity in a gridblock, t is the travel time through a gridblock, and q is the porosity of the gridblock. Thus, there is a value of F calculated for each gridblock through which a particle travels. The F-ratio calculations require assumptions on the average fracture frequency and the amount of the fracture surface that is actually in contact with zones of significant advection in the fracture itself. The S_A for each gridblock depends upon the material type as specified by the material fields generated by T-PROGS as listed in Table 5 (each fracture is composed of two parallel plates separated by the fracture aperture). Note that there are always three times the number of fractures per block listed in Table 1 to account for the three orthogonal directions and that the number of fractures per meter is the median of the values listed in Table 1 for that material. Although each material could be assigned a unique porosity, F would remain unchanged because the travel time, t, as calculated from MODPATH is inversely proportional to the porosity. The entire model domain is assigned a porosity of 0.3, which is consistent with the Quintessa model and previous models of the URL. For the calculation of F, the proportion of the fracture surface in contact with actively flowing water is assumed to be 5% (McKenna et al., 2001).

Table 5: Fracture frequency and corresponding surface area for each material type.

Material	Fractures/m	Fractures/m^3	S_A (m^2)
1	0.5	1.5	3
2	3	9	18
3	8	24	48
4	13	39	78
5	18	54	108

Box and whisker plots were generated to examine the statistics of *F* as a function of material type and particle. For the box and whisker plots of Figure 8 and Figure 9 the blue line is the median value, the yellow box indicates the range for the 25^{th} to 75^{th} percentile *F* values, the green whiskers indicate the 5^{th} and 95^{th} percentiles, and the red outliers are values that are greater than 150% of the whisker length. All outliers reside above the 95^{th} percentiles for all particles and all materials. These (few) high values are caused by long residence times in a gridblock with a low hydraulic gradient and/or low hydraulic conductivity, which is inherently a fairly rare occurrence because flowpaths tend to avoid these cells. Note that as expected, *F* increases with material type (because S_A increases with material type). Across different particle tracks, no such simple relations exist. However, it is important to note the similarities between Cases EF and NF in both figures. This is expected because when the fault is located in the east of the model domain, it impacts particle tracks minimally and this conceptual model is therefore nearly equivalent to the case with no fault (see Figure 5). Note that the low median values for Particles 3, 4, 7, and 8 of Case CF correspond to the high velocities noted in Figure 7.

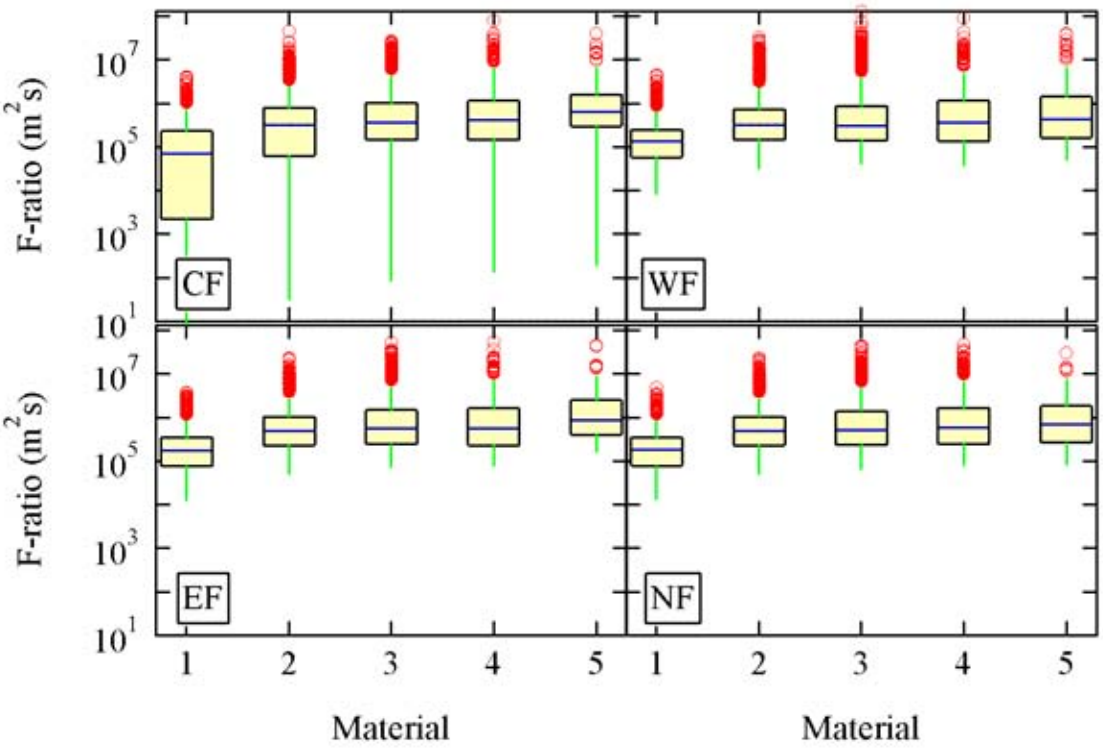

Figure 8: F-ratio distribution for each grid block of the four conceptual models and for each material.

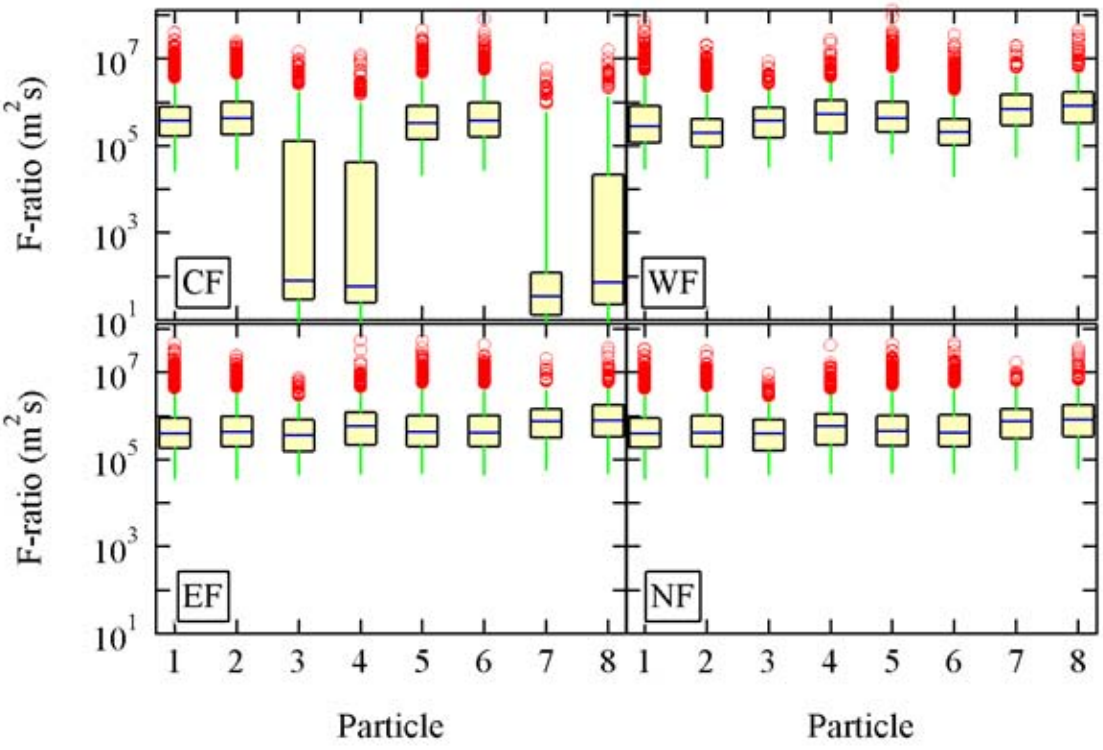

Figure 9: F-ratio distribution for each grid block of the four conceptual models and for each particle.

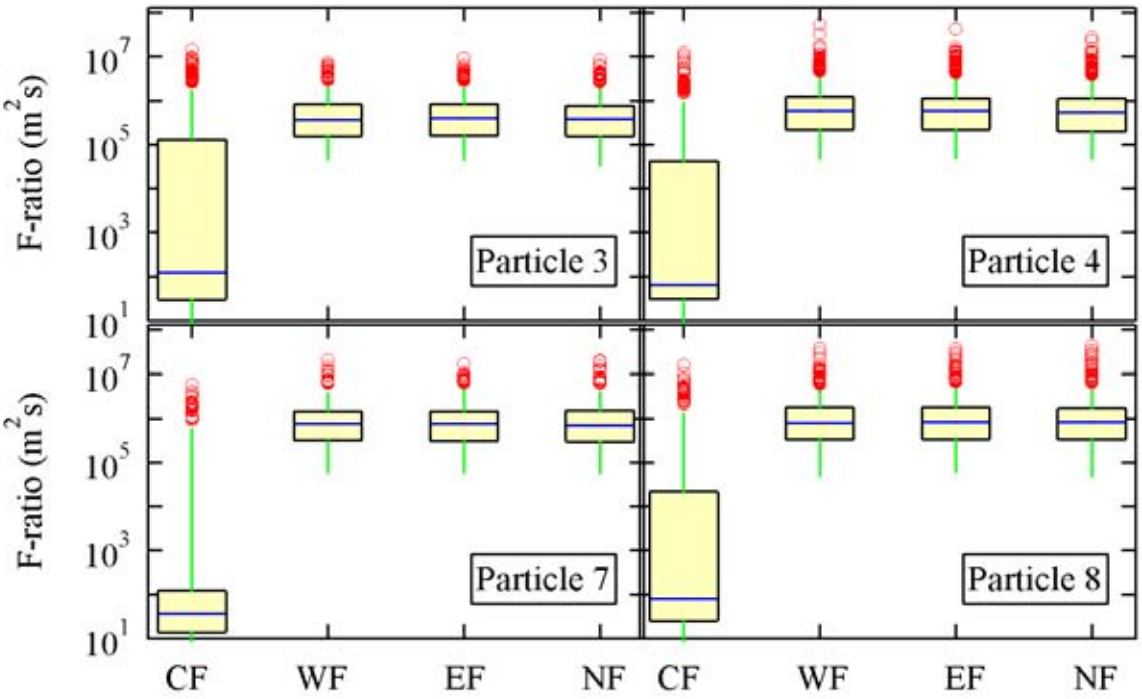

Figure 10: Comparison of F-ratios across different conceptual models for the particles most affected by the location of the fault (Particles 3, 4, 7, and 8).

Figure 10 compares F-ratios for the four particles that showed the most variation in Figure 9, Particles 3, 4, 7, and 8. Clearly, the conceptual model impacts the F-ratio with Case CF dramatically decreasing the calculated F values. Because the fault is modeled as a high hydraulic conductivity feature, when it is near the center of the model and can influence the particle paths, residence times decrease causing the F-ratios to increase. Cases WF, EF, and NF all show similar F-ratios because the fault is far from these particle paths and thus has a negligible effect on their transport.

In addition to F being calculated for each gridblock through which a particle passes, the time weighted F-ratio was also calculated for each of the fifty realizations (particle tracks) of the four conceptual models. The time weighting of the F-ratio is calculated as

$$F_{\text{weighted}} = \sum_{i=1}^{N} \frac{t_i}{t_{\text{total}}} F, \qquad (2)$$

where t_i is the time spent in cell i and t_{total} is the total travel time in the model domain through N cells. Box and whisker plots for each of the particles in each conceptual model are shown in Figure 11. As before, there is a strong similarity between Cases EF and NF. Overall, the high values for F indicate a potential for significant retardation (a great deal of diffusion and small velocities). Time weighted median F-ratios tend to be higher than median gridblock F-ratios because the averaging process weighs longer residence times more heavily and longer residence times correspond to higher F-ratios. Note also that in comparison to the range of F-ratios calculated for each gridblock (Figure 8) that time weighted F-ratios span a much smaller range because high and low gridblock values are damped through the averaging process along

a particle track (few outliers). The location of the fault has a notable impact on the weighted F-ratio distribution. Distributions in the CF and WF conceptual models (which significantly impact particle paths) are wider than for cases EF and NF where the fault has little impact on particle paths.

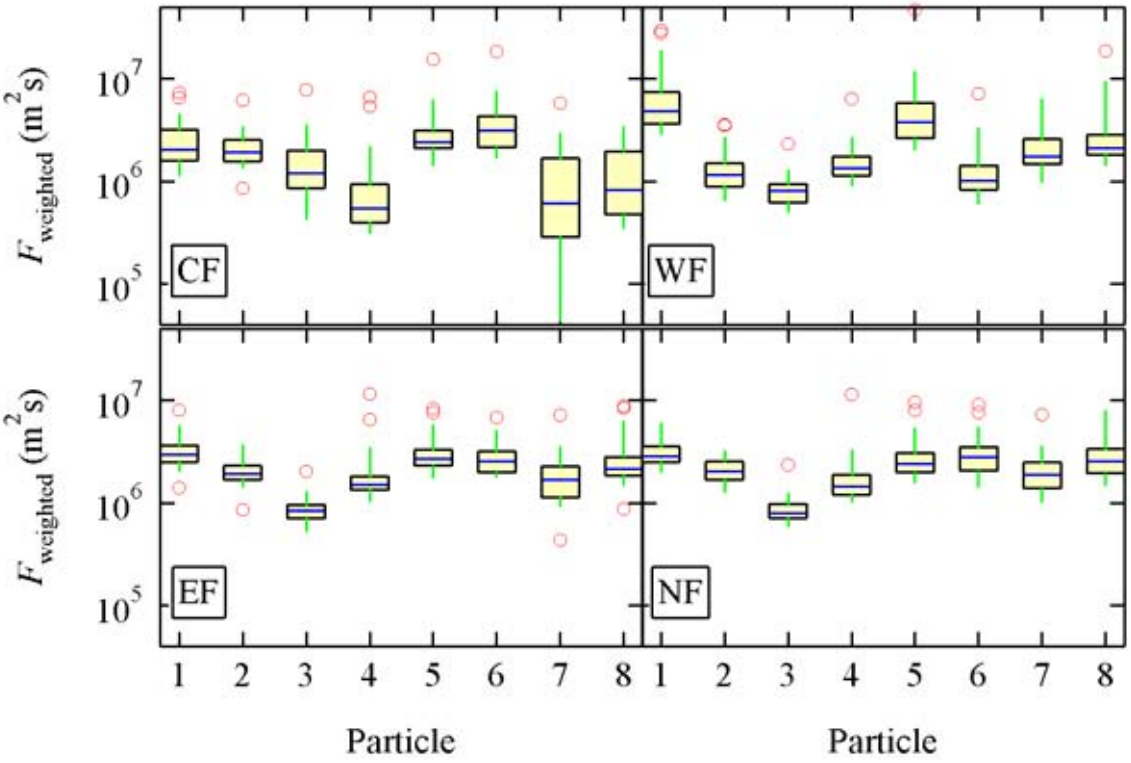

Figure 11: Box and whisker plots of the time weighted F-ratio for each particle and each conceptual model.

5. Discussion and Conclusions

Individual pathlengths and velocities for each particle and each realization are presented in Table 3. Several conclusions may be inferred from these results. For example, on average, particles released at 250 m depth travel 369 m further than particles released from the same boreholes at 500 m depth (due to the no flow bottom boundary of the model). However, average velocities are similar. Also, the location of the Oomagari Fault can significantly affect flow paths. Although in this modeling exercise the Oomagari Fault is modeled as a high conductivity feature, equally sensitive (although different) results would be expected from a low conductivity fault. Thus, the location and characteristics of the fault are critical to PA modeling.

The F-ratio as a measure of retarding (diffusion across surface area) to advancing forces (velocity) for this model indicates that significant retardation should be expected from both sorbing and non-sorbing radionuclides due to high values of *F*. It should be noted, however, that the presence of colloids could negate much of the diffusive retarding forces. For Case CF, particles released close to the fault show significantly lower F-ratios. This is because the high hydraulic conductivity fault acts as a conduit to flow and can result in increased velocities.

Again, the location and characteristics of the fault are important factors in assessing how well the site is expected to perform as a repository.

Overall, this modeling effort indicates that a more thorough identification of the fault location and properties would most significantly reduce the uncertainty in this PA model. Therefore, it is recommended that a thorough characterization of the Oomagari Fault be undertaken. In addition, it is important to identify primary material types expected within the model domain and to correlate parameters such as fracture frequency, and hydraulic conductivity.

One irregularity that must be addressed is that the modeled flow field is generally downward (–*z* direction) while at the Horonobe URL the flow field is typically upward (+*z* direction). This is due to the boundary conditions supplied by Quintessa from their regional-scale model. Because their model was intended to reproduce the conditions expected during dewatering of the site, the steady-state direction of the flow field is of little importance to that application. For PA modeling, however, it is critical to accurately represent steady-state conditions and, in practice, this model should be reworked (possibly with a different recharge boundary condition) to yield the expected upward flow field and the PA simulations could be run again.

6. References

Carle, S. F., *T-PROGS: Transition Probability Geostatistical Software Version 2.1*, Hydrologic Sciences Graduate Group, University of California, Davis, 1999.

EMRL, *GMS 4.0 Tutorial*, Environmental Modeling Research Laboratory, Brigham Young University, 2002.

McKenna, S. A., M. Eliassi, K. Inaba, and H. Saegusa, Steady-state groundwater flow modeling of the MIU Site area, SAND2001-3570C, Sandia National Laboratories, Albuquerque, NM, 2001.

SKI, SKI Site-94 Deep repository performance assessment project, Swedish Nuclear Power Inspectorate, SKI Report 97:5, Stockholm, Sweden, 1997.

Distribution
FOREIGN DISTRIBUTION:

Richard Jensen
4-33, Muramatsu, Tokai-Mura, Naka-Gun
Ibaraki, Japan 319-1194

Takanori Kunimaru
Miyazono-machi 1-8, Horonobe-cho,
Hokkaido, Japan 098-3207

Hiroshi Kurikami
Miyazono-machi 1-8, Horonobe-cho
Hokkaido, Japan 098-3207

Hitoshi Makino (5)
4-33, Muramatsu, Tokai-Mura, Naka-Gun
Ibaraki, Japan 319-1194

Keiichiro Wakasugi (2)
4-33, Muramatsu, Tokai-Mura, Naka-Gun
Ibaraki, Japan 319-1194

Shinichi Yamasaki (2)
Miyazono-machi 1-8, Horonobe-cho
Hokkaido, Japan 098-3207

Mikazu Yui
4-33, Muramatsu, Tokai-Mura, Naka-Gun
Ibaraki, Japan 319-1194

SANDIA INTERNAL:

1	MS 0719	6142	D. D. Drayer
1	MS 0735	6115	S. J. Altman
1	MS 0735	6115	R. E. Finley
4	MS 0735	6115	S. C. James
1	MS 0735	6115	T. S. Lowry
1	MS 0735	6115	S. A. McKenna
1	MS 0778	6852	B. W. Arnold
2	MS 0899	9616	Technical Library
1	MS 9018	8945-1	Central Technical Files
1	MS 1393	12100	E. K. Webb