ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Reliability Engineering and System Safety ▮ (▮▮▮▮) ▮▮▮–▮▮▮

RELIABILITY ENGINEERING & SYSTEM SAFETY

www.elsevier.com/locate/ress

Survey of sampling-based methods for uncertainty and sensitivity analysis

J.C. Helton[a,*], J.D. Johnson[b], C.J. Sallaberry[c], C.B. Storlie[d]

[a]Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287-1804, USA
[b]ProStat, Mesa, AZ 85204-5326, USA
[c]Sandia National Laboratories, Albuquerque, NM 87185-0776, USA
[d]Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, USA

Abstract

Sampling-based methods for uncertainty and sensitivity analysis are reviewed. The following topics are considered: (i) definition of probability distributions to characterize epistemic uncertainty in analysis inputs, (ii) generation of samples from uncertain analysis inputs, (iii) propagation of sampled inputs through an analysis, (iv) presentation of uncertainty analysis results, and (v) determination of sensitivity analysis results. Special attention is given to the determination of sensitivity analysis results, with brief descriptions and illustrations given for the following procedures/techniques: examination of scatterplots, correlation analysis, regression analysis, partial correlation analysis, rank transformations, statistical tests for patterns based on gridding, entropy tests for patterns based on gridding, nonparametric regression analysis, squared rank differences/rank correlation coefficient test, two-dimensional Kolmogorov–Smirnov test, tests for patterns based on distance measures, top down coefficient of concordance, and variance decomposition.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Aleatory uncertainty; Epistemic uncertainty; Latin hypercube sampling; Monte Carlo; Sensitivity analysis; Uncertainty analysis

1. Introduction

Uncertainty analysis and sensitivity analysis are essential parts of analyses for complex systems [1–14]. Specifically, uncertainty analysis refers to the determination of the uncertainty in analysis results that derives from uncertainty in analysis inputs, and sensitivity analysis refers to the determination of the contributions of individual uncertain analysis inputs to the uncertainty in analysis results. The uncertainty under consideration here is often referred to as epistemic uncertainty; alternative designations for this form of uncertainty include state of knowledge, subjective, reducible, and type B [15–24]. Epistemic uncertainty derives from a lack of knowledge about the appropriate value to use for a quantity that is assumed to have a fixed value in the context of a particular analysis. In the conceptual and computational organization of an analysis, epistemic uncertainty is generally considered to be distinct from aleatory uncertainty, which arises from an inherent randomness in the behavior of the system under study [15–24]. Alternative designations for aleatory uncertainty include variability, stochastic, irreducible, and type A.

A number of approaches to uncertainty and sensitivity analysis have been developed, including differential analysis [25–33], response surface methodology [34–43], Monte Carlo analysis [44–55], and variance decomposition procedures [56–60]. Overviews of these approaches are available in several reviews [61–68].

The focus of this presentation is on Monte Carlo (i.e., sampling-based) approaches to uncertainty and sensitivity analysis. Sampling-based approaches to uncertainty and sensitivity analysis are both effective and widely used [69–83]. Analyses of this type involve the generation and exploration of a mapping from uncertain analysis inputs to uncertain analysis results. The underlying idea is that analysis results $\mathbf{y}(\mathbf{x}) = [y_1(\mathbf{x}), y_2(\mathbf{x}), \ldots, y_{nY}(\mathbf{x})]$ are functions of uncertain analysis inputs $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]$. In

*Corresponding author. Sandia National Laboratories, Department 6849, MS 0779, Albuquerque, NM, 87185-0779, USA.
Tel.: +1 505 284 4808; fax: +1 505 844 2348.
E-mail address: jchelto@sandia.gov (J.C. Helton).

0951-8320/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2005.11.017

turn, uncertainty in **x** results in a corresponding uncertainty in **y(x)**. This leads to two questions: (i) what is the uncertainty in **y(x)** given the uncertainty in **x**? and (ii) how important are the individual elements of **x** with respect to the uncertainty in **y(x)**? The goal of uncertainty analysis is to answer the first question, and the goal of sensitivity analysis is to answer the second question. In practice, the implementation of an uncertainty analysis and the implementation of a sensitivity analysis are very closely connected on both a conceptual and a computational level.

The following sections summarize and illustrate the five basic components that underlie the implementation of a sampling-based uncertainty and sensitivity analysis: (i) definition of distributions $D_1, D_2, \ldots, D_{nX}$ that characterize the epistemic uncertainty in the elements $x_1, x_2, \ldots, x_{nX}$ of **x** (Section 2), (ii) generation of a sample $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{nS}$ from the **x**'s in consistency with the distributions $D_1, D_2, \ldots, D_{nX}$ (Section 3), (iii) propagation of the sample through the analysis to produce a mapping $[\mathbf{x}_i, \mathbf{y}(\mathbf{x}_i)]$, $i = 1, 2, \ldots, nS$, from analysis inputs to analysis results (Section 4), (iv) presentation of uncertainty analysis results (i.e., approximations to the distributions of the elements of **y** constructed from the corresponding elements of $\mathbf{y}(\mathbf{x}_i)$, $i = 1, 2, \ldots, nS$) (Section 5), and (v) determination of sensitivity analysis results (i.e., exploration of the mapping $[\mathbf{x}_i, \mathbf{y}(\mathbf{x}_i)]$, $i = 1, 2, \ldots, nS$) (Section 6). The presentation then ends with a concluding summary (Section 7).

Only probabilistic characterizations of uncertainty are considered in this presentation. Alternative uncertainty representations (e.g., evidence theory, possibility theory, fuzzy set theory, interval analysis) are active areas of research [84–92] but are outside the intended scope of this presentation.

2. Characterization of uncertainty

Definition of the distributions $D_1, D_2, \ldots, D_{nX}$ that characterize the epistemic uncertainty in the elements $x_1, x_2, \ldots, x_{nX}$ of **x** is the most important part of a sampling-based uncertainty and sensitivity analysis as these distributions determine both the uncertainty in **y** and the sensitivity of the elements of **y** to the elements of **x**. The distributions $D_1, D_2, \ldots, D_{nX}$ are typically defined through an expert review process [93–100], and their development can constitute a major analysis cost. A possible analysis strategy is to perform an initial exploratory analysis with rather crude definitions for $D_1, D_2, \ldots, D_{nX}$ and use sensitivity analysis to identify the most important analysis inputs; then, resources can be concentrated on characterizing the uncertainty in these inputs and a second presentation or decision-aiding analysis can be carried out with these improved uncertainty characterizations.

The scope of an expert review process can vary widely depending on the purpose of the analysis, the size of the analysis, and the resources available to carry out the analysis. At one extreme is a relatively small study in which a single analyst both develops the uncertainty characterizations (e.g., on the basis of personal knowledge or a cursory literature review) and carries out the analysis. At the other extreme, is a large analysis on which important societal decisions will be based and for which uncertainty characterizations are carried out for a large number of variables by teams of outside experts who support the analysts actually performing the analysis.

Given the breadth of analysis possibilities, it is beyond the scope of this presentation to provide an exhaustive review of how the distributions $D_1, D_2, \ldots, D_{nX}$ might be developed. However, as general guidance, it is best to avoid trying to obtain these distributions by specifying the defining parameters (e.g., mean and standard deviation) for a particular distribution type. Rather, distributions can be defined by specifying selected quantiles (e.g., 0.0, 0.1, 0.25, …, 0.9, 1.0) of the corresponding cumulative distribution functions (CDFs), which should keep the individual supplying the information in closer contact with the original sources of information or insight than is the case when a particular named distribution is specified (Fig. 1a). Distributions from multiple experts can be aggregated by averaging (Fig. 1b) [101].

This presentation draws most of its examples from an uncertainty and sensitivity analysis carried out for a two phase flow model (implemented in the BRAGFLO program) [102–104] in support of the 1996 Compliance Certification Application for the Waste Isolation Pilot Plant [105–107]. The uncertain variables considered in the example results (i.e., $x_1, x_2, \ldots, x_{nX}$ with $nX = 31$) and their associated distributions (i.e., $D_1, D_2, \ldots, D_{31}$) are summarized in Table 1. Additional information on the use of these variables in the two phase flow model and on the development of the associated uncertainty distributions is available in the original analysis documentation [102,108].

Additional information: Section 6.2, Refs. [46, 93–100,109–119]. As an example, Ref. [100] describes the approach used in the extensive expert review process that supported the US Nuclear Regulatory Commission's (NRC's) reassessment of the risk from commercial nuclear power plants (i.e., NUREG-1150; see Refs. [82,120–124]).

3. Generation of sample

Several sampling strategies are available, including random sampling, importance sampling, and Latin hypercube sampling [44,55] Latin hypercube sampling is very popular for use with computationally demanding models because its efficient stratification properties allow for the extraction of a large amount of uncertainty and sensitivity information with a relatively small sample size.

Latin hypercube sampling operates in the following manner to generate a sample of size nS from the distributions $D_1, D_2, \ldots, D_{nX}$ associated with the elements of $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]$. The range of each x_j is exhaustively divided into nS disjoint intervals of equal probability and one value x_{ij} is randomly selected from each interval. The nS values for x_1 are randomly paired without replacement

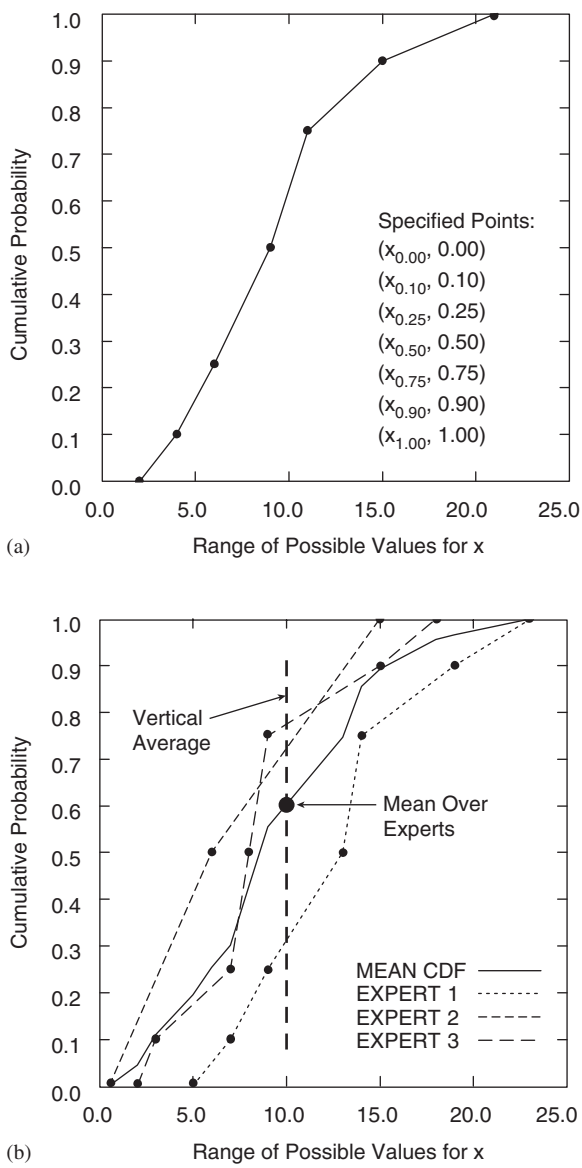

Fig. 1. Characterization of epistemic uncertainty: (a) construction of CDF from specified quantile values (Ref. [101, Fig. 4.1]), and (b) construction of mean CDF by vertical averaging of CDFs defined by individual experts with equal weight (i.e., $1/nE = 1/3$, where $nE = 3$ is the number of experts) given to each expert (Ref. [101, Fig. 4.2]).

with the nS value for x_2 to produce nS pairs. These pairs are then randomly combined without replacement with the nS values for x_3 to produce nS triples. This process is continued until a set of nS nX-tuples $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}]$, $i = 1, 2, \ldots, nS$, is obtained, with this set constituting the Latin hypercube sample (Fig. 2).

Latin hypercube sampling is a good choice for a sampling procedure when computationally demanding models are being studied. The popularity of Latin hypercube sampling recently led to the original article being designated a *Technometrics* classic in experimental design [125]. When the model is not computationally demanding, many model evaluations can be performed and random sampling works as well as Latin hypercube sampling.

If large sample sizes are required to provide appropriate coverage of low probability/high consequence subsets of values for $\mathbf{x}$, then importance sampling may be a more effective sampling procedure than either random or Latin hypercube sampling [126–134]. However, importance sampling complicates sensitivity analysis (Section 6) as the individual sample elements do not have equal weight (i.e., likelihood of occurrence). Often, some type of importance sampling is used to sample from aleatory uncertainty (e.g., possibly implemented through the use of event trees as is typically the case in probabilistic risk assessments for complex engineered facilities such as nuclear plants) and Latin hypercube sampling is used to sample from epistemic uncertainty. The NUREG-1150 analyses (see Refs. [82,120–124]) are an example of this approach to the propagation of uncertainty.

Control of correlations is an important aspect of sample generation. Specifically, correlated variables should have correlations close to their specified values, and uncorrelated variables should have correlations close to zero. In general, the imposition of complex correlation structures is not easy. However, Iman and Conover have developed a broadly applicable procedure to impose rank correlations on sampled values that (i) is distribution free (i.e., does not depend on the assumed marginal distributions for the sampled variables), (ii) can impose complex correlation structures involving multiple variables, (iii) works with both random and Latin hypercube sampling, and (iv) preserves the intervals used in Latin hypercube sampling [135,136]. Details on the implementation of the procedure are available in the original reference; [135] illustrative results are provided in Fig. 3 [137].

The analysis involving the variables in Table 1 used three independently generated (i.e., replicated) Latin hypercube samples of size $nS = 100$ each. The purpose of the replication was to provide a basis for testing the stability of uncertainty and sensitivity analysis results obtained with Latin hypercube sampling (Ref. [108, Sections 7 and 8]). The Iman/Conover restricted pairing technique indicated in the preceding paragraph was used to control correlations within the individual samples. The analyses with the three replicated samples were sufficiently similar that each analysis would have independently lead to the same insights with respect to model behavior [138]. However, to make full use of all model evaluations, final presentation results [103,104] were calculated with the three replicated samples pooled together to produce a single sample of size $nS = 300$.

Additional information: Ref. [46, Section 6.3], Refs. [44,50,54,55,139].

4. Propagation of sample through the analysis

Propagation of the sample through the analysis to produce the mapping $[\mathbf{x}_i, \mathbf{y}(\mathbf{x}_i)]$, $i = 1, 2, \ldots, nS$, from

Table 1
Uncertain variables $x_1, x_2, \ldots, x_{31}$ and associated uncertainty distributions $D_1, D_2, \ldots, D_{31}$ used in illustration of uncertainty and sensitivity analysis procedures for two phase flow model (Ref. [138, Table 1])

ANHBCEXP—Brooks–Corey pore distribution parameter for anhydrite (dimensionless). Distribution: Student's with 5 degree of freedom. Range: 0.491–0.842. Mean, median: 0.644.
ANHBCVGP—Pointer variable for selection of relative permeability model for use in anhydrite. Distribution: Discrete with 60% 0, 40% 1. Value of 0 implies Brooks–Corey model; value of 1 implies van Genuchten–Parker model.
ANHCOMP—Bulk compressibility of anhydrite (PA^{-1}). Distribution: Student's with 3 degrees of freedom. Range: 1.09×10^{-11}–$2.75 \times 10^{-10}\,Pa^{-1}$. Mean, median: $8.26 \times 10^{-11}\,Pa^{-1}$. Correlation: −0.99 rank correlation [23] with *ANHPRM*. Variable 21 in LHS.
ANHPRM—Logarithm of anhydrite permeability (m^2). Distribution: Student's with 5 degrees of freedom. Range: −21.0–−17.1 (i.e., permeability range is 1×10^{-21}–$1 \times 10^{-17.1}\,m^2$). Mean, median: −18.9. Correlation: −0.99 rank correlation with *ANHCOMP*.
ANRBRSAT—Residual brine saturation in anhydrite (dimensionless). Distribution: Student's with 5 degrees of freedom. Range: 7.85×10^{-3}–1.74×10^{-1}. Mean, median: 8.36×10^{-2}.
ANRGSSAT—Residual gas saturation in anhydrite (dimensionless). Distribution: Student's with 5 degrees of freedom. Range 1.39×10^{-2}–1.79×10^{-1}. Mean, median: 7.71×10^{-2}.
BHPRM—Logarithm of borehole permeability (m^2). Distribution: Uniform. Range: −14 to −11 (i.e., permeability range is 1×10^{-14}–$1 \times 10^{-11}\,m^2$). Mean, median: −12.5.
BPCOMP—Logarithm of bulk compressibility of brine pocket (Pa^{-1}). Distribution: triangular. Range: −11.3 to −8.00 (i.e., bulk compressibility range is $1 \times 10^{-11.3}$–$1 \times 10^{-8}\,Pa^{-1}$). Mean, mode: −9.80, −10.0. Correlation: −0.75 rank correlation with *BPPRM*.
BPINTPRS—Initial pressure in brine pocket (Pa). Distribution: triangular. Range: 1.11×10^7–1.70×10^7 Pa. Mean, mode: 1.36×10^7, 1.27×10^7 Pa.
BPPRM—Logarithm of intrinsic brine pocket permeability (m^2). Distribution: triangular. Range: −14.7 to −9.80 (i.e., permeability range is $1 \times 10^{-14.7}$ $-1 \times 10^{-9.80}\,m^2$). Mean, mode: −12.1, −11.8. Correlation: −0.75 rank correlation with *BPCOMP*.
BPVOL–Pointer variable for selection of brine pocket volume. Distribution: discrete, with integer values 1, 2, …, 32 equally likely.
HALCOMP—Bulk compressibility of halite (Pa−1). Distribution: Uniform. Range: 2.94×10^{-12}–$1.92 \times 10^{-10}\,Pa^{-1}$. Mean, median: $9.75 \times 10^{-11}\,Pa^{-1}$, $9.75 \times 10^{-11}\,Pa^{-1}$. Correlation: −0.99 rank correlation with *HALPRM*.
HALPOR—Halite porosity (dimensionless). Distribution: Piecewise uniform. Range: 1.0×10^{-3}–3×10^{-2}. Mean, median: 1.28×10^{-2}, 1.00×10^{-2}.
HALPRM–Logarithm of halite permeability (m^2). Distribution: Uniform. Range: −24 to −21 (i.e., permeability range is 1×10^{-24} to $1 \times 10^{-21}\,m^2$). Mean, median: −22.5, −22.5. Correlation: −0.99 rank correlation with *HALCOMP*.
SALPRES–Initial brine pressure, without the repository being present, at a reference point located in the center of the combined shafts at the elevation of the midpoint of MB 139 (Pa). Distribution: Uniform. Range: 1.104×10^7–1.389×10^7 Pa. 1.247×10^7 Pa.
SHBCEXP—Brooke-Corey pore distribution parameter for shaft (dimensionless). Distribution: Piecewise uniform. Range: 0.11–8.10. Mean, median: 2.52, 0.94.
SHPRMASP—Logarithm of permeability (m^2) of asphalt component of shaft seal (m^2). Distribution: Triangular. Range: −21–−18 (i.e., permeability range is 1×10^{-21}–$1 \times 10^{-18}\,m^2$). Mean, mode: −19.7, −20.0.
SHPRMCLY—Logarithm of permeability (m^2) for clay components of shaft. Distribution: Triangular. Range: −21–−17.3 (i.e., permeability range is 1×10^{-21}–$1 \times 10^{-17.3}\,m^2$). Mean, mode: −18.9, −18.3.
SHPRMCON—Same as *SHPRMASP* but for concrete component of shaft seal for 0 − 400 yr.Distribution: Triangular. Range: −17.0–−14.0 (i.e., permeability range is 1×10^{-17}–$1 \times 10^{-14}\,m^2$). Mean, mode: −15.3, −15.0.
SHPRMDRZ–Logarithm of permeability (m^2) of DRZ surrounding shaft. Distribution: Triangular. Range: −17.0–−14.0 (i.e., permeability range is 1×10^{-17}–$1 \times 10^{-14}\,m^2$). Mean, mode: −15.3, −15.0.
SHPRMHAL–Pointer variable (dimensionless) used to select permeability in crushed salt component of shaft seal at different times. Distribution: Uniform. Range: 0–1. Mean, mode: 0.5, 0.5. A distribution of permeability (m^2) in the crushed salt component of the shaft seal is defined for each of the following time intervals: [0, 10 yr], [10,25 yr], [25,50 yr], [50,100 yr], [100,200 yr], [200,10,000 yr]10,200. *SHPRMHAL* is used to select a permeability value from the cumulative distribution function for permeability for each of the preceding time intervals with result that a rank correlation of 1 exists between the permeabilities used for the individual time intervals.
SHRBRSAT—Residual brine saturation in shaft (dimensionless). Distribution: Uniform. Range: 0–0.4. Mean, median: 0.2, 0.2.
SHRGSSAT—Residual gas saturation in shaft (dimensionless). Distribution: Uniform. Range: 0–0.4. Mean, median: 0.2, 0.2.
WASTWICK—Increase in brine saturation of waste owing to capillary forces (dimensionless). Distribution: Uniform: Range: 0–1. Mean, median: 0.5, 0.5.
WFBETCEL—Scale factor used in definition of stoichiometric coefficient for microbial gas generation (dimensionless). Distribution: Uniform. Range: 0–1. Mean, median: 0.5, 0.5.
WGRCOR–Corrosion rate for steel under inundated conditions in the absence of CO_2 (m/s). Distribution: Uniform. Range: 0–1.58×10^{-14} m/s. Mean, median: 7.94×10^{-15} m/s, 7.94×10^{-15} m/s.
WGRMICH—Microbial degradation rate for cellulose under humid conditions (mol/kg s). Distribution: Uniform. Range: 0–1.27×10^{-9} mol/kg s. Mean, median: 6.34×10^{-10} mol/kg s, 6.34×10^{-10} mol/kg s.
WGRMICI—Microbial degradation rate for cellulose under inundated conditions (mol/kg s). Distribution: Uniform. Range: 3.17×10^{-10}–9.51×10^{-9} mol/kg s. Mean, median: 4.92×10^{-9} mol/kg s, 4.92×10^{-9} mol/kg s.
WMICDFLG–Pointer variable for microbial degradation of cellulose. Distribution: Discrete, with 50% 0, 25% 1, 25% 2, *WMICDFLG* = 0, 1, 2, implies no microbial degradation of cellulose, microbial degradation of only cellulose, microbial degradation of cellulose, plastic and rubber.
WRBRNSAT—Residual brine saturation in waste (dimensionless). Distribution: Uniform. Range: 0–0.552. Mean, median: 0.276, 0.276.
WRGSSAT—Residual gas saturation i n waste (dimensionless). Distribution: Uniform. Range: 0–0.15. Mean, median: 0.075, 0.075.

analysis inputs to analysis results is often the most computationally demanding part of a sampling-based uncertainty and sensitivity analysis. The details of this propagation are analysis specific and can range from very simple for analyses that involve a single model to very complicated for large analyses that involve complex systems of linked models [82,107].

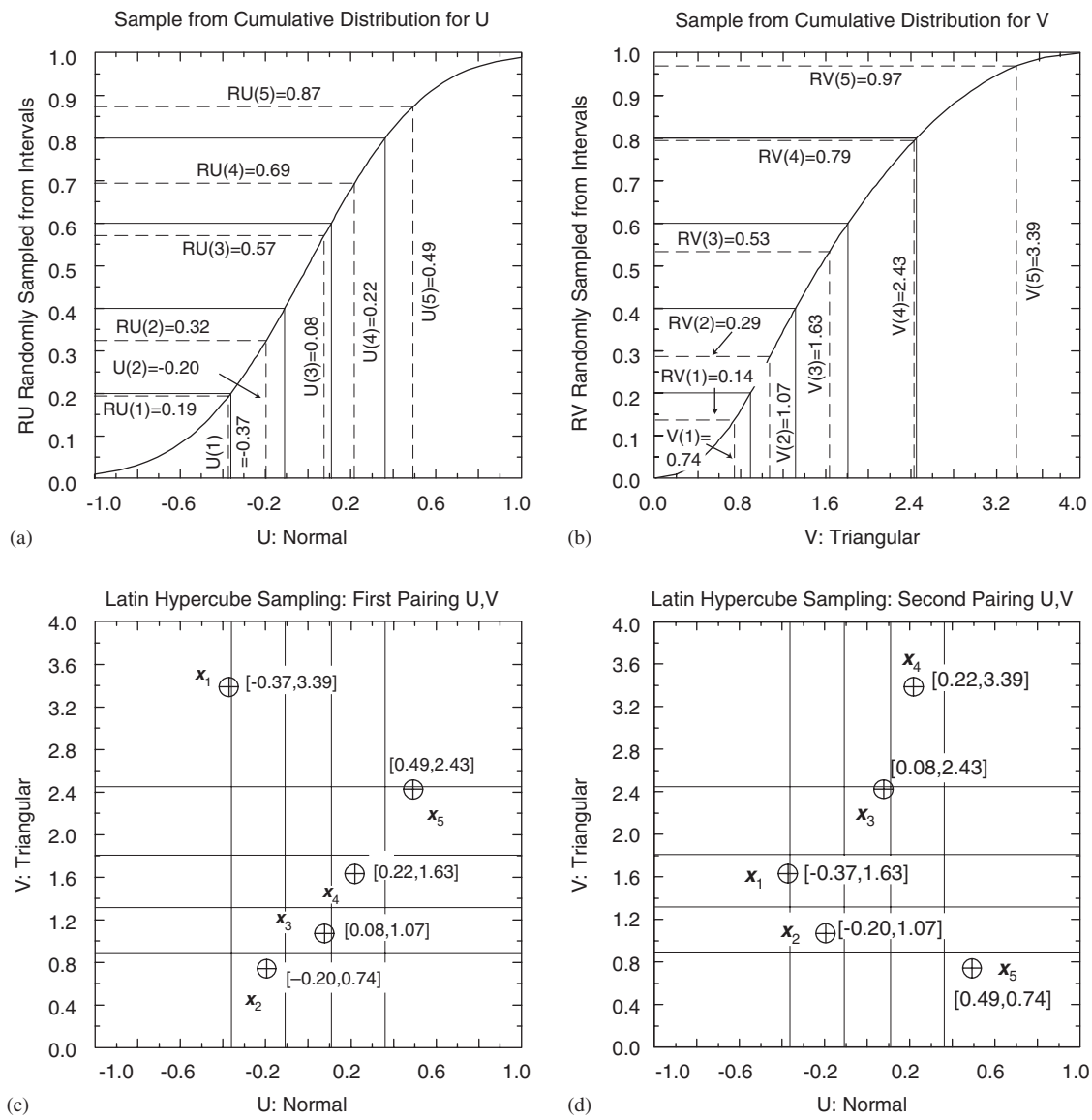

Fig. 2. Example of Latin hypercube sampling to generate a sample of size $nS = 5$ from $\mathbf{x} = [U, V]$ with U normal on $[-1, 1]$ (mean = 0.0; 0.01 quantile = −1; 0.99 quantile = 1) and V triangular on [0, 4] (mode = 1): (a, b) Upper frames illustrate sampling of values for U and V, and (c, d) Lower frames illustrate two different pairings of the sampled values of U and V in the construction of an LHS (Ref. [101, Fig. 5.3]).

When a single model is under consideration, this part of the analysis can involve little more than putting a DO loop around the model that (i) supplies the sampled input to the model, (ii) runs the model, and (iii) stores model results for later analysis. When more complex analyses with multiple models are involved, considerable sophistication may be required in this part of the analysis. Implementation of such analyses can involve (i) development of simplified models to approximate more complex models, (ii) clustering of results at model interfaces, (ii) reuse of model results through interpolation or linearity properties, and (iv) complex procedures for the storage and retrieval of analysis results.

Additional information: The NUREG-1150 analyses [82,93–100,109,119], the analyses carried out in support of the Compliance Certification Application for the Waste Isolation Pilot Plant [105–107], and analyses carried out in support of the Yucca Mountain Project's development of a facility for the deep geologic disposal of high level radioactive waste [140–142] provide examples of complex analyses that have used Latin hypercube sampling in the propagation of epistemic uncertainty.

5. Presentation of uncertainty analysis results

Presentation of uncertainty analysis results is generally straight forward and involves little more than displaying the results associated with the already calculated mapping $[\mathbf{x}_i,\ \mathbf{y}(\mathbf{x}_i)]$, $i = 1, 2, \ldots, nS$. Presentation possibilities include means and standard deviations, density functions, CDFs, complementary cumulative distribution functions (CCDFs), and box plots. Presentation formats such as

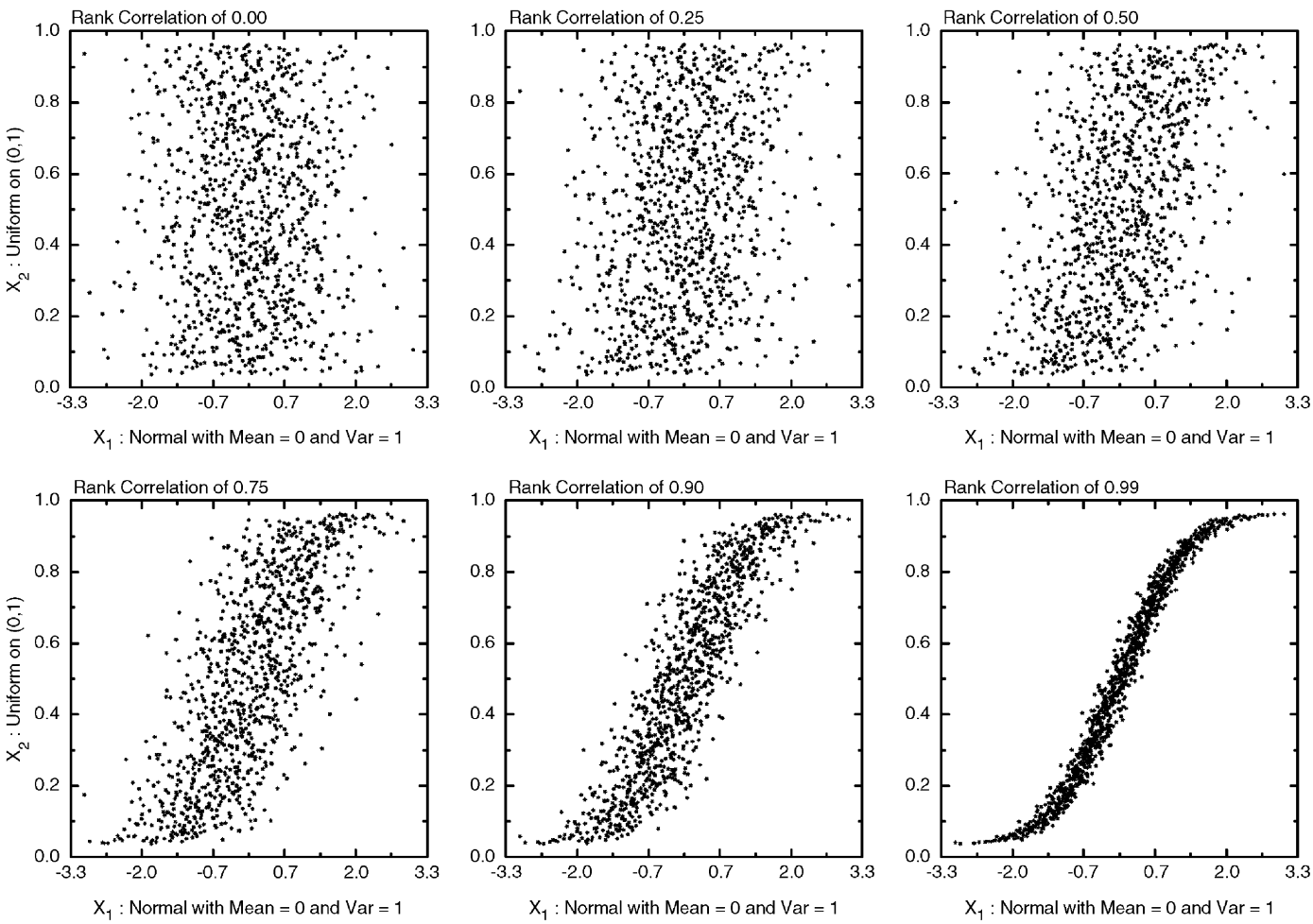

Fig. 3. Examples of rank correlations of 0.00, 0.25, 0.50, 0.75, 0.90 and 0.99 imposed with the Iman/Conover restricted pairing technique for an LHS of size $nS = 1000$ (Ref. [137, Fig. 5.1]).

CDFs (Fig. 4a), CCDFs (Fig. 4a) and box plots (Fig. 4b) are usually preferable to means and standard deviations because of the large amount of uncertainty information that is lost in the calculation of means and standard deviations (see Table 2 for definitions of dependent variables used to illustrate uncertainty and sensitivity analysis procedures). Owing to their flattened shape, box plots are particularly useful when it is desired to the display and compare the uncertainty in a number of related variables.

The representational challenge is more complex when the analysis outcome of interest is a function rather than a scalar. For example, time-dependent system properties are common analysis outcomes. As another example, a CCDF that summarizes the effects of aleatory uncertainty is a standard analysis outcome in risk assessments. An effective display format for such analysis outcomes is to use two plot frames, with first frame displaying the analysis results for the individual sample elements and the second frame displaying summary results for the outcomes in the first frame (e.g., quantiles and means) (Fig. 5).

Additional information: Ref. [46, Section 6.4]; Refs. [143,144].

6. Determination of sensitivity analysis results

Determination of sensitivity analysis results is usually more demanding than the presentation of uncertainty analysis results due to the need to actually explore the mapping $[\mathbf{x}_i, \mathbf{y}(\mathbf{x}_i)]$, $i = 1, 2, \ldots, nS$, to assess the effects of individual elements of $\mathbf{x}$ on the elements of $\mathbf{y}$. A number of approaches to sensitivity analysis that can be used in conjunction with a sampling-based uncertainty analysis are briefly summarized in this section. In this summary, (i) x_j is an element of $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]$, (ii) y is an element of $\mathbf{y}(\mathbf{x}) = [y_1(\mathbf{x}), y_2(\mathbf{x}), \ldots, y_{nY}(\mathbf{x})]$, (iii) $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}]$, $i = 1, 2, \ldots, nS$, is a random or Latin hypercube sample from the possible values for $\mathbf{x}$ generated in consistency with the joint distribution assigned to the x_j's, (iv) $\mathbf{y}_i = \mathbf{y}(\mathbf{x}_i)$ for $i = 1, 2, \ldots, nS$, and (v) x_{ij} and y_i are elements of $\mathbf{x}_i$ and $\mathbf{y}_i$, respectively. Sensitivity analyses usually consider the effects of all elements of $\mathbf{x}$ on individual elements of $\mathbf{y}$;

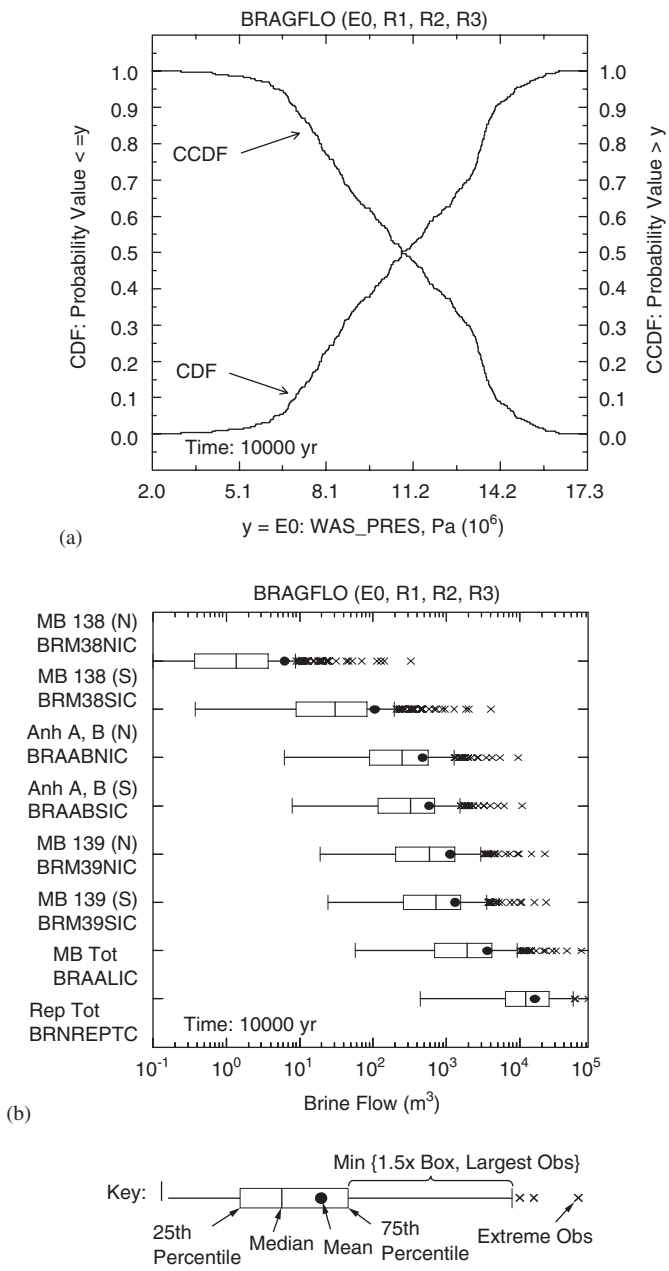

Fig. 4. Representation of uncertainty in scalar-valued analysis results: (a) CDFs and CCDFs (Ref. [101, Fig. 7.2]) and (b) box plots (Ref. [101, Fig. 7.4]).

for this reason and for notational simplification, the subscripted variables x_j, $j = 1, 2, \ldots, nX$, are used to represent the elements of $\mathbf{x}$ but the unsubscripted variable y is used to represent an arbitrary element of $\mathbf{y}$.

6.1. Scatterplots

A plot of the points $[x_{ij}, y_i]$ for $i = 1, 2, \ldots, nS$ (i.e., a scatterplot of y versus x_j) can reveal nonlinear or other unexpected relationships between analysis inputs and analysis results (Fig. 6). Scatterplots are a natural starting point in a complex analysis that can help in the development of a sensitivity analysis strategy using one or more additional techniques. Often, the examination of scatterplots is all that is needed to understand the relationships between the uncertainty in analysis inputs and the uncertainty in analysis results [145].

Most analyses start with two dimensional scatterplots. However, when strong three-way interactions between variables are present, three-dimensional scatterplots (i.e., scatterplots involving three variables) can provide informative displays of analysis results (Fig. 7). The three-dimensional scatterplot in Fig. 7 involves one sampled variable (i.e., $x_j = WPRTDIAM$) and two calculated variables (i.e., $y_k = WAS_PRES$ and $y_l = REL_VOL$). The result in Fig. 7 was calculated by a model that uses the calculated value for WAS_PRES under undisturbed conditions as an input and then determines the volume of material (i.e., REL_VOL) released to the surface at the time of a drilling intrusion due to a pressure-driven spallings event; $WPRTDIAM$ is one of the uncertain (i.e., sampled) variables used in this calculation [145] Specifically, Fig. 7 contains a plot of the points (x_{ij}, y_{ik}, y_{il}) for $i = 1, 2, \ldots, nS$. As examination of Fig. 7 shows, (i) WAS_PRES acts as a switch that determines if REL_PRES is nonzero, and (ii) $WPRTDIAM$ determines the magnitude of the nonzero values for REL_PRES. Because of the large number of possible three-way variable combinations in most analyses, some initial insights with respect to variable interactions usually needs to be developed before a reasonable selection of three-dimensional scatterplots can be made.

Additional information: Ref. [46, Section 6.6.1]; see Ref. [146] for additional plotting formats, including cobweb plots which provide a representation of multi-dimensional results (e.g., $[\mathbf{x}_i, y_i] = [x_{i1}, x_{i2}, \ldots, x_{i,nX}, y_i]$, $i = 1, 2, \ldots, nS$) in a two-dimensional plot.

6.2. Correlation

Correlation provides a measure of the strength of the linear relationship between x_j and y. Specifically, the (Pearson or sample) correlation coefficient (CC) $c(x_j, y)$ between x_j and y is defined by

$$c(x_j, y) = \frac{\sum_{i=1}^{nS}(x_{ij} - \bar{x}_j)(y_i - \bar{y})}{\left[\sum_{i=1}^{nS}(x_{ij} - \bar{x}_j)^2\right]^{1/2}\left[\sum_{i=1}^{nS}(y_i - \bar{y})^2\right]^{1/2}}, \tag{1}$$

where

$$\bar{x}_j = \sum_{i=1}^{nS} x_{ij}/nS \text{ and } \bar{y} = \sum_{i=1}^{nS} y_i/nS.$$

The CC $c(x_j, y)$ has a value between -1 and 1, with a positive value indicating that x_j and y tend to increase and decrease together and a negative value indicating that x_j and y tend to move in opposite directions. Further, gradations in the absolute value of $c(x_j, y)$ between 0 and 1 correspond to a trend from no linear relationship between x_j and y to an exact linear relationship between

Table 2
Definition of dependent variables calculated by BRAGFLO program for two phase flow and used in the illustration of uncertainty and sensitivity analysis procedures

BNBHDNUZ—Cumulative brine flow (m^3) down borehole at Market Bed (MB) 138 (i.e., from cell 223 to cell 575 in Ref. [102, Fig. 3]).
BRAABNIC—Cumulative brine flow (m^3) out of north anhydrites A and B into disturbed rock zone (DRZ) (i.e., from cells 556 to cell 527 in Ref. [102, Fig. 3]).
BRAABSIC—Cumulative brine flow (m^3) out of south anhydrites A and B into DRZ (i.e., from cell 555 to cell 482 in Ref. [102, Fig. 3]).
BRAALIC—Cumulative brine flow (m^3) out of all MBs into DRZ (i.e.,
BRAALIC = *BRM38NIC*+*BRAABNIC*+*BRM39NIC*+*BRM38SIC*+*BRAABSIC*+*BRM39SIC*).
BRM38NIC—Cumulative brine flow (m^3) out of north MB138 into DRZ (i.e., from cells 588 to 587 in Ref. [102, Fig. 3]).
BRM38SIC—Cumulative brine flow (m^3) out of south MB138 into DRZ (i.e., from cell 571 to cell 572 in Ref. [102, Fig. 3]).
BRM39NIC–Cumulative brine flow (m^3) out of north MB139 to DRZ (i.e., from cells 540 to 465 in Ref. [102, Fig. 3]).
BRM39SIC—Cumulative brine flow (m^3) out of south MB139 into DRZ (i.e., from cell 539 to cell 436 in Fig. 3, Ref. [102]).
BRNREPTC—Cumulative brine flow (m^3) into repository (i.e., into regions corresponding to cells 596–625, 638–640 in Ref. [102, Fig. 3]).
REP_SATB—Brine saturation in upper waste panels (i.e., average brine saturation calculated over cells 617–625 in Ref. [102, Fig. 3]).
WAS_PRES—Pressure (Pa) in lower waste panel (i.e., average pressure calculated over cells 596–616 in Ref. [102, Fig. 3]).
WAS_SATB—Brine saturation in lower waste panel (i.e., average brine saturation calculated over cells 596–616 in Ref. [102, Fig. 3])

The designator E0 is used to indicate results calculated for undisturbed conditions, and the designator E2 is used to indicate results calculated for disturbed conditions due to a drilling intrusion that penetrates the lower waste panel of the repository 1000 yr after repository closure. Further, the designator *R*1 indicates results calculated for the first of the three replicated Latin hypercube samples described in Section 3, and the designators *R*1, *R*2, *R*3 collectively are used to indicate results calculated with the three replicates pooled together.

x_j and y. As an example, the CCs associated with the scatterplots in Fig. 8 are $c(HALPOR,\ REP_SATB) = 0.75$ (Fig. 8a) and $c(WGRCOR,\ REP_SATB) = -0.41$ (Fig. 8b).

The CC $c(x_j,\ y)$ is closely related to results obtained in a linear regression relating y to x_j. Specifically, $c(x_j,\ y)$ is equal to the standardized regression coefficient (SRC) in the indicated regression, and the absolute value of $c(x_j,\ y)$ is equal to the square root of the corresponding R^2 value (see Section 6.3). As a correlation of 0 only indicates the absence of a linear association between x_j and y, it does not preclude the existence of a well-defined nonlinear relationship between x_j and y (e.g., $y = \sin x_j$).

Additional information: Ref. [46, Section 6.6.4].

6.3. Regression analysis

Regression analysis provides an algebraic representation of the relationships between y and one or more of the x_j's. Unless stated otherwise, regression analysis is usually assumed to involve the construction of linear models of the form

$$\hat{y} = b_0 + b_j x_j \tag{2}$$

for a single independent variable (i.e., x_j) and

$$\hat{y} = b_0 + \sum_{j=1}^{nX} b_j x_j \tag{3}$$

for multiple independent variables (i.e., $x_1,\ x_2,\ \ldots,\ x_{nX}$). The regression coefficients in Eqs. (2) and (3) are determined such that the sums

$$\sum_{i=1}^{nS} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{nS} [y_i - (b_0 + b_j x_{ij})]^2 \tag{4}$$

and

$$\sum_{i=1}^{nS} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{nS} \left[y_i - \left(b_0 + \sum_{j=1}^{nS} b_j x_{ij} \right) \right]^2, \tag{5}$$

respectively, are minimized. As a result, the regression models in Eqs. (2) and (3) are often referred to as least-squares models due to the minimization of the sums of squares in Eqs. (4) and (5).

As important property of least squares regression models is the equality

$$\sum_{i=1}^{nS} (y_i - \bar{y})^2 = \sum_{i=1}^{nS} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{nS} (\hat{y}_i - y_i)^2. \tag{6}$$

For notational convenience, the preceding equality is often written

$$\mathrm{SS_{tot}} = \mathrm{SS_{reg}} + \mathrm{SS_{res}}, \tag{7}$$

where

$$\mathrm{SS_{tot}} = \sum_{i=1}^{nS} (y_i - \bar{y})^2, \quad \mathrm{SS_{reg}} = \sum_{i=1}^{nS} (\hat{y}_i - \bar{y})^2,$$

$$\mathrm{SS_{res}} = \sum_{i=1}^{nS} (\hat{y}_i - y)^2$$

and the three preceding summations are called the total sum of squares ($\mathrm{SS_{tot}}$), regression sum of squares ($\mathrm{SS_{reg}}$) and residual sum of squares ($\mathrm{SS_{res}}$), respectively.

Since $\mathrm{SS_{res}}$ provides a measure of variability about the regression model, the ratio

$$R^2 = \mathrm{SS_{reg}}/\mathrm{SS_{tot}} = \sum_{i=1}^{nS} (\hat{y}_i - \bar{y})^2 \Big/ \sum_{i=1}^{nS} (y_i - \bar{y})^2 \tag{8}$$

provides a measure of the extent to which the regression model can match the observed data. Specifically, when the

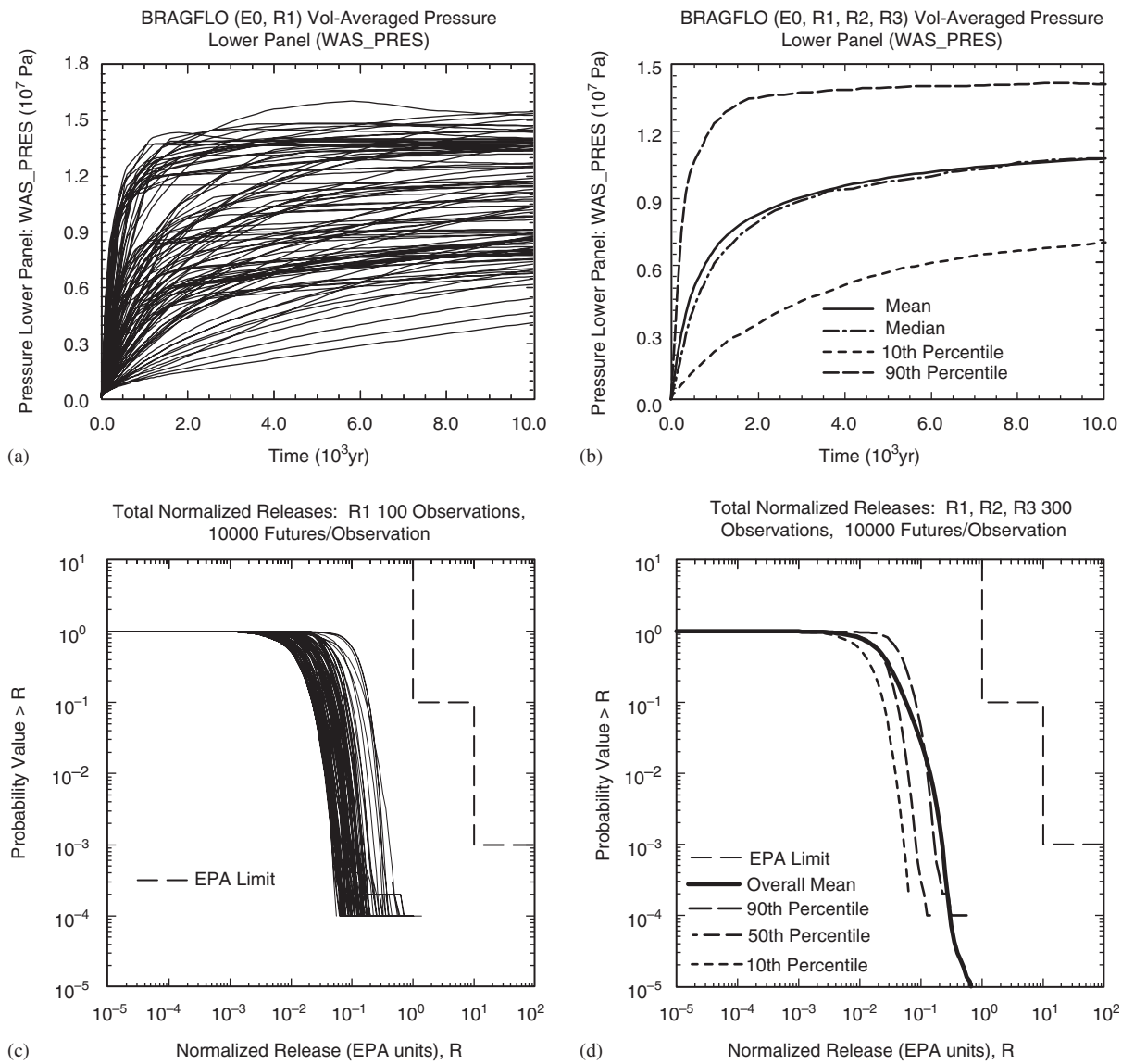

Fig. 5. Representation of uncertainty in analysis results that are functions: (a, b) Pressure as a function of time (Ref. [101, Figs. 7.5, 7.9]), and (c, d) effects of aleatory uncertainty summarized as a CCDF (Ref. [101, Fig. 10.5]).

variation about the regression model is small (i.e., SS_{res} is small relative to SS_{reg}), then the corresponding R^2 value is close to 1, which indicates that the regression model is accounting for most of the uncertainty in y. Conversely, an R^2 value close to 0 indicates that the regression model is not very successful in accounting for the uncertainty in y. When the individual x_j in the regression model in Eq. (3) are independent, the R^2 value for the regression model can be expressed as

$$R^2 = SS_{reg}/SS_{tot} = R_1^2 + R_2^2 + \cdots + R_{nX}^2, \tag{9}$$

where R_j^2 is the R^2 value that results from regressing y on only x_j. Thus, R_j^2 is equal to the contribution of x_j to the R^2 value for the regression model in Eq. (3) when the x_j's are independent.

The regression coefficients b_j, $j = 1, 2, \ldots, nX$, are not very useful in sensitivity analysis because each b_j is influenced by the units in which x_j is expressed and also does not incorporate any information on the distribution assigned to x_j. Because of this, the regression models in Eqs. (2) and (3) are usually reformulated as

$$(\hat{y} - \bar{y})/\hat{s} = (b_j \hat{s}_j/\hat{s})(x_j - \bar{x}_j)/\hat{s}_j \tag{10}$$

and

$$(\hat{y} - \bar{y})/\hat{s} = \sum_{j=1}^{nX} (b_j \hat{s}_j/\hat{s})(x_j - \bar{x}_j)/\hat{s}_j, \tag{11}$$

respectively, where

$$\hat{s} = \left[\sum_{i=1}^{nS} (y_i - \bar{y})^2/(nS-1)\right]^{1/2},$$

$$\hat{s}_j = \left[\sum_{i=1}^{nS} (x_{ij} - \bar{x}_j)^2/(nS-1)\right]^{1/2}$$

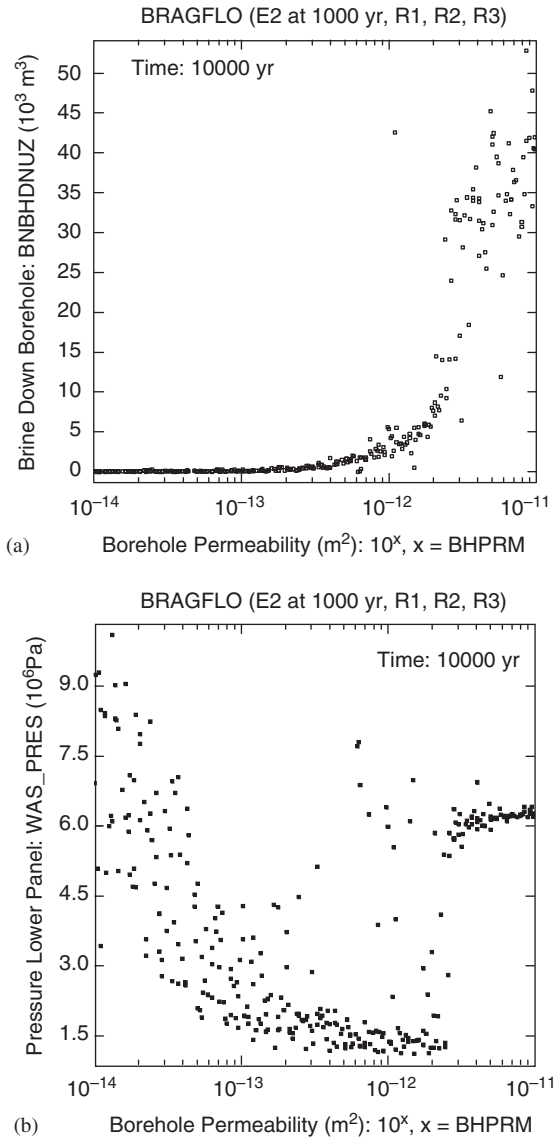

Fig. 6. Examples of scatterplots obtained in a sampling-based uncertainty/sensitivity analysis (Ref. [101, Figs. 8.1, 8.2]).

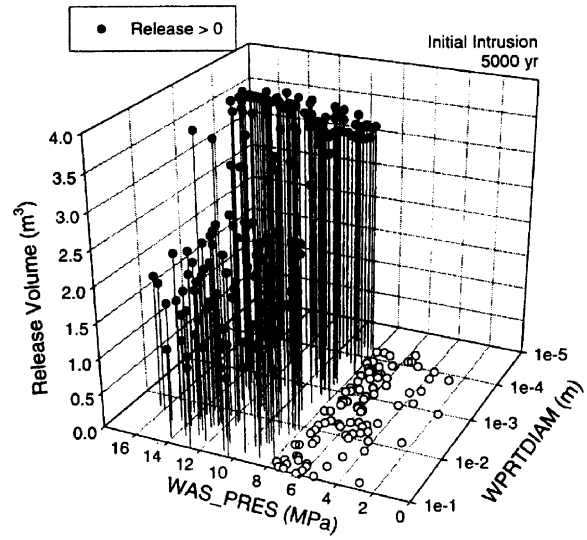

Fig. 7. Example of three-dimensional scatterplot obtained in a sampling-based uncertainty/sensitivity analysis (Ref. [145, Fig. 13]).

and $\bar{y}$ and $\bar{x}_j$ are defined in conjunction with Eq. (1). The coefficients $b_j\hat{s}_j/\hat{s}$ in Eqs. (10) and (11) are referred to as SRCs.

When the regression models in Eqs. (2) and (10) involving only x_j are under consideration, the SRC $b_j\hat{s}_j/\hat{s}$ provides a measure of variable importance based on the effect on y relative to the standard deviation $\hat{s}$ of y of moving x_j away from its expected value $\bar{x}_j$ by a fixed fraction of its standard deviation $\hat{s}_j$. Further, when the x_j's are independent, the inclusion or exclusion of an individual x_j from the regression models in Eqs. (3) and (11) has no effect on the SRCs for the remaining variables in the model. Thus, as long as the x_j's are independent, the SRCs $b_j\hat{s}_j/\hat{s}$ in Eq. (11) provide a useful measure of variable importance, with (i) the absolute values of the coefficients $b_j\hat{s}_j/\hat{s}$ providing a comparative measure of variable importance (i.e., variable x_u is more important than variable x_v if $|b_u\hat{s}_u/\hat{s}| > |b_v\hat{s}_v/\hat{s}|$) and (ii) the sign of $b_j\hat{s}_j/\hat{s}$ indicating whether x_j and y tend to move in the same direction or in opposite directions. However, when x_j's are not independent, SRCs do not provide reliable indications of variable importance (Ref. [46, Section 6.6.7]).

For purposes of sensitivity analysis, there is usually no reason to construct a regression model containing all the uncertain variables (i.e., x_1, x_2, …, x_{nX}) as indicated in Eqs. (3) and (11). Rather, a more appropriate procedure is to construct regression models in a stepwise manner. With this procedure, a regression model is first constructed with the most influential variable (e.g., $\tilde{x}_1$ as determined based on R^2 values for regression models containing only single variables). Then, a regression model is constructed with $\tilde{x}_1$ and the next most influential variable (e.g., $\tilde{x}_2$ as determined based on R^2 values for regression models containing $\tilde{x}_1$ and each of the remaining variables). The process then repeats to determine $\tilde{x}_3$ in a similar manner and continues until no more variables with an identifiable effect on y can be found. Variable importance (i.e., sensitivity) is then indicated by the order in which variables are selected in the stepwise process, the changes in cumulative R^2 values as additional variables are added to the regression model, and the SRCs for the variables in the final regression model. An example of a sensitivity analysis of this form is presented in Table 3.

A display of regression results of the form shown in Table 3 is very unwieldy when results at a sequence of times are under consideration. In this situation, a more compact display of regression results is provided by plotting SRCs as functions of time for all x_j that appear to have a significant effect on y at some point in the time interval under consideration (Fig. 9a).

This section only considers linear regression models. However, linear regression models also include models of

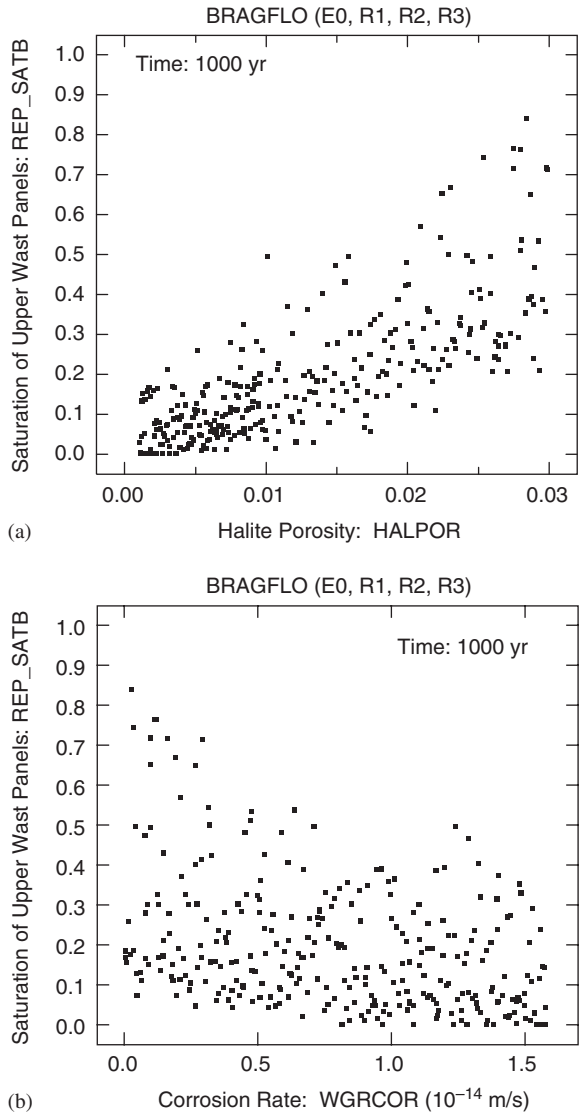

Fig. 8. Illustration of correlation coefficients: (a) $c(x_j, y) = 0.75$ with $x_j = HALPOR$ and $y = REP_SATB$ (left frame), and (b) $c(x_j, y) = -0.41$ with $x_j = WGRCOR$ and $y = REP_SATB$ (right frame).

Table 3
Example of stepwise regression analysis to identify uncertain variables affecting the uncertainty in pressure (*WAS_PRES*) at 10,000 yr in Fig. 5a (Ref. [101, Table 8.6])

Step[a]	Variable[b]	SRC[c]	R^{2d}
1	*WMICDFLG*	0.718	0.508
2	*HALPOR*	0.466	0.732
3	*WGRCOR*	0.246	0.792
4	*ANHPRM*	0.129	0.809
5	*SHRGSSAT*	0.070	0.814
6	*SALPRES*	0.063	0.818

[a]Steps in stepwise regression analysis.
[b]Variables listed in the order of selection in regression analysis.
[c]SRCs for variables in final regression model.
[d]Cumulative R^2 value with entry of each variable into regression model.

forms such as

$$\hat{y} = b_0 + \sum_{j=1}^{nX} b_j f_j(x_j) + \sum_{j=1}^{nX} \sum_{l=j}^{nX} b_{jl} f_{jl}(x_j, x_l). \tag{12}$$

This inclusion exists because the preceding model is linear in its coefficients (i.e., b_0, the b_j, the b_{jl}); in essence, the indicated transformations involving the x_j (i.e., $f_j(x_j)$, $f_{jl}(x_j, x_l)$) are simply defining a new set of analysis inputs to be used in a regression-based sensitivity analysis. Results can be improved in some analyses by well-chosen variable transformations of the form indicated in Eq. (12). However, in large analyses involving many uncertain analysis inputs (i.e., x_j) and many possibly time-dependent analysis results (i.e., many different elements of **y**), the a priori determination of suitable transformations can be difficult. Also, care can be taken to suitably account for any correlations that may be introduced by the chosen transformations (i.e., $f_j(x_j)$ and $f_{jl}(x_j, x_l)$ may be highly correlated).

Nonlinear regression provides an alternative to linear regression that can be useful in some analyses. In nonlinear regression, at least some of the model coefficients are operated on by nonlinear functions. For example,

$$\hat{y} = b_0 + b_1 \exp(b_2 x_1) + b_3 \sin(b_4 x_2) \tag{13}$$

is a nonlinear model because b_2 and b_4 appear in expressions that are operated on by nonlinear functions. A major challenge in the use of nonlinear regression in sensitivity analysis is the determination of a suitable form for the nonlinear regression model. The following two alternatives to nonlinear regression for use in the presence of nonlinear relationships between model inputs (i.e., the x_j) and model results (i.e., the elements of **y**) that place fewer a priori demands on the analyst are described later in this presentation: rank transformations (Section 6.5) and nonparametric regression (Section 6.8).

Additional information: Ref. [46, Sections 6.6.2, 6.6.3, and 6.6.5]. Further, general information on regression analysis is available in a number of texts (e.g., Refs. [147–151]).

6.4. Partial correlation

The partial correlation coefficient (PCC) between x_j and y can be defined in the following manner. First, the two regression models indicated below are constructed:

$$\hat{x}_j = c_0 + \sum_{\substack{p=1 \\ p \neq j}}^{nX} c_p x_p \quad \text{and} \quad \hat{y} = b_0 + \sum_{\substack{p=1 \\ p \neq j}}^{nX} b_p x_p. \tag{14}$$

Then, the results of the two preceding regressions are used to define the new variables $x_j - \hat{x}_j$ and $y - \hat{y}$. The PCC between x_j and y is the CC $c(x_j - \hat{x}_j, y - \hat{y})$ (see Eq. (1)) between $x_j - \hat{x}_j$ and $y - \hat{y}$. As for SRCs, PCCs are often defined for variables that are functions of time and presented as time-dependent plots (Fig. 9b).

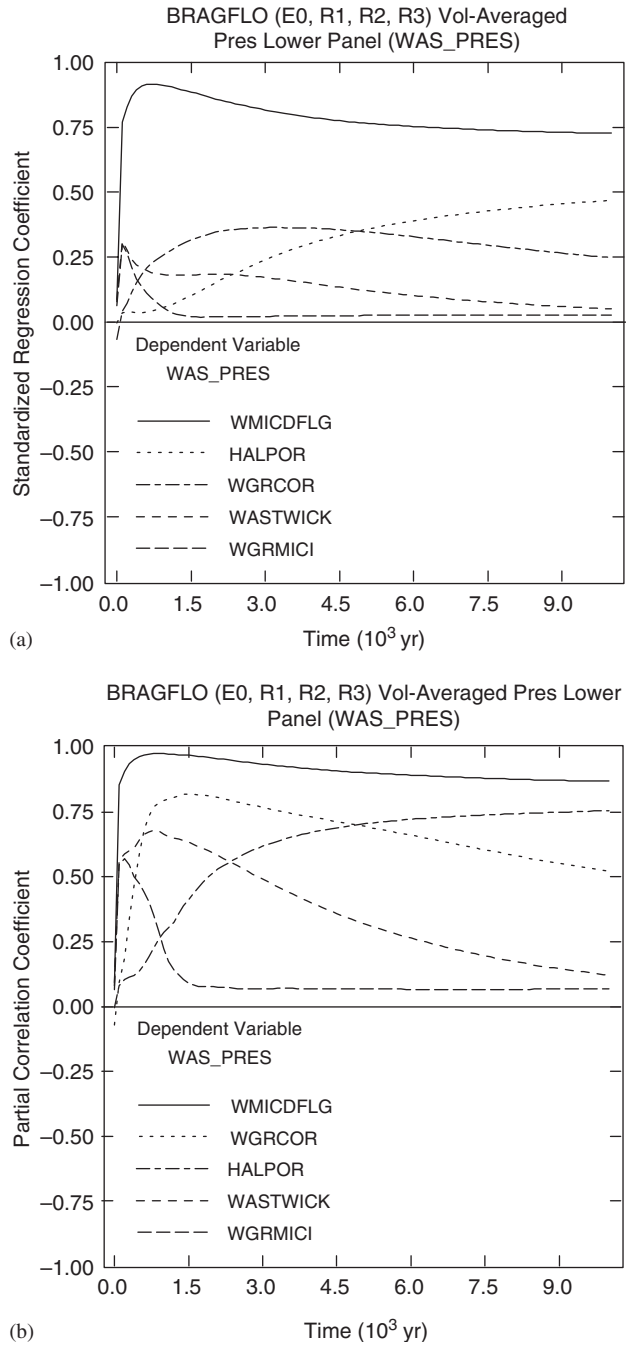

Fig. 9. Time-dependent sensitivity analysis results for uncertain pressure curves in Fig. 5a: (a) SRCs as a function of time, and (b) PCCs as a function of time (Ref. [101, Fig. 8.3]).

The PCC characterizes the linear relationship between x_j and y after a correction has been made for the linear effects on y of the remaining elements of $\mathbf{x}$, and the SRC characterize the effect on y that results from perturbing x_j by a fixed fraction of its standard deviation. Thus, PCCs and SRCs provide related, but not identical, measures of variable importance. In particular, the PCC between x_j and y provides a measure of variable importance that tends to exclude the effects of the other elements of $\mathbf{x}$, the assumed distribution for x_j, and the magnitude of the impact of the uncertainty in x_j on the uncertainty in y. In contrast, the SRC relating x_j to y is more influenced by the distribution assigned to x_j and the magnitude of the impact of the uncertainty in x_j on the uncertainty in y. However, when the elements of $\mathbf{x}$ are independent, PCCs and SRCs give the same rankings of variable importance. Specifically, an ordering of variable importance based on the absolute value of PCCs is the same as an ordering based on either the absolute value of CCs or the absolute value of SRCs (Ref. [46, Section 6.6.4]). A cosmetic benefit of using PCCs is that PCCs tend to be spread out in value more than SRCs and thus produce results that are easier to read (e.g., compare Figs. 9a and b); however, the downside to this is that a variable can appear to have a larger effect on the uncertainty in y than is actually the case.

As for analyses based on SRCs, analyses based on PCCs can give very misleading results when correlations exist between the elements of $\mathbf{x}$. Specifically, if $\mathbf{x}$ contains two highly correlated variables, then each variable will cancel the other's effect when PCCs with y are calculated.

Additional information: Ref. [46, Section 6.6.4]; Ref. [152].

6.5. Rank transformations

A rank transformation can be used to convert a nonlinear but monotonic relationship between the x_j and y into a linear relationship. With this transformation, the values for the x_j and y are replaced by their corresponding ranks. Specifically, the smallest value for a variable is assigned a rank of 1; the next largest value is assigned a rank of 2; tied values are assigned their average rank; and so on up to the largest value, which is assigned a rank of nS. Use of the rank transformation results in rank (i.e., Spearman) correlation coefficients (RCCs), rank regressions, standardized rank regression coefficients (SRRCs) and partial rank correlation coefficients (PRCCs). In the presence of nonlinear but monotonic relationships between the x_j and y, use of the rank transform can substantially improve the resolution of sensitivity analysis results (Table 4).

Additional information: Ref. [46, Section 6.6.6]; Ref. [153].

6.6. Statistical tests for patterns based on gridding

Analyses based on raw or rank-transformed data can fail when the underlying relationships between the x_j and y are nonlinear and nonmonotonic (Fig. 10). The scatterplot in Fig. 6b is for the pressure at 10,000 yr in Fig. 10a versus the uncertain variable *BHPRM*. The partial correlation analyses summarized in Fig. 10b fail at later times because the pattern appearing in Fig. 6b is too complex to be captured with a partial correlation analysis based on raw or rank-transformed data; analyses with SRCs or SRRCs also fail for the same reason. An alternative analysis strategy for

Table 4
Comparison of stepwise regression analyses with raw and rank-transformed data for cumulative brine inflow to vicinity of repository over 10,000 yr from anhydrite marker beds (*BRAALIC*) under undisturbed (i.e., E0) conditions in Fig. 4b (Ref. [101, Table 8.8])

Step[a]	Raw data			Rank-transformed data		
	Variable[b]	SRC[c]	R^{2d}	Variable[b]	SRRC[e]	R^{2d}
1	*ANHPRM*	0.562	0.320	*WMICDFLG*	−0.656	0.425
2	*WMICDFLG*	−0.309	0.423	*ANHPRM*	0.593	0.766
3	*WGRCOR*	−0.164	0.449	*HALPOR*	−0.155	0.802
4	*WASTWICK*	−0.145	0.471	*WGRCOR*	−0.152	0.824
5	*ANHBCEXP*	−0.120	0.486	*HALPRM*	0.143	0.845
6	*HALPOR*	−0.101	0.496	*SALPRES*	0.120	0.860
7				*WASTWICK*	−0.010	0.869

[a]Steps in stepwise regression analysis.
[b]Variables listed in order of selection in regression analysis.
[c]SRCs for variables in final regression model.
[d]Cumulative R^2 value with entry of each variable into regression model.
[e]SRRCs for variables in final regression model.

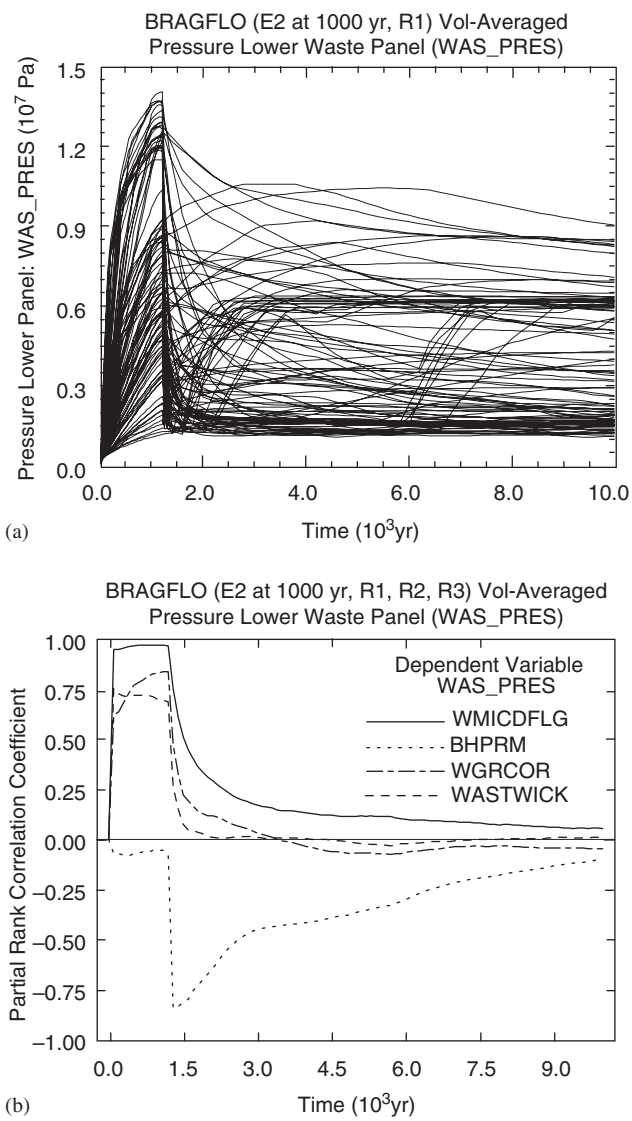

Fig. 10. Illustration of failure of a sensitivity analysis based on rank-transformed data: (a) pressures as a function of time and (b) PRCCs as a function of time (Ref. [101, Fig. 8.7]).

situations of this type is to place grids on the scatterplot for y and x_j and then perform various statistical tests to determine if the distribution of points across the grid cells appears to be nonrandom. Appearance of a nonrandom pattern indicates that x_j has an effect on y. Possibilities include tests for (i) common means (CMNs), (ii) common distributions or locations (CLs), (iii) common medians (CMDs), and (iv) statistical independence (SI). Descriptions of these tests follow.

The CMNs test is based on dividing the values of x_j (i.e., x_{ij}, $i = 1, 2, \ldots, nS$) into nI classes and then testing to determine if y has a CMN across these classes (Ref. [154, Section 3.1]). The required classes are obtained by dividing the range of x_j into a sequence of mutually exclusive and exhaustive subintervals containing equal numbers of sampled values (Fig. 11a). If x_j is discrete, individual classes are defined for each of the distinct values. For notational convenience, let c, $c = 1, 2, \ldots, nI$, designate the individual classes into which the values of x_j have been divided; let X_c designate the set such that $i \in \mathrm{X}_c$ only if x_{ij} belongs to class c; and let nI_c equal the number of elements contained in X_c (i.e., the number of x_{ij}'s associated with class c).

The F-test can be used to test for the equality of the mean values of y for the classes into which the values of x_j have been divided (e.g., the intervals defined on the abscissa of the scatterplot in Fig. 11a). Specifically, if the y values conditional on each class of x_j values are normally distributed with equal expected values, then

$$F = \frac{\left[\sum_{c=1}^{nI} nI_c \bar{y}_c^2 - nS\bar{y}^2\right] \Big/ (nI - 1)}{\left[\sum_{i=1}^{nS} y_i^2 - \sum_{c=1}^{nI} nI_c \bar{y}_c^2\right] \Big/ (nS - nI)} \tag{15}$$

follows an F-distribution with (nI–1, nS–nI) degrees of freedom, where $\bar{y}_c = \sum_{i \in X_c} y_i / nI_c$ and $\bar{y}$ is defined in conjunction with Eq. (1). Given that the indicated

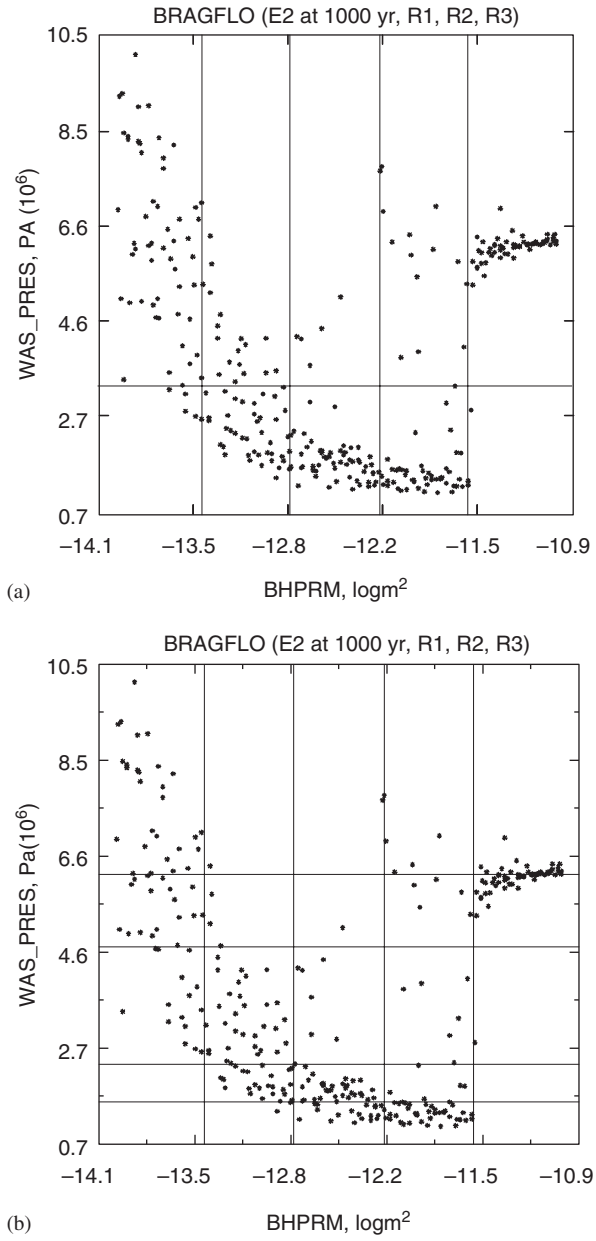

Fig. 11. Grids used to test for nonrandom patterns: (a) partitioning of range of x_j for CMNs and CLs tests and ranges of x_j and y for CMDs test (Ref. [101, Fig. 8.8]), and (b) partitioning of ranges of x_j and y for SI (Ref. [101, Fig. 8.9]).

assumptions hold, the probability $\text{prob}_F(\tilde{F} > F|\ nI\text{–}1, nS\text{–}nI)$ of obtaining an F-statistic of value $\tilde{F}$ that exceeds the value of F in Eq. (15) can be obtained from an F-distribution with (nI–1, nS–nI) degrees of freedom. A low probability (i.e., p-value) of obtaining a larger value for F suggests that the observed pattern involving x_j and y did not arise by chance and hence that x_j has an effect on the behavior of y.

The CLs test employs the Kruskal–Wallis test statistic T, which is based on rank-transformed data and uses the same classes of x_j values as the F-statistic in Eq. (15) Ref. [155, pp. 229–230]). Specifically,

$$T = \left[\sum_{c=1}^{nI}(R_c^2/nI_c) - nS(nS+1)^2/4\right] \Big/ s^2, \tag{16}$$

where

$$R_c = \sum_{i \in X_c} r(y_i), s^2 = \left[\sum_{i=1}^{nS} r(y_i)^2 - nS(nS+1)^2/4\right] \Big/ (nS-1)$$

and $r(y_i)$ denotes the rank of y_i. If the y values conditional on each class of x_j values have the same distribution, then the statistic T in Eq. (16) approximately follows a χ^2 distribution with nI–1 degrees of freedom (Ref. [155, pp. 230–231]). Thus, the probability $\text{prob}_{\chi^2}(\tilde{T} > T|nI-1)$ of obtaining a value $\tilde{T}$ that exceeds T in the presence of identical y distributions for the individual classes can be obtained from a χ^2 distribution with nI–1 degrees of freedom. A small value for $\text{prob}_{\chi^2}(\tilde{T} > T|nX-1)$ (i.e., a p-value) indicates that the values for y's conditional on individual classes have different distributions and thus, most likely, different means and medians. Hence, a small p-value indicates that x_j has an effect on y.

The CMDs test is based on the χ^2-test for contingency tables, which can be used to test for the equality of the median values of y for the classes into which the values of x_j have been divided (Ref. [155, pp. 143–178]). First, the median $y_{0.5}$ for y is estimated for all nS observations. Specifically,

$$y_{0.5} = \begin{cases} y_{(0.5nS)} & \text{if } 0.5nS \text{ is an integer,} \\ [y_{([0.5nS])} + y_{([0.5nS]+1)}]/2 & \text{otherwise,} \end{cases} \tag{17}$$

where $y_{(i)}$, $i = 1, 2, \ldots, nS$, denotes the ordering of the y-values such that $y_{(i)} \leqslant y_{(i+1)}$ and [~] designates the greatest integer function. The individual classes of x_j values are then further subdivided on the basis of whether y values fall above or below $y_{0.5}$ (Fig. 11a). For class c, let nI_{1c} equal the number of y values that exceed $y_{0.5}$, and let nI_{2c} equal the number of y values that are less than or equal to $y_{0.5}$.

The result of this partitioning is a $2 \times nI$ contingency table with nI_{rc} observations in each cell (i.e., in cell (r, c), where r and c designate "row" and "column," respectively, in the corresponding contingency table). The following statistic can now be defined:

$$T = \sum_{c=1}^{nI}\sum_{r=1}^{2}(nI_{rc} - nE_{rc})^2/nE_{rc}, \tag{18}$$

where

$$nE_{rc} = \left(\sum_{p=1}^{2} nI_{pc}/nS\right)\left(\sum_{q=1}^{nI} nI_{rq}/nS\right)nS = \left(\sum_{p=1}^{2} nI_{pc}\right)\left(\sum_{q=1}^{nI} nI_{rq}\right) \Big/ nS$$

and corresponds to the expected number of observations in cell (r, c). If the individual classes of x_j values have equal

medians, then T approximately follows a χ^2 distribution with $(nI-1)(2-1) = nI-1$ degrees of freedom (Ref. [155, p. 156]). Thus, the probability of obtaining a value $\tilde{T}$ that exceeds T in the presence of equal medians is given by $\text{prob}_{\chi^2}(\tilde{T} > T|nI-1)$. A small value (i.e., p-value) for $\text{prob}_{\chi^2}(\tilde{T} > T|nI-1)$ indicates that the y's conditional on individual classes have different medians and hence that x_j has an influence on y.

The SI test also uses the χ^2-test to indicate if the pattern appearing in a scatterplot appears to be nonrandom. The SI test uses the same partitioning of x_j values as used for the CMNs, CLs and CMDs tests. In addition, the y values are also partitioned in a manner analogous to that used for the x_j values (Fig. 11b). For notational convenience, let r, $r = 1, 2, \ldots, nD$, designate the individual classes into which the values of y are divided; let Y_r designate the set such that $i \in \mathrm{Y}_r$ only if y_i belongs to class r; and let nD_r equal the number of elements contained in Y_r (i.e., the number of y_i's associated with class r).

The partitioning of x_j and y into nI and nD classes in turn partitions (x_j, y) into $nI\ nD$ classes (Fig. 11a), where (x_{ij}, y_i) belongs to class (r, c) only if x_{ij} belongs to class c of the x_j values (i.e., $i \in \mathrm{X}_c$) and y_i belongs to class r of the y values (i.e., $i \in Y_r$). For notational convenience, let O_{rc} denote the set such that $x_{ij} \in O_{rc}$ only if $i \in X_c$ (i.e., x_{ij} is in class c of x_j values) and also $i \in Y_r$ (i.e., y_i is in class r of y values), and let nO_{rc} equal the number of elements contained in O_{rc}. Further, if x_j and y are independent, then

$$nE_{rc} = (nD_r/nS)(nI_c/nS)nS = nD_r\, nI_c/nS \tag{19}$$

is an estimate of the expected number of observations (x_j, y) that should fall in class (r, c).

The following statistic can be defined:

$$T = \sum_{c=1}^{nI}\sum_{r=1}^{nD}(nO_{rc} - nE_{rc})^2/nE_{rc}. \tag{20}$$

Asymptotically, T follows a χ^2-distribution with $(nI–1)(nD–1)$ degrees of freedom when x_j and y are independent (Ref. [155, pp. 158–153]). Thus, $\text{prob}_{\chi^2}[\tilde{T} > T|(nI-1)(nD-1)]$ is the probability (i.e., p-value) of obtaining a value of $\tilde{T}$ that exceeds T when x_j and y are independent. A small p-value indicates that the pattern in the scatterplot arose from some underlying relationship involving x_j and y rather than from chance alone. As shown by comparison of Eqs. (18) and (20), the CMDs and SI tests differ only in the partitionings used for the y values.

The four tests described in this section are illustrated in Table 5 for $y = WAS_PRES$ at 10,000 yr under undisturbed conditions (Fig. 5a) and disturbed conditions (Fig. 10a). Scatterplots illustrating the partitioning for $x_j = BHPRM$ and $y = WAS_PRES$ under disturbed conditions are given in Fig. 11. For perspective, rankings based on CCs and RCCs are also presented in Table 4. The relationships between $y = WAS_PRES$ and the dominant sampled variables under undisturbed conditions are fairly linear, with the result that all ranking procedures (i.e., CMNs, CLs, CMDs, SI, CCs, RCCs) give the same ordering of variable importance for the top four variables. In contrast, the relationship between $y = WAS_PRES$ and $x_j = BHPRM$ under disturbed conditions is both nonlinear and nonmonotonic (Fig. 11), with the result that the tests based on gridding (i.e., CMNs, CLs, CMDs, SI) all identify $BHPRM$ as being the dominant variable influencing the uncertainty in WAS_PRES; in contrast, the effect of $BHPRM$ was completely missed by tests based on CCs and RCCs. (Table 5).

The CMNs, CLs, CMDs and SI tests discussed in this section are all based on p-values that derive from statistical tests predicated on assumptions that are certainly not satisfied in their entirety in sampling-based sensitivity analyses. Thus, it is possible that the violation of these assumptions could be leading to misrankings of variable importance. Such a possibility can be explored by using a Monte Carlo procedure to assess if the use of formal statistical procedures to determine p-values is producing misleading results (Ref. [156]; Ref. [157, Section 14.5]). Specifically, nR samples of the form

$$(x_{ij}, y_i), \quad i = 1, 2, \ldots, nS \tag{21}$$

can be generated by pairing the nS values for x_j randomly and without replacement with the nS values for y. This random assignment is repeated nR times to produce nR samples of the form in Eq. (21) for each uncertain input x_j under consideration. In this example, $nR = 10,000$ and $nS = 300$. For a given procedure (i.e., CMNs, CLs, CMDs, SI), each of the nR samples can be used to calculate the value of the statistic used to determine the corresponding p-value. The resulting empirical distribution of the statistic can then be used to estimate the p-value for the statistic actually observed in the analysis. Comparison of the p-value obtained for a given set of statistical assumptions with the p-value obtained from the empirical distribution of the corresponding statistic provides an indication of the robustness of the variable rankings with respect to possible deviations from the assumptions underlying the formal statistical procedure. As examination of Table 6 shows, the variable rankings illustrated in this section are quite robust with respect to possible deviations from the underlying statistical assumptions on which they are predicated.

Additional Information: Ref. [46, Sections 6.6.8 and 6.6.9]; Refs. [47,158–160].

6.7. Entropy tests for patterns based on gridding

Measures of entropy provide another grid-based procedure to assess the strength of nonlinear relationships between the x_j and y. Specifically, the following quantities can be defined (Ref. [157, pp. 480–484]):

$$H(y) = -\sum_{r=1}^{nD}(nD_r/nS)\ln(nD_r/nS), \tag{22}$$

Table 5
Comparison of statistical tests for patterns based on gridding for pressure (*WAS_PRES*) at 10,000 yr under undistributed (i.e., *E*0) conditions (Fig. 5a) and disturbed (i.e., *E*2) conditions (Fig. 10a) (adapted from Ref. [47, Tables 4 and 21])

Variable[a]	CMNs: 1×5[b]		CLs: 1×5[c]		CMDs: 2×5[d]		SI: 5×5[e]		CCs[f]		RCCs[g]	
	Rank	*p*-Val	*Rank*	p-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a)												
WMICDFLG	1	0.0000	1	0.0000	1	0.0000	1	0.0000	1	0.0000	1	0.0000
HALPOR	2	0.0000	2	0.0000	2	0.0000	2	0.0000	2	0.0000	2	0.0000
WGRCOR	3	0.0000	3	0.0000	3	0.0025	3	0.0003	3	0.0000	3	0.0000
ANHPRM	4	0.0195	4	0.0187	4	0.0663	4	0.0049	4	0.0241	4	0.0268
ANHBCVGP	18	0.8062	16	0.7686	14	0.6442	5	0.0194	20	0.8084	15	0.7686
Pressure, disturbed (i.e., *E*2) conditions at 10,00 yr (Fig. 10a)												
BHPRM	1	0.0000	1	0.0000	1	0.0000	1	0.0000	10	0.3651	6	0.1704
HALPRM	2	0.0000	2	0.0000	2	0.0000	2	0.0002	1	0.0000	1	0.0000
ANHPRM	3	0.0002	3	0.0000	3	0.0007	4	0.0049	2	0.0000	2	0.0000
ANHBCEXP	4	0.0405	4	0.0602	4	0.0595	14	0.4414	7	0.1786	8	0.2373
HALPOR	5	0.0415	5	0.0940	5	0.0700	11	0.3142	3	0.0090	3	0.0184
WGRCOR	17	0.5428	9	0.2242	14.5	0.5249	3	0.0002	20	0.7676	17	0.6560

[a]Table includes only variables that had a *p*-value less than 0.05 for at least one of the procedures although the variable rankings for a specific procedure are based on the *p*-values obtained for that procedure for all variables considered in the analysis (see Table 1; variable *BHPRM* not included in analyses for undisturbed conditions).
[b]Variable ranks and *p*-values for CMNs test with 1×5 grid; see Eq. (15). Exceptions for CMNs, CLs, CMDs and SI tests: because variables *ANHBCVGP* and *WMICDFLG* are discrete with 2 and 3 values, respectively (see Table 1), $nI=2$ and 3 rather than 5 for these two variables.
[c]Variable ranks and *p*-values for CLs test with 1×5 grid; see Eq. (16).
[d]Variable ranks and *p*-values for CMDs test with 2×5 grid; see Eq. (18).
[e]Variable ranks and *p*-values for SI test with 5×5 grid; see Eq. (20).
[f]Variable ranks and *p*-values for CC; see Eq. (24), Ref. [47].
[g]Variable ranks and *p*-values for RCC; see Eq. (38), Ref. [47].

$$H(x_j) = -\sum_{c=1}^{nI}(nI_c/nS)\ln(nI_c/nS), \tag{23}$$

$$H(y,x_j) = -\sum_{r=1}^{nD}\sum_{c=1}^{nI}(nO_{rc}/nS)\ln(nO_{rc}/nS), \tag{24}$$

$$\begin{aligned} H(x_j|y) &= \sum_{r=1}^{nD}\left\{\frac{nD_r}{nS}\right\} \times\left\{-\sum_{c=1}^{nI}[(nO_{rc}/nS)/nD_r/nS] \times\ln[(nO_{rc}/nS)/nD_r/nS]\right\} \\ &= -\sum_{r=1}^{nD}\sum_{c=1}^{nI}(nO_{rc}/nS)\ln(nO_{rc}/nD_r) \\ &= H(y,x_j) - H(y), \end{aligned} \tag{25}$$

$$\begin{aligned} H(y|x_j) &= \sum_{c=1}^{nI}\left\{\frac{nI_c}{nS}\right\} \times\left\{-\sum_{r=1}^{nD}[(nO_{rc}/nS)/(nI_c/nS)] \times\ln[(nO_{rc}/nS)/(nI_c/nS)]\right\} \\ &= -\sum_{c=1}^{nI}\sum_{r=1}^{nD}(nO_{rc}/nS)\ln(nO_{rc}/nI_c) \\ &= H(y,x_j) - H(x_j), \end{aligned} \tag{26}$$

$$\begin{aligned} U(x_j|y) &= [H(x_j) - H(x_j|y)]/H(x_j) \\ &= [H(y) + H(x_j) - H(y,x_j)]/H(x_j), \end{aligned} \tag{27}$$

$$\begin{aligned} U(y|x_j) &= [H(y) - H(y|x_j)]/H(y) \\ &= [H(y) + H(x_j) - H(y,x_j)]/H(y), \end{aligned} \tag{28}$$

$$\begin{aligned} U(y,x_j) &= 2[H(y) + H(x_j) - H(y,x_j)]/[H(y) + H(x_j)] \\ &= [H(y)U(y|x_j) + H(x)U(x_j|y)]/[H(y) + H(x_j)], \end{aligned} \tag{29}$$

where (i) $H(y)$ and $H(x_j)$ are estimates of the entropy associated with y and x_j, respectively, (ii) $H(y, x_j)$ is an estimate of the entropy associated with y and x_j, (iii) $H(x_j|y)$ and $H(y|x_j)$ are estimates of the expected entropy of x_j conditional on y and the expected entropy of x_j conditional on y, respectively, (iv) $U(x_j|y)$ and $U(y|x_j)$ are measures (i.e., uncertainty coefficients) of the contributions of y to the entropy associated with x_j and of x_j to the entropy associated with y, respectively, (v) $U(y,x)$ is an entropy-based measure of the strength of the association between x_j and y, (vi) the remaining expressions are the same as defined in Section 6.6, and (vii) the defined

Table 6
Comparison of variable rankings obtained with formal statistical procedures and Monte Carlo procedures for statistical tests for patterns based on gridding for pressure (*WAS_PRES*) at 10,000 yr under undisturbed (i.e., E0) conditions (Adapted from Ref. [47, Table 8]; see Ref. [47, Table 23], for a similar comparison for pressure at 10,000 yr under disturbed (i.e., E2) conditions)

Variable name[a]	CMN: 1×5[b]		CMNMC: 1×5[c]		Variable name[a]	CL: 1×5[b]		CLMC: 1×5[c]	
	Rank	*p*-Val	Rank	*p*-Val		Rank	*p*-Val	Rank	*p*-Val
WMICDFLG	1.0	0.0000	2.0	0.0000	*WMICDFLG*	1.0	0.0000	2.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0000	*HALPOR*	2.0	0.0000	2.0	0.0000
WGRCOR	3.0	0.0000	2.0	0.0000	*WGRCOR*	3.0	0.0000	2.0	0.0000
ANHPRM	4.0	0.0195	4.0	0.0214	*ANHPRM*	4.0	0.0187	4.0	0.0212
SHPRMASP	5.0	0.1439	5.0	0.1495	*SHPRMASP*	5.0	0.1237	5.0	0.1277
WRBRNSAT	6.0	0.1506	6.0	0.1526	*WRBRNSAT*	6.0	0.2042	6.0	0.2053
SHRGSSAT	7.0	0.2488	7.0	0.2497	*ANRBRSAT*	7.0	0.2710	7.0	0.2710
ANRBRSAT	8.0	0.3034	8.0	0.3027	*SHRGSSAT*	8.0	0.3153	8.0	0.3167
…	…	…	…	…	…	…	…	…	…
WGRMICI	23.0	0.9705	23.0	0.9717	*WGRMICI*	23.0	0.9649	23.0	0.9663
WGRMICH	24.0	0.9975	24.0	0.9973.	*WGRMICH*	24.0	0.9865	24.0	0.9839
TDCC[d]			0.970		TDCC[d]			0.971	
Variable name[a]	CMD: 2×5[b]		CMDMC: 2×5[c]		Variable name[a]	SI: 5×5[b]		SIMC: 5×5[c]	
	Rank	*p*-Val	Rank	*p*-Val		Rank	*p*-Val	Rank	*p*-Val
WMICDFLG	1.0	0.0000	1.5	0.0000	*WMICDFLG*	1.0	0.0000	1.5	0.0000
HALPOR	2.0	0.0000	1.5	0.0000	*HALPOR*	2.0	0.0000	1.5	0.0000
WGRCOR	3.0	0.0025	3.0	0.0018	*WGRCOR*	3.0	0.0003	3.0	0.0003
ANHPRM	4.0	0.0663	4.0	0.0690	*ANHPRM*	4.0	0.0049	4.0	0.0038
SHPRMASP	5.0	0.2427	5.0	0.2401	*ANHBCVGP*	5.0	0.0194	5.0	0.0178
SHPRMCON	6.0	0.2674	6.0	0.2718	*WRGSSAT*	6.0	0.1229	6.0	0.1196
ANRBRSAT	7.0	0.3386	7.0	0.3329	*SHPRMCON*	7.0	0.1487	7.0	0.1529
HALPRM	8.0	0.3883	8.0	0.3967	*WASTWICK*	8.0	0.1850	8.0	0.1829
…	…	…	…	…	…	…	…	…	…
WGRMICH	23.0	0.9554	23.0	0.9439	*WGRMICH*	23.0	0.9437	23.0	0.9429
WGRMICI	24.0	0.9702	24.0	0.9664	*ANRGSSAT*	24.0	0.9763	24.0	0.9791
TDCC[d]			0.986		TDCC[d]			0.988	

[a]Twenty-four (24) variables included in analysis; highly correlated variables and variables not relevant to E0 conditions not included.
[b]Variable rankings obtained with a maximum of five classes of x values (i.e., $nI = 5$; see footnote b, Table 5) and analytic determination of p-values.
[c]Variable rankings obtained with a maximum of five classes of x values (i.e., $nI = 5$; see footnote b, Table 5) and Monte Carlo determination of p-values.
[d]Top down coefficient of concordance (TDCC, see Section 6.12) with variable rankings obtained with a maximum of five classes of x values (i.e., $nI = 5$; see footnote b, Table 5) and analytic determination of p-values.

quantities in Eqs. (22)–(29) are conditional on the grid structure in use.

The quantities $U(y|x_j)$ and $U(y, x_j)$ can be used as sensitivity measures, with $U(y|x_j)$ providing a measure of the effect of the uncertainty in x_j on the uncertainty in y and $U(y, x_j)$ providing a measure of the joint behavior of x_j and y. Both quantities equal zero when there is no relationship between y and x_j that is identifiable with the grid structure in use and equal one when there is a perfect association between y and x_j with the grid structure in use. Values between zero and one are indicative of intermediate levels of association. Specifically,

$$U(y|x_j) = U(y, x_j) = 0 \tag{30}$$

if

$$nO_{rc} = nS/(nD_r nI_c) \tag{31}$$

for $r = 1, 2, \ldots, nD$ and $c = 1, 2, \ldots, nI$, and

$$U(y|x_j) = U(y, x_j) = 1 \tag{32}$$

if each interval of values for x_j is associated with only one interval of values for y and each interval of values for y is associated with only one interval of values for x_j. Necessary, but not sufficient, conditions for the equality in Eq. (31) are (i) $nI = nD$, and (ii) $nIc = nDc$, $c = 1, 2, \ldots, nI(= nD)$.

When the nI and nD intervals into which the values for x_j and y are divided contain equal numbers of sampled values (i.e., nS/nI and nS/nD values for the intervals associated with x_j and y, respectively), then the following simpler expressions result:

$$H(x_j) = \ln(nI), H(y) = \ln(nD), \tag{33}$$

$$H(y|x_j) = H(y, x_j) - \ln(nI), H(x|y) = H(y, x_j) - \ln(nD), \tag{34}$$

$$U(y|x_j) = [\ln(nI) + \ln(nD) - H(y, x_j)]/\ln(nD), \tag{35}$$

$$U(x_j|y) = [\ln(nI) + \ln(nD) - H(y, x_j)]/\ln(nI), \tag{36}$$

$$U(y, x_j) = 2[\ln(nI) + \ln(nD) - H(y, x_j)]/[\ln(nI) + \ln(nD)]. \quad (37)$$

Further,

$$U(y|x_j) = U(x_j|y) = U(y, x_j) = 2 - H(y, x_j)/\ln(nI) \quad (38)$$

if $nI = nD$.

As shown by comparison of Eqs. (35) and (37), use of either $U(y|x_j)$ or $U(y, x_j)$ will produce identical rankings of variable importance based on the size of $H(y, x_j)$ when the same values for nI and nD and also for $nIc = nS/nI$ and $nDr = nS/nD$ are used in the determination of $U(y|x_j)$ and $U(y, x_j)$ for each of the independent variables under consideration. Specifically, $U(y|x_j)$ and $U(y, x_j)$ increase in size as the entropy $H(y, x_j)$ associated with joint distribution for x_j and y decreases. Thus, $U(y|x_j)$ and $U(y, x_j)$ are really sensitivity measures that quantify variable importance on the basis of the entropy $H(y, x_j)$ associated with x_j and y. Specifically, the smaller the entropy $H(y, x_j)$, the more important x_j is assessed to be in affecting the value of y. As shown in Eq. (38), $U(y|x_j)$ and $U(y, x_j)$ have identical numerical values when $nI = nD$ and $nI_c = nD_r = nS/nD$.

A closely related measure of association is given by

$$R(y, x_j) = \{1 - \exp(-2[H(x_j) + H(y) - H(y, x_j)])\}^{1/2}, \quad (39)$$

which has (i) a value of zero if there is no association between x_j and y in the sense indicated in Eq. (30), (ii) a value that approaches one as nI and nD increase if there is perfect association between x_j and y in the sense indicated in conjunction with Eq. (32), and (iii) intermediate values for intermediate levels of association (Ref. [161]). If x_j and y have a bivariate normal distribution, then $R(y, x_j)$ approaches the absolute value of the correlation coefficient between x_j and y as the sample and grid sizes increase [161].

As suggested by Mishra and Knowlton [162], the SI test (i.e., a χ^2-test on the same grid used to define entropy measures) can be used to identify important variables, and then the entropy measures $U(y, x_j)$, $U(y|x_j)$ and $R(y, x_j)$ can be used to provide a numerical representation of variable importance. The result of this approach is illustrated in Table 7, with the top two sets of results corresponding to the use of $nI = nD = 5$, and the lower two sets corresponding to the use of $nI = 10$ and $nD = 5$. As should be the case, the values for $U(y, x_j)$ and $U(y|x_j)$ are the same when $nI = nD$ and are somewhat different when $nI \neq nD$. Further, there is little difference in the variable rankings based on the SI test and on the entropy measures $U(y, x_j)$, $U(y|x_j)$ and $R(y, x_j)$. Although $U(y, x_j)$, $U(y|x_j)$ and $R(y, x_j)$ result in the same rankings of variable importance because of the underlying dependence on $H(y, x_j)$, the normalization associated with the definition of $R(y, x_j)$ produces results that are more widely spread over the interval [0,1]. Although not presented, similar normalizations referred to as Cramer's V and the contingency coefficient, respectively, are also possible for the χ^2statistic T in Eq. (20) associated with SI test (see Ref. [157, Section 13.6]). The right-most columns in Table 7 labeled "KS Test" and "KSMC Test" relate to a sensitivity analysis procedure based on a two-dimensional Kolmogorov–Smirnov (KS) test that will be discussed in Section 6.10.

The similarity between the ranking of variable importance with the SI test and with entropy-based measures is quite striking (Table 8). For all practical purposes, the χ^2 statistic T defined in Eq. (20) associated with the SI test and the entropy-based measures $U(y, x_j)$, $U(y|x_j)$ and $R(y, x_j)$ defined in Eqs. (28), (29) and (39) give the same rankings of variable importance. However, when discrete variables such as *ANHBCVGP* and *WMICDFLG* are under consideration, there can be some differences between rankings based on p-values for the χ^2 statistic and rankings based on either the χ^2 statistic itself or entropy measures because of the effects of the resultant different degrees of freedom associated with different variables on the p-values for the χ^2 statistic. Clearly, there is a close algebraic connection between T and the entropy-based measures $U(y, x_j)$, $U(y|x_j)$ and $R(y, x_j)$. As previously illustrated, p-values for the χ^2 statistic provide a way to discern influential from noninfluential variables for both the SI test and the entropy-based measures. Although not illustrated, the Monte Carlo procedure discussed in conjunction with Eq. (21) and Table 6 for the empirical determination of p-values could be used to directly determine p-values for $U(y, x_j)$, $U(y|x_j)$ and $R(y, x_j)$.

Additional information: Ref. [157, pp. 480–484]; Refs. [161–164].

6.8. Nonparametric regression

There are drawbacks to the parametric regression techniques indicated in Section 6.3 that can reduce their effectiveness in some sensitivity analyses. First, it is necessary to provide an a priori specification of the form of the regression model (e.g., linear as in Eqs. (3) and (12), nonlinear as in Eq. (13), or linear with rank transformed data as discussed in Section 6.5). Unfortunately, when complex patterns of behavior are present, it can be difficult to determine the appropriate form for a regression model. Such determinations can be a particular challenge in exploratory analyses that can involve 10 s or even 100 s of analysis results, with each result potentially requiring the specification of a different regression model. Second, the specified form for the regression is required to hold across the entire mapping from analysis inputs to analysis results, which makes the representation of local behavior and/or asymptotes difficult. In addition, the grid-based procedures discussed in Sections 6.6 and 6.7 have the drawback that the associated sensitivity results can be dependent on the particular grid selected for use. Unfortunately, the most appropriate grid for use with these procedures is not always apparent.

Nonparametric regression procedures provide an alternative to parametric regression procedures and grid-based procedures that can mitigate the potential problems

Table 7
Examples of entropy measures to identify uncertain variables affecting the uncertainty in pressure (*WAS_PRES*) at 10,000 yr under undisturbed (i.e., E0) Conditions (Fig. 5a) and Disturbed (i.e., E2) conditions (Fig. 10a)

Variable[a]	SI test[b]			Entropy[c]		Cond. entropy[d]		*R*-statistic[e]	
	χ^2	*p*-Value	Rank	$U(y, x_j)$	Rank	$U(y\|x_j)$	Rank	$R(y, x_j)$	Rank
Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a): $nI = 5$, $nD = 5$									
WMICDFLG	198.6	0.0000	1	0.2868	1	0.2361	1	0.7296	1
HALPOR	127.2	0.0000	2	0.1350	2	0.1350	2	0.5930	2
WGRCOR	42.5	0.0003	3	0.0485	3	0.0485	3	0.3800	3
ANHPRM	34.3	0.0049	4	0.0420	4	0.0420	4	0.3560	4
ANHBCVGP	11.7	0.0194	5	0.0172	15.5	0.0123	25	0.1970	25
Pressure, disturbed (i.e., *E*2) conditions at 10,000 yr (Fig. 10a): $nI = 5$, $nD = 5$									
BHPRM	337.2	0.0000	1	0.3700	1	0.3700	1	0.8340	1
HALPRM	43.7	0.0002	2	0.0526	2	0.0526	2	0.3940	2
WGRCOR	43.7	0.0002	3	0.0456	3	0.0456	3	0.3690	3
ANHPRM	34.3	0.0049	4	0.0405	4	0.0405	4	0.3500	4
Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a): $nI = 10$, $nD = 5$									
WMICDFLG	198.6	0.0000	1	0.1868	1	0.2361	1	0.7296	1
HALPOR	140.2	0.0000	2	0.1240	2	0.1510	2	0.6200	2
WGRCOR	56.3	0.0167	3	0.0515	4.	0.0626	4	0.4270	4
ANHPRM	53.3	0.0314	4	0.0547	3	0.0664	3	0.4390	3
Pressure, disturbed (i.e., *E*2) conditions at 10,000 yr (Fig. 10a); $nI = 10$, $nD = 5$									
BHPRM	402.3	0.0000	1	0.3490	1	0.4240	1	0.8630	1
WGRCOR	69.0	0.0008	2	0.0616	2	0.0749	2	0.4630	2
HALPRM	63.0	0.0035	3	0.0601	3	0.0731	3	0.4580	3
ANHPRM	63.0	0.0035	4	0.0594	4	0.0722	4	0.4550	4

[a]Table includes only variables that had a *p*-value less than 0.05 for SI test.
[b]χ^2 value, *p*-value and variable rank for SI test with 5×5 grid for $nI = 5$, $nD = 5$ and 10×5 grid for $nI = 10$, $nD = 5$; see Eq. (20). Exception: because variables *ANHBCVGP* and *WMICDFLG* are discrete with 2 and 3 values, respectively (see Table 1), $nI = 2$ and 3 rather than 5 for these two variables.
[c]Entropy $U(y, x_j)$ and variable rank; see Eq. (29).
[d]Conditional entropy and variable rank; see Eq. (28).
[e]*R*-statistic ($R(y, x_j)$ and variable rank; see Eq. (39).

indicated in the preceding paragraph. With nonparametric regression procedures, an a priori specification of the exact algebraic form of the regression model is not required. Rather, an iterative procedure is used to construct a model that captures the relationships that are present in the mapping between analysis inputs and a particular analysis result. This iterative construction procedure does not require the use of a grid and produces a model that can represent local patterns of behavior. Nonparametric regression is often referred to as smoothing. Popular nonparametric regression procedures include (i) locally weighted regression (LOESS), (ii) generalized additive models (GAMs), (iii) projection pursuit regression (PP_REG), and (iv) recursive partitioning regression (RP_REG). These procedures are briefly described below.

The LOESS technique is based on the assumption that the relationship between y and $\mathbf{x}$ is of the form

$$y = f(\mathbf{x}) = \alpha(\mathbf{x}) + \beta(\mathbf{x})\mathbf{x}, \tag{40}$$

where $\boldsymbol{\beta}(\mathbf{x}) = [\beta 1(\mathbf{x}), \beta 2(\mathbf{x}), \ldots, \beta nX(\mathbf{x})]$ and $\mathbf{x} = [x1, x2, \ldots, xnX]^{\mathrm{T}}$. In turn, an approximate relationship of the form

$$\hat{y} = \hat{f}(\mathbf{x}) = \hat{\alpha}(\mathbf{x}) + \hat{\beta}(\mathbf{x})\mathbf{x} \tag{41}$$

is sought with LOESS. The quantities $\hat{\alpha}(\mathbf{x})$ and $\hat{\boldsymbol{\beta}}(\mathbf{x})$ for a given value of $\mathbf{x}$ are defined to be the values for α and $\boldsymbol{\beta} = [\beta 1, \beta 2, \ldots, \beta nX]$ that minimize the sum

$$\sum_{i=1}^{nS} (\alpha + \boldsymbol{\beta}\mathbf{x}_i - y_i)^2 \left[1 - \left(\frac{\|\mathbf{x} - \mathbf{x}_i\|}{d_r(\mathbf{x})}\right)^3\right]^3 I_{[0,d_r(\mathbf{x}))}(\|\mathbf{x} - \mathbf{x}_i\|), \tag{42}$$

where (i) $d_r(\mathbf{x})$ is the distance to the rth nearest neighbor (NN) of $\mathbf{x}$ in nX-dimensional Eulidean space, (ii) $I_{[0,dr(\mathbf{x}))}(\|\mathbf{x}–\mathbf{x}_i\|$ equals 1 if $\|\mathbf{x}–\mathbf{x}_i\| < d_r(\mathbf{x})$ and equals 0 otherwise, and (iii) the individual independent variables (i.e., x_1, x_2, …, x_{nX}) are normalized to mean zero and standard deviation one so that the value of the norm $\| \cdot \|$ is not dominated by the units used for these variables. The determination of α and $\boldsymbol{\beta}$ is straightforward with the use of appropriate matrix techniques (Ref. [165, p. 139]).

For GAMs, the function $f(\mathbf{x})$ is assumed to have the form

$$f(\mathbf{x}) = \sum_{j=1}^{nX} f_j(x_j), \tag{43}$$

Table 8
Detailed comparison of χ^2 statistic T and entropy $U(y, x_j)$ used to identify uncertain variables affecting the uncertainty in pressure (*WAS_PRES*) at 10,000 yr under undisturbed (i.e., *E*0) conditions (Fig. 5a) and disturbed (i.e., *E*2) conditions (Fig. 10a)

Variable[a]	Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a): $nI = 5$, $nD = 5$				Variable[a]	Pressure, disturbed (i.e., *E*2) conditions at 10,000 yr (Fig. 10a): $nI = 5$, $nD = 5$			
	SI test			Entropy		SI test			Entropy
	χ^2 [b]	df[c]	*p*-Value	$U(y, x_j)$[d]		χ^2 [b]	df[c]	*p*−Value	$U(y, x_j)$[d]
WMICDFLG	198.6 (1.0)	8	0.0000 (1.0)	0.2868 (1.0)	*BHPRM*	337.2 (1.0)	16	0.0000 (1.0)	0.3700 (1.0)
HALPOR	127.0 (2.0)	16	0.0000 (1.0)	0.1350 (2.0)	*WGRCOR*	43.7 (2.0)	16	0.0002 (2.0)	0.0456 (3.0)
WGRCOR	42.5 (3.0)	16	0.0003 (3.0)	0.0485 (3.0)	*HALPRM*	43.7 (3.0)	16	0.0002 (3.0)	0.0526 (2.0)
ANHPRM	34.3 (4.0)	16	0.0049 (4.0)	0.0420 (4.0)	*ANHPRM*	34.3 (4.0)	16	0.0049 (4.0)	0.0405 (4.0)
WRGSSAT	22.7 (5.0)	16	0.1229 (6.0)	0.0230 (5.0)	*SHRGSSAT*	25.0 (5.0)	16	0.0698 (5.0)	0.0268 (5.0)
SHPRMCON	21.8 (6.0)	16	0.1487 (7.0)	0.0228 (6.0)	*SHBCEXP*	23.5 (6.0)	16	0.1010 (6.0)	0.0260 (6.0)
WASTWICK	20.8 (7.0)	16	0.1850 (8.0)	0.0223 (7.0)	*WGRMICI*	20.5 (7.0)	16	0.1985 (7.0)	0.0213 (7.0)
SHBCEXP	19.5 (8.0)	16	0.2436 (9.0)	0.0212 (8.0)	*WRBRNSAT*	19.5 (8.0)	16	0.2436 (9.0)	0.0198 (8.0)
SHPRNHAL	19.3 (9.0)	16	0.2518 (10.0)	0.0200 (10.0)	*ANRBRSAT*	19.3 (9.0)	16	0.2518 (10.0)	0.0197 (9.0)
SHPRMSAP	19.2 (10.0)	16	0.2601 (11.0)	0.0190 (12.0)	*SHRBRSAT*	18.2 (10.5)	16	0.3142 (11.5)	0.0186 (11.0)
SHPRMDRZ	18.2 (11.0)	16	0.3142 (12.0)	0.0204 (9.0)	*HALPOR*	18.2 (10.5)	16	0.3142 (11.5)	0.0190 (10.0)
WGRMICI	18.0 (12.0)	16	0.3239 (13.0)	0.0191 (11.0)	*WFBETCEL*	16.8 (12.0)	16	0.3965 (13.0)	0.0175 (12.0)
ANHBCEXP	17.7 (13.0)	16	0.3438 (14.0)	0.0179 (13.5)	*ANHBCEXP*	16.2 (13.0)	16	0.4414 (14.0)	0.0170 (13.0)
WFBETCEL	17.0 (14.0)	16	0.3856 (15.0)	0.0179 (13.5)	*WASTWICK*	15.2 (14.0)	16	0.5125 (15.0)	0.0164 (14.0)
SHRBRSAT	16.3 (15.0)	16	0.4299 (16.0)	0.0169 (17.0)	*WGRMICH*	14.7 (15.0)	16	0.5492 (16.0)	0.0148 (15.5)
ANRBRSAT	15.7 (16.0)	16	0.4765 (17.0)	0.0172 (15.5)	*SHPRMDRZ*	13.8 (16.0)	16	0.6111 (17.0)	0.0148 (15.5)
HALPRM	13.7 (17.0)	16	0.6235 (18.0)	0.0156 (18.0)	*SHPRMCLY*	13.3 (18.0)	16	0.6482 (19.0)	0.0133 (20.0)
SHRGSSAT	13.3 (18.0)	16	0.6482 (19.0)	0.0141 (19.0)	*ANRGSSAT*	13.3 (18.0)	16	0.6482 (19.0)	0.0137 (19.0)
WRBRNSAT	12.8 (19.0)	16	0.6849 (20.0)	0.0131 (20.0)	*SHPRMSAP*	13.3 (18.0)	16	0.6482 (19.0)	0.0145 (17.5)
SALPRES	11.8 (20.0)	16	0.7554 (21.0)	0.0125 (21.0)	*SALPRES*	12.5 (20.0)	16	0.7089 (21.0)	0.0145 (17.5)
ANHBCVGP	11.7 (21.0)	4	0.0197 (5.0)	0.0172 (15.5)	*WRGSSAT*	10.2 (21.0)	16	0.8578 (22.0)	0.0102 (21.0)
SHPRMCLY	8.7 (22.0)	16	0.9265 (22.0)	0.0093 (22.0)	*SHPRNHAL*	9.2 (22.0)	16	0.9064 (24.0)	0.0099 (22.0)
WGRMICH	8.2 (23.0)	16	0.9437 (23.0)	0.0085 (23.0)	*SHPRMCON*	5.8 (23.0)	16	0.9898 (25.0)	0.0059 (24.0)
ANRGSSAT	6.8 (24.0)	16	0.9763 (24.0)	0.0072 (24.0)	*ANHBCVGP*	5.5 (24.0)	4	0.2427 (8.0)	0.0080 (23.0)
					WMICDFLG	3.7 (25.0)	8	0.8859 (23.0)	0.0045 (25.0)

p-Value for χ^2 statistic and variable rank based on *p*-value for χ^2 statistic.
[a]Variables ordered by χ^2 statistic for SI test.
[b]χ^2 statistic for SI test with 5×5 grid (see footnote b, Tables 5 and 7, and Eq. (20)) and variable rank based on values of χ^2 statistic.
[c]Degrees of freedom for χ^2 statistic.
[d]Entropy $U(y, x_j)$ based on 5×5 grid (see footnote b, Tables 5 and 7, and Eq. (29)) and variable rank based on $U(y, x_j)$.

where the f_j are arbitrary functions that will be determined as part of the analysis process. In turn, the observed values for y are assumed to be of the form

$$y_i = f(\mathbf{x}_i) = \sum_{j=1}^{nX} f_j(x_{ij}). \tag{44}$$

Given initial estimates, $\hat{f}_2, \hat{f}_3, \ldots, \hat{f}_{nX}$ for $f_2, f_3, \ldots, f_{nX}$, an estimate $\hat{f}_1$ for f_1 can be obtained through use of the relationship

$$y_i - \sum_{j=2}^{nX} \hat{f}_j(x_{ij}) \cong f_1(x_{i1}) \tag{45}$$

for $i = 1, 2, \ldots, nS$. In particular, a scatterplot smoother (e.g., LOESS with only one independent variable) can be used to smooth the partial residuals on the left-hand side of Eq. (45) across x_1. This produces an estimate $\hat{f}_1$ for f_1 defined across the range of values for x_1. Given this estimate for f_1, the estimate $\hat{f}_2$ for f_2 can be refined in the same manner across the range of values for x_2 with $\hat{f}_1, \hat{f}_3, \hat{f}_4, \ldots, \hat{f}_{nX}$. This procedure then continues and repetitively cycles through the variables. The cycling continues until convergence is achieved. The result is $\hat{f}_j$ defined at $x_{1j}, x_{2j}, \ldots, x_{nS,j}$ for $j = 1, 2, \ldots, nX$. Additional detail is available elsewhere (Ref. [166, pp. 90–91]; Ref. [167, pp. 300–302]).

The PP_REG procedure involves both dimension reduction and additive modeling and is based on the assumption that $f(\mathbf{x})$ has the form

$$f(\mathbf{x}) = \sum_{s=1}^{nD} g_s(\alpha_s x), \tag{46}$$

where $\boldsymbol{\alpha}_s = [\boldsymbol{\alpha}_{1s}, \boldsymbol{\alpha}_{2s}, \ldots, \boldsymbol{\alpha}_{nX,s}]$, $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]^{\mathrm{T}}$, $\boldsymbol{\alpha}_s\mathbf{x}$ corresponds to a linear combination of the elements of $\mathbf{x}$, and g_s is an arbitrary function. Values for g_s, $\boldsymbol{\alpha}_s$ and nD are determined as part of the analysis procedure. The expression in Eq. (46) is an additive model with the quantities $\boldsymbol{\alpha}_s\mathbf{x}$ replacing the elements x_j of $\mathbf{x}$ as the independent variables. Further, this expression involves a reduction in dimension as nD is usually smaller than nX. The entities $\hat{\boldsymbol{\alpha}}_1, \hat{\boldsymbol{\alpha}}_2, \ldots, \hat{\boldsymbol{\alpha}}_{nD}$ and $\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_{nD}$ are estimated as part of

the construction process. This is accomplished by first estimating $\boldsymbol{\alpha}_1$ and g_1. Specifically, $\hat{\boldsymbol{\alpha}}_1$ and $\hat{g}_1$ are defined to be the values for $\boldsymbol{\alpha}$ and $g_{\boldsymbol{\alpha}}$ that minimize the sum

$$\sum_{i=1}^{nS}[y_i - g_{\alpha}(\alpha x_i)]^2, \tag{47}$$

where $\boldsymbol{\alpha} \in R^{nX}$, $\|\boldsymbol{\alpha}\| = 1$, and $g_{\boldsymbol{\alpha}}$ is the outcome of using a scatterplot smoother (e.g., LOESS) on the points $[y_i, \boldsymbol{\alpha}\mathbf{x}_i]$, $i = 1, 2, \ldots, nS$. Once $\hat{\boldsymbol{\alpha}}_1$ and $\hat{g}_1$ are estimated, the partial residuals $y_i - g_1(\hat{\boldsymbol{\alpha}}_1\mathbf{x}_i)$, $i = 1, 2, \ldots, nS$, are used to obtain $\hat{\boldsymbol{\alpha}}_2$ and $\hat{g}_2$. Specifically, $\hat{\boldsymbol{\alpha}}_2$ and $\hat{g}_2$ are defined to be the values for $\boldsymbol{\alpha}$ and $g_{\boldsymbol{\alpha}}$ that minimize the sum

$$\sum_{i=1}^{nS}\{[y_i - \hat{g}_1(\hat{\boldsymbol{\alpha}}_1\mathbf{x}_i) - g_{\boldsymbol{\alpha}}(\alpha x_i)]\}^2, \tag{48}$$

where $\boldsymbol{\alpha} \in R^{nX}$, $||\boldsymbol{\alpha}|| = 1$, and $g_{\boldsymbol{\alpha}}$ is the outcome of using a scatterplot smoother on the points $[y_i - \hat{g}_1(\hat{\boldsymbol{\alpha}}_1\ \mathbf{x}_i),\ \boldsymbol{\alpha}\mathbf{x}_i]$, $i = 1, 2, \ldots, nS$. This process continues until no appreciable improvement based on a relative error criterion is observed.

The RP_REG procedure is based on splitting the data into subgroups where observations within each subgroup are more homogeneous than they are over the set of all observations. Then, $f(\mathbf{x})$ is estimated with regression models defined for each subgroup. Specifically, $f(\mathbf{x})$ is estimated by

$$\hat{f}(\mathbf{x}) = \sum_{s=1}^{nP}(\hat{\alpha}_s + \hat{\boldsymbol{\beta}}_s\mathbf{x})I_s(\mathbf{x}), \tag{49}$$

where (i) A_s, $s = 1, 2, \ldots, nP$, designate the subgroups into which the data are partitioned, (ii) $= {}_s + \boldsymbol{\beta}_s\mathbf{x}$ is the least-squares approximation to y associated with A_s, and (iii) I_s is the indicator functions such $I_s(\mathbf{x}) = 1$ if $\mathbf{x}$ is associated with A_s and $I_s(\mathbf{x}) = 0$ otherwise. The subgroups A_s, $s = 1, 2, \ldots, nP$, are developed algorithmically from the observations $[\mathbf{x}_i, y_i]$, $i = 1, 2, \ldots, nS$.

The preceding procedures can all be carried out in a stepwise manner to determine variable importance, with (i) the most important variable $\tilde{x}_1$ being the variable that results in the single-variable model with the most predictive capability, (ii) the second-most important variable $\tilde{x}_2$ being the variable that in conjunction with $\tilde{x}_1$ results in the two-variable model with the most predictive capability, and so on until (iii) some stopping criteria is reached that indicates that the consideration of additional variables does not produce models with improved predictive capability. Order of selection in the stepwise construction process and fraction of variability explained (i.e., R^2 as defined in Eq. (8)) can be used to indicate variable importance. The F-statistic with appropriate degrees of freedom (a topic too complicated for consideration here; see Ref. [168] and Ref. [169, Section 3.13]) can be used to determine a stopping point in the stepwise variable selection procedure.

Nonparametric regression procedures are illustrated in Table 9 for the pressures in Figs. 5a and 10a at 10,000 yr. For comparison, Table 9 also contains results obtained with parametric regression procedures, with LIN_REG indicating linear regression (see Eq. (3)), RANK_REG indicating rank regression (see Section 6.5), and RS_REG indicating response surface regression (i.e., the regression model in Eq. (12) with $f_j(x_j) = x_j$ and $f_{jl}(x_j, x_l) = x_jx_l$). For the result in Fig. 5a (i.e., pressure at 10,000 yr under undisturbed conditions), the relationship between pressure and the dominant independent variables is fairly monotonic, with the result that all the regression procedures perform reasonably well (i.e., R^2 values between 0.80 and 0.97 for the first five variables selected in the individual regressions). As shown in Fig. 6b, there is a strong nonlinear relationship between the result in Fig. 10a (i.e., pressure at 10,000 yr under disturbed conditions) and the variable *BHPRM*. The stepwise regressions with the four nonparametric procedures all identify *BHPRM* as the most important variable. In contrast, the linear regressions with raw and rank-transformed data fail to identify an effect for *BHPRM*. For this particular variable, the parametric response surface regression (i.e., RS_REG in Table 9) also performs well and results in a regression model with an R^2 value of 0.87; however, in many situations the nonparametric regression procedures will outperform response surface regression.

Additional information: A more detailed discussion of the use of nonparametric regression in sensitivity is given in Ref. [168]. General discussions of nonparametric regression procedures appear in Refs. [165–167,169]. The use of regression trees [170] in sensitivity analysis is discussed and illustrated in Ref. [171].

6.9. Squared rank differences/rank correlation coefficient (SRD/RCC) test

The SRD/RCC test is the result of combining a test for nonrandomness in the relationship between an independent and a dependent variable called the SRD test with the Spearman RCC [172]. This test is effective at identifying linear and very general nonlinear patterns in analysis results. However, unlike the regression procedures introduced in Sections 6.3 and 6.8, the SRD/RCC test does not involve the development of a model that approximates the relationship between independent and dependent variables. Further, unlike the grid-based procedures introduced in Sections 6.6 and 6.7, the SRD/RCC test does not require the introduction and use of a grid.

A brief description of the SRD/RCC test follows. The test is used to assess the relationships between individual elements x_j of $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]$ and a predicted variable y of interest for a random or LHS and a functional relationship of the form $y = f(\mathbf{x})$. The SRD component of the test is based on the statistic

$$Q_j = \sum_{i=1}^{nS-1}(r_{i+1,j} - r_{ij})^2, \tag{50}$$

Table 9
Comparison of variable rankings obtained with parametric regression (i.e., LIN_REG, RANK_REG, RS_REG), nonparametric regression (i.e., LOESS, PP_REG, RP_REG, GAMs), and the squared rank differences/rank correlation (SRD/RCC) test for pressure at (*WAS_PRES*) 10,000 yr under undisturbed (i.e., *E*0) conditions (Fig. 5a) and disturbed (i.e., *E*2) conditions (Fig. 10a)

Variable[a]	R^{2b}	df[c]	*p*-Val[d]	Variable	R^2	df	*p*-Val	Variable	R^2	df	*p*-Val
Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a)											
	LIN_REG				RANK_REG				RS_REG		
WMICDFLG	0.5076	1.0	0.0000	*WMICDFLG*	0.5226	1.0	0.0000	*WMICDFLG*	0.5098	2.0	0.0000
HALPOR	0.7316	1.0	0.0000	*HALPOR*	0.7320	1.0	0.0000	*HALPOR*	0.7462	3.0	0.0000
WGRCOR	0.7923	1.0	0.0000	*WGRCOR*	0.7859	1.0	0.0000	*WGRCOR*	0.8812	4.0	0.0000
ANHPRM	0.8088	1.0	0.0000	*ANHPRM*	0.7975	1.0	0.0001	*ANHPRM*	0.9160	5.0	0.0000
SHRGSSAT	0.8137	1.0	0.0056	*SALPRES*	0.8027	1.0	0.0058	*WASTWICK*	0.9304	6.0	0.0000
SALPRES	0.8177	1.0	0.0119	*SHRGSSAT*	0.8064	1.0	0.0187	*SALPRES*	0.9383	7.0	0.0000
	LOESS				PP_REG			*ANHBCEXP*	0.9427	8.0	0.0119
WMICDFLG	0.5098	2.0	0.0000	*WMICDFLG*	0.5098	2.0	0.0000		RP_REG		
HALPOR	0.7662	6.1	0.0000	*HALPOR*	0.7617	5.4	0.0000	*WMICDFLG*	0.5076	1.0	0.0000
WGRCOR	0.9186	33.1	0.0000	*WGRCOR*	0.9236	21.5	0.0000	*HALPOR*	0.8205	17.0	0.0000
ANHPRM	0.9477	25.1	0.0000	*ANHPRM*	0.9623	11.3	0.0000	*WGRCOR*	0.9220	3.0	0.0000
	GAM			*WASTWICK*	0.9711	10.1	0.0000	*ANHPRM*	0.9662	16.0	0.0000
WMICDFLG	0.5098	2.0	0.0000	*ANYBCVGP*	0.9755	9.1	0.0000	*WASTWICK*	0.9823	40.0	0.0000
HALPOR	0.7448	4.0	0.0000	*WRBRNSAT*	0.9813	10.5	0.0000	SRD/RCC TEST			
WGRCOR	0.8556	4.0	0.0000	*WFBETCEL*	0.9851	11.6	0.0000	*WMICDFLG*	NA[e]	4.0	0.0000
ANHPRM	0.8854	4.0	0.0000	*HALPRM*	0.9874	9.3	0.0000	*HALPOR*	NA	4.0	0.0000
WASTWICK	0.8921	4.0	0.0019	*SALPRES*	0.9901	8.2	0.0000	*WGRCOR*	NA	4.0	0.0001
SHRGSSAT	0.9007	10.0	0.0116	*SHPRMCLY*	0.9929	13.3	0.0000				
SALPRES	0.9042	1.0	0.0018	*SHRBRSAT*	0.9944	9.4	0.0000				
				SHPRMDRZ	0.9969	10.1	0.0000				
Pressure, disturbed (i.e., *E*2) conditions at 10,000 yr (Fig. 10a)											
	LIN_REG				RANK_REG				RS_REG		
HALPRM	0.1410	1.0	0.0000	*HALPRM*	0.1289	1.0	0.0000	*BHPRM*	0.6098	2.0	0.0000
ANHPRM	0.1999	1.0	0.0000	*ANHPRM*	0.1866	1.0	0.0000	*HALPRM*	0.7006	3.0	0.0000
HALPOR	0.2203	1.0	0.0057	*HALPOR*	0.2049	1.0	0.0094	*ANHPRM*	0.7902	4.0	0.0000
LOESS				PP_REG				*HALPOR*	0.8291	5.0	0.0000
BHPRM	0.6625	8.8	0.0000	*BHPRM*	0.6646	9.0	0.0000	*ANHBCVGP*	0.8400	6.0	0.0023
ANHPRM	0.7321	12.8	0.0000	*ANHPRM*	0.7603	10.7	0.0000	*WGRCOR*	0.8532	7.0	0.0013
HALPRM	0.7894	10.5	0.0000	*HALPRM*	0.8440	9.8	0.0000	*SHRBRSAT*	0.8654	8.0	0.0030
ANHBCVGP	0.8286	28.9	0.0058	*HALPOR*	0.8965	10.4	0.0000		RP_REG		
	GAM							*BHPRM*	0.7163	17.0	0.0000
BHPRM	0.6654	10.0	0.0000					*HALPRM*	0.8474	15.0	0.0000
ANHPRM	0.7555	4.0	0.0000					*ANHPRM*	0.8894	-9.0	0.0000
HALPRM	0.8242	2.0	0.0000					*ANRGSSAT*	0.9726	81.0	0.0000
HALPOR	0.8590	2.0	0.0000					SRD/RCC TEST			
								BHPRM	NA	4.0	0.0000
								HALPRM	NA	4.0	0.0000
								ANHPRM	NA	4.0	0.0001
								SHPRMDRZ	NA	4.0	0.0150

[a]Variables listed in order of selection.
[b]Cumulative R^2 value with entry of each variable into model.
[c]Incremental degrees of freedom with entry of each variable into model for all cases except SRD/RCC test; df fixed at 4.0 for all variables for SRD/RCC test.
[d]*p*-Value for model with addition of each new variable. Stepwise procedure terminates at a *p*-value of 0.02.
[e]NA indicates that result is not applicable.

where r_{ij}, $i = 1, 2, \ldots, nS$, is the rank of y obtained with the sample element in which x_j has rank i. Under the null hypothesis of no relationship between x_j and y, the quantity

$$S_j = \{Q_j - [nS(nS^2 - 1)/6]\} / \left\{ \sqrt{nS^5}/6 \right\} \tag{51}$$

approximately follows a standard normal distribution for $nS > 40$. Thus, a p-value p_{rj} indicative of the strength of the nonlinear relationship between x_j and y can be obtained from Q_j. Specifically, p_{rj} is the probability that a value $\tilde{Q}_j > Q_j$ would occur due to chance if there was no relationship between x_j and y.

The RCC component of the test is based on the rank (i.e., Spearman) correlation coefficient

$$rc(x_j, y) = \frac{\sum_{i=1}^{nS}[r(x_{ij}) - (nS+1)/2][r(y_i) - (nS+1)/2]}{\left\{\sum_{i=1}^{nS}[r(x_{ij}) - (nS+1)/2]^2\right\}^{1/2}\left\{\sum_{i=1}^{nS}[r(y_i) - (nS+1)/2]^2\right\}^{1/2}}, \tag{52}$$

where $r(x_{ij})$ and $r(y_i)$ are the ranks associated x_j and y for sample element i. Under the null hypothesis of no rank correlation between x_j and y, the quantity $rc(x_j, y)$ has a known distribution (Ref. [155, Table 10]). Thus, a p-value p_{cj} indicative of the strength of the monotonic relationship between x_j and y can be obtained from $rc(x_j, y)$.

The SRD/RCC test is obtained from combining the p-values p_{rj} and p_{cj} to obtain the statistic

$$\chi_4^2 = -2[\ln(p_{rf}) + \ln(p_{cj})], \tag{53}$$

which has a χ^2-square distribution with four degrees of freedom. The p-value associated with χ_4^2 constitutes the SRD/RCC test for the strength of the relationship between x_j and y.

Results obtained with SRD/RCC test are illustrated in Table 9. Like the nonparametric regression procedures, the SRD/RCC test is able to identify the nonlinear effect associated with *BHPRM* for the result in Fig. 10a (i.e., pressure at 10,000 yr under disturbed conditions), which is completely missed with the linear regression procedures with raw and rank-transformed data.

Additional information: A detailed description of the SRD/RCC test and the determination of the associated p-value is available in the original article [172].

6.10. Two-dimensional Kolmogorov–Smirnov (KS) test

The two dimensional KS test provides a way to test for a pattern in a scatterplot without the use of a grid [173–175]. With this test, each point $[x_{ij}, y_i]$ in the sample $[x_{ij}, y_i]$, $i = 1, 2, \ldots, nS$, is used to divide the $x_j y$ plane into four quadrants (Fig. 12):

$$Q_{i1} = \{(x_j, y) : x_{ij} < x_j, y_i < y\}, \tag{54}$$

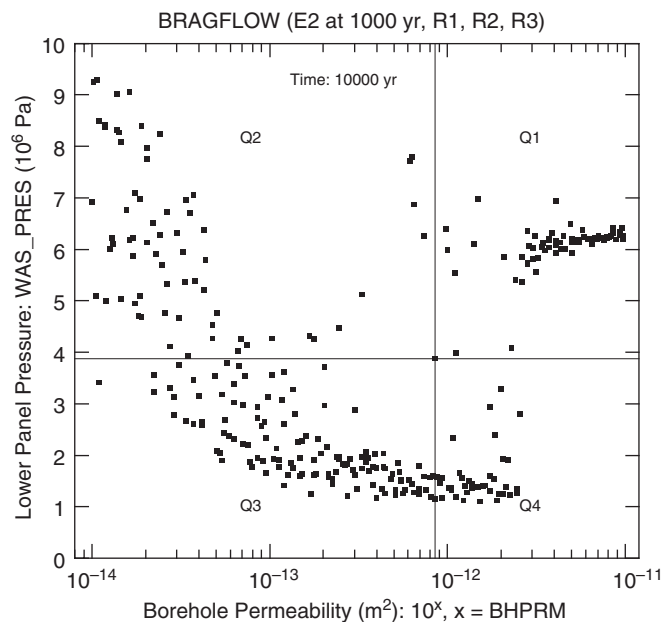

Fig. 12. Illustration of quadrants used with the two-dimensional KS test for the variable *WAS_PRES* at 10,000 yr.

Table 10
Comparison of formal statistical and Monte Carlo Determination of p-values for the SI Test and the two dimensional KS test for pressure (*WAS_PRES*) at 10,000 yr under undisturbed (i.e., *E*0) conditions (Fig. 5a) and disturbed (i.e., *E*2) conditions (Fig. 10a)

Variable[a]	SI test: 5 × 5[b]		SIMC test: 5 × 5[c]		KS test[d]		KSMC test[e]	
	p-Value	Rank	p-Value	Rank	p-Value	Rank	p-Value	Rank
Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a)								
WMICDFLG	0.0000	1	0.0000	1.5	0.0001	1	0.0000	1.5
HALPOR	0.0000	2	0.0000	1.5	0.0077	2	0.0000	1.5
WGRCOR	0.0003	3	0.0003	3	0.2979	3	0.0002	3
ANHPRM	0.0049	4	0.0031	4	0.8228	4	0.0257	4
ANHBCVGP	0.0194	5	0.0181	5	1.0000	24	0.4975	16
Pressure, disturbed (i.e., *E*2) conditions at 10,000 yr (Fig. 10a)								
BHPRM	0.0000	1	0.0000	1	0.0048	1	0.0000	1.5
HALPRM	0.0002	2	0.0003	3	0.1302	2	0.0000	1.5
WGRCOR	0.0002	3	0.0001	2	0.9609	5	0.1540	6
ANHPRM	0.0049	4	0.0039	4	0.6102	3	0.0023	3
HALPOR	0.3142	12	0.3164	12	0.7830	4	0.0178	4

[a]Variables ordered by p-values for SI test. Table includes only variables that had a p-value less than 0.05 for at least one of the procedures.
[b]p-Values and variable ranks for SI test with 5 × 5 grid (see Footnote b in Tables 5 and 7) determined from χ^2 distribution; see Eq. (20).
[c]p-Values and variable ranks for SI test with 5 × 5 grid (see Footnote b in Tables 5 and 7) determined with Monte Carlo procedure; see discussion associated with Eq. (21).
[d]p-Values and variable ranks for KS test determined from Eq. (61).
[e]p-Values and variable ranks for KS test determined with Monte Carlo procedure; see discussion associated with Eq. (21).

$$Q_{i2} = \{(x_j, y) : x_j < x_{ij}, y_i < y\}, \tag{55}$$

$$Q_{i3} = \{(x_j, y) : x_j < x_{ij}, y < y_i\}, \tag{56}$$

$$Q_{i4} = \{(x_j, y) : x_{ij} < x_j, y < y_i\}. \tag{57}$$

In turn, two fractions are defined for each quadrant:

$$fE_{ik} = \text{expected fraction of observations in quadrant } Q_{ik} \text{ if there is no relationship between } x_j \text{ and } y, \tag{58}$$

$$fO_{ik} = \text{observed fraction of observations in quadrant } Q_{ik}. \tag{59}$$

The quantity

$$D = \max\{|fE_{ik} - fO_{ik}|, k = 1,\ 2,\ 3,\ 4,\ i = 1,\ 2,\ \ldots,\ nS\} \tag{60}$$

is the KS statistic for the scatterplot.

The probability prob($\tilde{D} > D$) of exceeding D given that there is no relationship between x_j and y can be approximated by

$$\text{prob}(\tilde{D} > D) \cong Q_{\text{KS}}\left(\frac{D\sqrt{nS}}{1 + [1 - c(x_j, y)]^{1/2}\left[0.25 - 0.75/\sqrt{nS}\right]}\right), \tag{61}$$

where Q_{KS} is the function defined by

$$Q_{\text{KS}}(\lambda) = 2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-2j^2\lambda^2) \tag{62}$$

and $c(x_j, y)$ is the estimated CC between x_j and y (Ref. [157, Section 14.7]). Alternatively, prob($\tilde{D} > D$) can be estimated by a Monte Carlo procedure in which D is repeatedly estimated with randomly shuffled values (without replacement) of the x_{ij}'s and y_i's as previously illustrated in conjunction with Eq. (21) and Table 6 for the CMNs, CLs, CMDs and SI tests.

The result of applying the KS test is illustrated in Table 10, with p-values being calculated as indicated in Eq. (61) and also calculated with the previously indicated Monte Carlo procedure. This table also presents the results of using the SI test with a 5×5 grid. The direct calculation of p-values as indicated in Eq. (61) performs rather poorly and produces p-values that are much larger than those obtained with the Monte Carlo procedure. In contrast, the Monte Carlo calculation of p-values for the KS test produces results that are generally similar to, but not the same as, the results obtained with the SI test. In particular, the KS test with Monte Carlo calculation of p-values and the SI test agree on the most important variables but show some differences on the less-important variables.

Additional information: Ref. [157]; Refs. [173–175].

6.11. Tests for patterns based on distance measures

Tests for patterns based on distance measures provide possible alternatives to tests based on gridding as described in Sections 6.6 and 6.7. Distance-based tests for patterns have a potential advantage over grid-based tests in that they do not require the definition and use of a grid that can possibly influence the outcome of the test. Such tests have a long history of use in the ecological sciences [176–189].

Three distance-based tests will be illustrated: NN test, total distance (TD) test, and coefficient of aggregation (CA) test. Each of these tests involves the consideration of a set of points of the form [x_{ij}, y_i], $i = 1, 2, \ldots, nS$. Further, the x_{ij}'s and y_i's are assumed to be normalized to mean zero and standard deviation one.

The NN test [190] is based on the statistic

$$d_j = \sum_{i=1}^{nS} d_{ij}/nS, \tag{63}$$

where d_{ij} is the distance from the point (x_{ij}, y_i) to its NN among the points (x_{kj}, y_k) for $k = 1, 2, \ldots, nS$ and $k \neq i$. If x_j has an effect on y, then the value for d_j should tend to be smaller than would be the case if x_j had no effect on y. Determination of values $\tilde{d}_j$ for samples ($\tilde{x}_{ij}$, $\tilde{y}_i$), $i = 1, 2, \ldots, nS$, obtained by randomly pairing, without replacement, the values for the x_{ij}'s and y_i's in the original sample allows the determination of a distribution for d_j under the null hypothesis that there is no relationship between x_j and y. Thus, conditional on the observed distributions for x_j and y, the probability (i.e., a p-value) of obtaining a smaller value $\tilde{d}_j$ than the observed value d_j by chance alone can be determined. A small value for this probability (e.g., <0.01) indicates that x_j does indeed have an effect on y.

The TD test is a variant of the NN test and is based on the statistic

$$d_{tj} = \sum_{i=1}^{nS}\sum_{k=i+1}^{nS} d_{ik}/nD, \tag{64}$$

where d_{ik} is the distance between the points (x_{ij}, y_i) and (x_{kj}, y_k) and $nD = nS(nS-1)/2$ is the total number of distances d_{ik}. As for the NN statistic d_j, the value for d_{tj} will tend to be smaller than would otherwise be the case if x_j has an effect on y. Similarly to d_j, a Monte Carlo procedure can be used to develop a distribution for d_{tj} under the assumption that x_j has no effect on y. Then, conditional on the observed distributions for x_j and y, the probability of obtaining a smaller value for d_{tj} by chance alone can be estimated.

The CA test [179,191] is based on the statistic

$$A_j = \sum_{i=1}^{nS}\tilde{d}_{ij}^2 \Bigg/ \left[\sum_{i=1}^{nS} d_{ij}^2 + \sum_{i=1}^{nS}\tilde{d}_{ij}^2\right], \tag{65}$$

where d_{ij} is defined the same as in Eq. (63) for the NN test and $\tilde{d}_{ij}$ is defined similarly but for a sample ($\tilde{x}_{ij}$, $\tilde{y}_i$), $i = 1, 2, \ldots, nS$, obtained by randomly permuting the values for the x_{ij}'s and y_i's in the sample (x_{ij}, y_i), $i = 1, 2, \ldots, nS$. If x_j has an effect on y, then the value for A_j will tend to be larger than would otherwise be the case

because of the presence of $\sum_i d_{ij}^2$ in the denominator in the definition of A_j. A Monte Carlo procedure involving repeated calculations of A_j with two different random permutations of the x_{ij}'s and y_i's in the sample (x_{ij}, y_i), $i = 1, 2, \ldots, nS$, can be used to estimate a distribution for A_j under the assumption that x_j has no effect on y. Then, conditional on the observed distributions for x_j and y, the probability of obtaining a larger value for $\tilde{A}_j$ for A_j than the observed value by chance alone can be estimated.

The SI, NN, TD and CA tests are illustrated in Table 11. On the whole, the results obtained with the distance-based tests show considerable disagreement with results obtained with the SI test and also with other grid-based techniques illustrated in Table 5. Of the distance-based tests, the TD test compares best with results obtained with the grid-based techniques. Thus, this comparison suggests that the NN, TD and CA tests are less effective sensitivity analysis procedures than some of the other techniques introduced in this survey. However, the idea of using a grid-free, distance-based measure of sensitivity is very appealing. It is certainly possible that more appropriate distance-based measures of sensitivity can be found than those used in the presented tests. This is an area that merits additional investigation. For example, the use of rank-transformed data might yield more informative results.

Additional information: Refs. [176–189]; Ref. [192, Section 8.2.5].

6.12. Top down coefficient of concordance (TDCC)

The TDCC was introduced by Iman and Conover as a way to test agreement between different sensitivity analysis procedures [193]. However, it also provides a way to identify significant sets of variables in a sampling-based sensitivity analysis that does not rely on statistical tests predicated on distributional assumptions that may not be satisfied. In this application, the TDCC is used in a stepwise manner to test for agreement of sensitivity results obtained when a particular sensitivity analysis procedures is applied individually to each sample in a sequence of replicated samples of the same size. The significant variables are those which the TDCC indicates are identified as being important across all replicates.

The TDCC is based on the consideration of arrays of the form

$$\begin{array}{cccccc} & R_1 & R_2 & \ldots & R_{nR} \\ x_1 & r(O_{11}) & r(O_{12}) & \ldots & r(O_{1,nR}) \\ x_2 & r(O_{21}) & r(O_{22}) & \ldots & r(O_{2,nR}) \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ x_{nX} & r(O_{nX,1}) & r(O_{nX,2}) & \ldots & r(O_{nX,nR}), \end{array} \tag{66}$$

where (i) $x_1, x_2, \ldots, x_{nX}$ are the variables under consideration, (ii) $R_1, R_2, \ldots, R_{nR}$ designate the replicates, (iii) O_{jk} is the outcome (i.e., sensitivity measure) for variable x_j and replicate R_k, and (iv) $r(O_{jk})$, $j = 1, 2, \ldots, nX$, are the ranks assigned to the outcomes associated with replicate R_k. In the assigning of ranks, (i) a rank of 1 is assigned to the outcome O_{jk} with the largest value for $|O_{jk}|$, (ii) a rank of 2 is assigned the outcome O_{jk} with the second largest value for $|O_{jk}|$, and so on, and (iii) averaged ranks are assigned to equal values of O_{jk}. This is the reverse of the procedure used to assign ranks for use in rank regression.

Table 11
Comparison of tests for patterns based on distance measures for pressure (*WAS_PRES*) at 10,000 yr under undisturbed (i.e., *E*0) conditions (Fig. 5a) and disturbed (i.e., *E*2) conditions (Fig. 10a)

Variable[a]	SI test: 5×5[b]		NN test[c]		TD test[d]		CA test[e]	
	p-Value	Rank	*p*-Value	Rank	*p*-Value	Rank	*p*-Value	Rank
Pressure, undisturbed (i.e., *E*0) conditions at 10,000 yr (Fig. 5a)								
WMICDFLG	0.0000	1	0.0001	2	0.0000	2	0.6664	21
HALPOR	0.0000	2	0.0000	1	0.0000	2	0.0014	1
WGRCOR	0.0003	3	0.0327	3	0.0000	2	0.0049	2
ANHPRM	0.0049	4	0.3669	15	0.6348	21	0.3302	9
ANHBCVGP	0.0194	5	0.4745	7	0.4563	14	0.7544	24
Pressure, disturbed (i.e., *E*2) conditions at 10,000 yr (Fig. 10a)								
BHPRM	0.0000	1	0.0000	1	0.0000	2	0.0020	2
HALPRM	0.0002	2	0.3511	13	0.0000	2	0.0752	4
WGRCOR	0.0002	3	0.0095	2	0.7210	22	0.3420	12
ANHPRM	0.0049	4	0.0732	4	0.0000	2	0.0018	1
HALPOR	0.3142	12	0.2280	8	0.0210	4	0.2245	9

[a]Variables ordered by *p*-values for SI test. Table includes only variables that had a *p*-value less than 0.05 for at least one of the procedures.
[b]*p*-Values and variable ranks for SI test with 5×5 grid (see footnote b in Tables 5 and 7) determined from χ^2 distribution; see Eq. (20).
[c]*p*-Values and variable ranks for NN test (see Eq. (63)) determined with Monte Carlo procedures; see discussion associated with Eq. (21).
[d]Same as c but for TD test (see Eq. (64)).
[e]Same as c but for CA test (see Eq. (65)).

The TDCC is a measure of agreement between multiple rankings that emphasizes agreement between rankings assigned to important variables and deemphasizes disagreement between rankings assigned to less important/unimportant variables. For the TDCC, the ranks $r(O_{jk})$ in Eq. (66) are replaced by the corresponding Savage scores $ss(O_{ij})$, where

$$ss(O_{jk}) = \sum_{j=r(O_{jk})}^{nX} 1/j \tag{67}$$

and average Savage scores are assigned in the event of ties. The result is an array of the form

$$\begin{array}{cccccc} & R_1 & R_2 & \dots & R_{nR} \\ x_1 & \mathrm{ss}(O_{11}) & \mathrm{ss}(O_{12}) & \dots & \mathrm{ss}(O_{1,nR}) \\ x_2 & \mathrm{ss}(O_{21}) & \mathrm{ss}(O_{22}) & \dots & \mathrm{ss}(O_{2,nR}) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ x_{nX} & \mathrm{ss}(O_{nX,1}) & \mathrm{ss}(O_{nX,2}) & \dots & \mathrm{ss}(O_{nX,nR}), \end{array} \tag{68}$$

which has the same form as the array in Eq. (66) except that the ranks $r(O_{jk})$ have been replaced by the corresponding Savage scores $ss(O_{jk})$.

The TDCC is defined by

$$C_T = \frac{\{\sum_{j=1}^{nX}[\sum_{k=1}^{nR}\mathrm{ss}(O_{jk})]^2 - nR^2nX\}}{\{nR^2(nX - \sum_{j=1}^{nX}1/j)\}} \tag{69}$$

and is equivalent to Kendall's coefficient of concordance (Ref. [155, p. 305]) calculated with Savage scores rather than ranks. Under repeated random assignment of the integers in the columns of Eq. (66),

$$T = nR(nX-1)C_T \tag{70}$$

approximately follows a χ^2-distribution with nX–1 degrees of freedom and thus provides the basis for a statistical test of agreement.

The procedure to identify a significant set of variables with the TDCC operates in the following manner: (i) The sensitivity analysis technique in use (e.g., stepwise regression analysis) is applied to each replicate to rank variable importance. (ii) The TDCC is applied to the variable rankings obtained with each replicate to determine if there is a significant agreement between the replicates (e.g., as defined by a specified p-value for the TDCC). (iii) If there is significant agreement, the top ranked variable (i.e., rank 1) for each replicate is removed from consideration for all replicates; this results in the removal of one variable if all replicates assign the same variable a rank of 1 and more than one variable if different variables are assigned a rank of 1 in different replicates. (iv) A new sensitivity analysis is then performed for each replicate with the remaining variables, the remaining variables are reranked for each replicate, and Steps (ii) and (iii) are repeated with the reduced set of variables. (v) The process is continued until the deleted variable result in the analysis reaches a point at which the TDCC indicates that there is no significant agreement between the variable rankings obtained with the individual replicates. (vi) At this point, the analysis ends, and the significant set of variables are those deleted before the TDCC indicated no significant agreement between the variable rankings obtained with the individual replicates.

This procedure is illustrated for rank regression analysis with the three replicated random samples (i.e., RS1, RS2, RS3) from the variables in Table 1 for cumulative brine flow into the repository (*BRNREPTC*) at 1000 yr. The individual regression analyses all rank *HALPOR* as the most important variable (Table 12) and have a TDCC of 0.80 with a p-value of 5.2E-5 (Table 13). As a result, *HALPOR* is removed from consideration, which reduces the number of independent variables from 29 to 28. A new rank regression is then performed for each replicate with the remaining 28 variables, and the variables are reranked (i.e., from 1 to 28) on the basis of their SRRCs, with *ANHPRM* having a rank of 1 in one replicate and *WMICDFLG* having a rank of 1 in two replicates. For this new ranking (i.e., without *HALPOR*), the TDCC has a value of 0.71 with a p-value of 5.0E-4 (Table 13). As this is considered to be significant agreement, *ANHPRM* and *WMICDFLG* are dropped; the remaining 26 variables are reranked; new regressions are performed for each replicate; and a resultant TDCC of 0.46 with a p-value of 9.8E-2 is calculated (Table 13). If a p-value of 9.8E-2 is considered to be insignificant, then the analysis ends, and the set of significant variables is taken to be {*HALPOR*, *ANHPRM*, *WMICDFLG*}.

If a p-value of 9.8E-2 is considered to be significant (e.g., if the analysis was using 0.1 as the p-value above which the analysis stopped), then the analysis would continue with the top ranked variables in the individual replicates being dropped (i.e., *SALPRES*, *HALPRM*, *BPPRM*) and the TDCC recalculated for the remaining 23 variables. This process would continue until either an insignificant value for the TDCC was obtained or all variables were dropped, with the latter being an unlikely outcome.

Additional information: Refs. [138,193]. Content of this section is an adaptation of material contained in of Ref. [138, Sections 5 and 6].

6.13. Variance decomposition

An informative, but potentially computationally expensive, sensitivity analysis procedure is based on a complete variance decomposition of the uncertainty associated with y [56–59] With this procedure, the variance $V(y)$ of y is expressed as

$$V(y) = \sum_{j=1}^{nX} V_j + \sum_{j=1}^{nX}\sum_{k=j+1}^{nX} V_{jk} + \cdots + V_{12,\dots,nX}, \tag{71}$$

where V_j is the contribution of x_j to $V(y)$, V_{jk} is the contribution of the interaction of x_j and x_k to $V(y)$, and so on up to $V_{12,\dots,nX}$, which is the contribution of the interaction of x_1, x_2, …, x_{nX} to $V(y)$. Sensitivity measures

Table 12
Sensitivity analysis results based on SRRCs for three replicated random samples (RS1 RS2, RS3) of size 100 for cumulative brine flow into repository (*BRNREPTC*) at 1000 yr under undisturbed (i.e., *E*0) conditions (adapted from Ref. [138, Table 8])

Variable[a]	RS1[b]	RS2	RS3
HALPOR	9.93E−01(1)[c]	9.67E−01(1)	9.73E−01(1)
WMICDFLG	−9.72E-02(2)	−6.92E−02(4)	−1.13E−01(2)
ANHPRM	6.49E-02(3)	1.33E−01(2)	9.84E−02(3)
SALPRES	−4.00E−02(4)	$-2.70^{E}-03$(26)	−1.41E−02(13)
HALPRM	3.53E−02(5)	7.67E−02(3)	4.05E−02(5)
WRBRNSAT	−3.08E−02(6)	$-1.79^{E}-02$(14)	9.13E−03(17)
WASTWICK	−2.82E−02(7)	−2.27E−02(10)	−4.47E−03(21)
BPCOMP	−2.61E−02(8)	2.36E−02(9)	−8.05E−04(29)
SHPRMDRZ	2.29E−02(9)	$-1.37^{E}-02$(17)	2.58E−02(8)
BPPRM	−1.85E−02(10)	1.27E−02(19)	5.08E−02(4)
…	…	…	…
BPVOL	−1.58E−03(27)	6.54E−03(23)	4.64E−03(20)
ANHBCEXP	−1.30E−03(28)	4.32E−03(25)	2.88E−02(6)
WRGSSAT	−1.19E−03(29)	1.32E−02(18)	−5.33E−03(19)

[a]Variables in regression model ordered by SRRCs for sample RS1.
[b]SRRC in model containing all variables for indicated sample.
[c]Variable rank based on absolute value of SRRC for indicated sample.

Table 13
Sensitivity analysis with the TDCC for three replicated random samples of size 100 for cumulative brine flow into repository (*BRNREPTC*) at 1000 yr under undisturbed (i.e., *E*0) conditions (adapted from Ref. [138, Table 9])

Step[a]	TDCC[b]	*p*-Value[c]	Variable(s) removed[d]
1	0.80	5.2E-05	*HALPOR*
2	0.71	5.0E-04	*WMICDFLG, ANHPRM*
3	0.46	9.8E-02	*SALPRES, HALPRM, BPPRM*

[a]Steps in analysis.
[b]TDCC at beginning of step.
[c]*p*-Value for TDCC at beginning of step.
[d]Variable(s) removed at end of step.

are provided by

$$s_j = V_j/V(y) \text{ and } s_{jT} = \frac{V_j + \sum_{\substack{k=1\\k\neq j}}^{nX} V_{jk} + \cdots + V_{12,\ldots,nX}}{V(v)}, \tag{72}$$

where s_j is the fraction of $V(y)$ contributed by x_j alone and s_{jT} is the fraction of $V(y)$ contributed x_j and interactions of x_j with other variables.

The contributions to variance V_j, V_{jk}, …, $V_{12,\cdots,nX}$ in Eqs. (71) and (72) are defined by multidimensional integrals involving $y = f(\mathbf{x})$ and the individual elements x_j of $\mathbf{x}$. Specifically,

$$E(y) = \int_X f(\mathbf{x}) \prod_{j=1}^{nX} d_j(x_j) \prod_{j=1}^{nX} dx_j, \tag{73}$$

$$V(y) = \int_X [f(x) - E(y)]^2 \prod_{j=1}^{nX} d_j(x_j) \prod_{j=1}^{nX} dx_j = \int_X f^2(x) \prod_{j=1}^{nX} d_j(x_j) \prod_{j=1}^{nX} dx_j - E^2(y), \tag{74}$$

$$V_j = \int_{X_j} \left[\int_{X_{-j}} f(x) \prod_{\substack{k=1\\k\neq j}}^{nX} d_k(x_k) \prod_{\substack{k=1\\k\neq j}}^{nX} d_k x_k \right]^2 \times d_j(x_j) dx_j - E^2(y), \tag{75}$$

$$V_{jk} = \int_{X_j} \int_{X_k} \left[\int_{X_{-j,k}} f(x) \prod_{\substack{l=1\\l\neq j,k}}^{nX} d_l(x_l) \prod_{\substack{l=1\\l\neq j,k}}^{nX} dx_l \right]^2 \times d_j(x_j) d_k(x_k) dx_j dx_k - E^2(y) - V_j - V_k \tag{76}$$

and

$$V_j + \sum_{\substack{k=1\\k\neq j}}^{nX} V_{jk} + \cdots + V_{12,\ldots,nX} = V(f) - \left\{ \left[\int_{X_{-j}} \int_{X_j} \int_{\tilde{X}_j} f(\mathbf{x})\, f(\tilde{\mathbf{x}}) d_j(\tilde{x}_j) d_j(x_j) \prod_{\substack{k=1\\k\neq j}}^{nX} d_k(x_k) \right] \times d\tilde{x}_j dx_j \prod_{\substack{k=1\\k\neq j}}^{nX} dx_k - E^2(y) \right\}, \tag{77}$$

where (i) X_j is the sample space for x_j, $d_j(x_j)$ is the density function for x_j and the resultant quantities

$$X = \prod_{j=1}^{nX} X_j \text{ and } d(\mathbf{x}) = \prod_{j=1}^{nX} d_j(x_j)$$

are the sample space and density function, respectively, for $\mathbf{x}$, (ii) X_{-j} and $X_{-\mathrm{j,k}}$ correspond to the reduced sample spaces defined by

$$X_{-j} = \prod_{\substack{k=1\\k\neq j}}^{nX} X_k \text{ and } X_{-j,k} = \prod_{\substack{l=1\\l\neq j,k}}^{nX} X_l$$

and (iii) $X_j = \tilde{X}_j$ in Eq. (77) with the value for $\tilde{x}_j \in \tilde{X}_j$ replacing the value for $x_j \in X_j$ in the vector $\tilde{\mathbf{x}}$ (i.e., the variables x_j and $\tilde{x}_j$ associated with X_j and $\tilde{X}_j$ have identical distributions but are assumed to be independent and the vectors $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are the same except that x_j appears as element j in $\mathbf{x}$ and $\tilde{x}_j$ appears as element j in $\tilde{\mathbf{x}}$).

As a result, the determination of s_j and s_{jT} is a problem in the evaluation of multidimensional integrals. In practice, this evaluation is carried out with sampling-based methods of the form indicated in the following algorithm.

Step 1: Generate a random or LHS

$$\mathbf{x}_i = [x_{i1},\ x_{i2}, \ldots,\ x_{i,nX}], \quad i = 1,\ 2,\ \ldots,\ nS \tag{78}$$

from $\mathbf{x} = [x1, x2, \ldots, xnX]$ in consistency with the distributions assigned to the individual x_j.

Step 2: Estimate the mean and variance for y with the approximations

$$\hat{E}(y) = \sum_{i=1}^{nS} f(x_i)/nS \tag{79}$$

and

$$\hat{V}(y) = \sum_{i=1}^{nS} [f(x_i) - \hat{E}(y)]^2/nS = \sum_{i=1}^{nS} f^2(x_i)/nS - \hat{E}^2(y). \tag{80}$$

The estimation of $\hat{E}(y)$ and $\hat{V}(y)$ requires nS evaluations of the function f.

Step 3: Generate a second random or LHS

$$\mathbf{r}_i = [r_{i1},\ r_{i2},\ \ldots,\ r_{i,nX}], \quad i = 1,\ 2,\ \ldots,\ nS \tag{81}$$

by randomly permuting, without replacement, the individual variable values associated with the sample generated in Step 1.

Step 4: For each variable x_j, generate a reordering

$$\mathbf{r}_{ij} = [r_{ij1},\ r_{ij2},\ \ldots,\ r_{ij,nX}], \quad i = 1,\ 2,\ \ldots,\ nS \tag{82}$$

of the sample generated in Step 3 such that $r_{ijj} = x_{ij}$. This step only involves a change in the numbering associated with the sample generated in Step 3 for each x_j; no changes to the sample itself are involved.

Step 5: For each variable x_j, estimate s_j by

$$s_j \cong \left[\sum_{i=1}^{nS} f(\mathbf{x}_i) f(\mathbf{r}_{ij})/nS - \hat{E}^2(y)\right] \Big/ \hat{V}(y). \tag{83}$$

The estimation of s_j for all x_j requires only nS additional evaluations of the function f as a result of the efficient reuse of the function evaluations for the sample generated in Step 3.

Step 6: For each variable x_j, generate an additional sample

$$\mathbf{x}_{ij} = [x_{ij1},\ x_{ij2},\ \ldots,\ x_{ij,nX}], \quad i = 1,\ 2,\ \ldots,\ nS, \tag{84}$$

where x_{ijj} is generated as a random or LHS from x_j and $x_{ijk} = x_{ik}$ for $k \neq j$. The sample generated for x_j in this step differs from the sample generated in Step 1 only in the values associated with x_j.

Step 7: Estimate s_{jT} by

$$s_{jT} \cong \sum_{i=1}^{nS} f(\mathbf{x}_i)[f(\mathbf{x}_i) - f(\mathbf{x}_{ij})]/[nS\hat{V}(y)] \tag{85}$$

for each x_j. The estimation of s_{jT} for all x_j requires an additional $(nX)(nS)$ evaluations of the function f.

Although the sensitivity measures s_j and s_{jT} provide valuable sensitivity information, their determination can be computationally expensive due to the large number of function evaluations that could be required. Specifically, $2(nS)$, $(nX+1)(nS)$ and $(nX+2)(nS)$ function evaluations are required to estimate s_j, s_{jT} and both s_j and s_{jT}, respectively, for nX uncertain variables. Further, because integrals are being approximated, the basic sample size nS required for the preceding algorithm to produce acceptable approximations to s_j and s_{jT} is likely to be larger than the sample sizes required for other sampling-based sensitivity measures.

Sensitivity analysis based on variance decomposition is illustrated with a simple test function introduced as part of a review of uncertainty and sensitivity analysis procedures (Ref. [194, Model 9]). Specifically, this test function is defined by

$$\begin{aligned} y &= f(x), x = [x_1,\ x_2,\ x_3] \\ &= \sin x_1 + A \sin^2 x_2 + B x_3^4 \sin x_1 \end{aligned} \tag{86}$$

with $A = 7$, $B = 0.1$ and each x_j uniform on $[-\pi, \pi]$. Unfortunately, the fluid flow model that has been used to illustrate other sensitivity analysis procedures is too computationally demanding for use with the procedures discussed in this section. Values of s_j and s_{jT} obtained with a base sample size of $nS = 10,000$ are

$$S_1 = 0.30,\ s_2 = 0.46,\ s_3 = 0.00 \tag{87}$$

and

$$s_{1\,T} = 0.53,\ s_{2\,T} = 0.45,\ s_{3\,T} = 0.23. \tag{88}$$

Further, results obtained with different values for nS are illustrated in Table 14 and suggest that the approximations of the integrals appearing in the definitions of s_j and s_{jT} are close to being converged with $nS = 10,000$.

For perspective, sensitivity results based on CCs, RCCs, CMNs, CLs, CMDs and SI are presented in Table 15 and scatterplots for x_1, x_2 and x_3 are given in Fig. 13. The model in Eq. (86) was constructed to have patterns that would be difficult to identify with regression-based sensitivity analysis procedures. Thus, although x_2 is a major contributor to the uncertainty in y, this effect is completely missed by the analyses based on CCs and RCCs in Table 15 owing to the oscillatory relationship between x_2 and y (Fig. 13b). Similarly, the CMDs test does not identify x_3 as having an effect on y owing to the constancy of the median values for y across the range of x_3 (Fig. 13c). Of the tests presented in Table 15, the SI test has the best performance and gives a reasonable indication of the importance of x_1, x_2 and x_3 with respect to the uncertainty in y for $nS = 100$ and 1000. This is not surprising as the SI test is effective at identifying nonlinear relationships. Fullest representation of the effects of x_1, x_2 and x_3 on uncertainty in y is given by the variance decomposition results in Eqs. (87) and (88). However, this enhanced resolution comes at a cost as the results in Eqs. (87) and (88) required more function evaluations (i.e., $nS = 10,000$) than the SI results (i.e., $nS = 100$ and $nS = 1000$) in Table 15.

Additional Information: Refs. [56–60,195–210].

Table 14
Evaluation of variance decompositions s_j and s_{jT} for model in Eq. (86) with different sample sizes

nS[a]	$\hat{E}(y)$[b]	$\hat{V}(y)$[c]	$\hat{s}_1$[d]	$\hat{s}_2$[d]	$\hat{s}_3$[d]	$\hat{s}_{1T}$[e]	$\hat{s}_{2T}$[e]	$\hat{s}_{3T}$[e]
10	3.7	16.5	0.70	0.65	−0.04	0.84	−0.09	−0.24
100	3.9	13.1	0.10	0.37	−0.24	0.79	0.80	0.45
1000	3.5	14.2	0.30	0.44	−0.02	0.56	0.53	0.24
10,000	3.5	14.0	0.30	0.46	0.00	0.53	0.45	0.23
100,000	3.5	13.9	0.32	0.44	−0.00	0.56	0.44	0.24
1,000,000	3.5	13.8	0.32	0.44	0.00	0.56	0.44	0.24

[a]Sample size.
[b]Estimate for expected value of y; see Eqs. (74) and (79).
[c]Estimate for variance of y; see Eqs. (74) and (80).
[d]Estimate for contribution of x_j, $j = 1, 2, 3$, to variance of y; see Eqs. (72) and (83).
[e]Estimate for contribution of x_j, $j = 1, 2, 3$, and its interactions with the other two variables to the variance of y; see Eqs. (72) and (85).

Table 15
Sensitivity results based on CCs, RCCs, CMNs, CLs, CMDs and SI for model in Eq. (86) (Ref. [101, Table 9.14])

Variable name[a]	CC[b]		RCC[c]		CMN: 1×5[d]		CL: 1×5[e]		CMD: 2×5[f]		SI: 5×5[g]	
	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val
Sample size $nLHS = 100$												
x_1	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	2.0	0.0001	1.0	0.0000
x_3	2.0	0.5667	2.0	0.6361	3.0	0.6917	3.0	0.5495	3.0	0.9384	3.0	0.0615
x_2	3.0	0.8327	3.0	0.8393	2.0	0.0000	2.0	0.0000	1.0	0.0000	2.0	0.0008
Sample size $nLHS = 1000$												
x_1	1.0	0.0000	1.0	0.0000	1.5	0.0000	1.5	0.0000	2.0	0.0000	1.5	0.0000
x_3	2.0	0.0162	2.0	0.0187	3.0	0.0438	3.0	0.0347	3.0	0.1446	3.0	0.0000
x_2	3.0	0.9799	3.0	0.9999	1.5	0.0000	1.5	0.0000	1.0	0.0000	1.5	0.0000

[a]Variables ordered by p-values for CCs.
[b]Ranks and p-values for CCs; see Eq. (24), Ref. [47].
[c]Ranks and p-values for RCCs; see Eq. (38), Ref. [47].
[d]Ranks and p-values for CMNs test with 1×5 grid; see Eq. (15).
[e]Ranks and p-values for CLs test with 1×5 grid; see Eq. (16)
[f]Ranks and p-values for CMDs test with 2×5 grid; see Eq. (18).
[g]Ranks and p-values for SI test with 5×5 grid; see Eq. (20).

7. Summary

Sampling-based uncertainty and sensitivity analysis is widely used, and as a result, is a fairly mature area of study. However, there remain a number of important challenges and areas for additional study. For example, there is a need for sensitivity analysis procedures that are more effective at revealing nonlinear relations than the parametric regression procedures (Section 6.3) and partial correlation procedures (Section 6.4) currently in wide use. Among the approaches to sensitivity analysis described in the preceding section, statistical tests for patterns based on gridding (Section 6.6), nonparametric regression (Section 6.8), the SRD/rank correlation test (Section 6.9), the two dimensional KS test (Section 6.10), and complete variance decomposition (Section 6.13) have not been as widely used as approaches based on parametric regression and partial correlation and merit additional investigation and use. As another example, sampling-based procedures for uncertainty and sensitivity analysis usually use probability as the model, or representation, for uncertainty. However, when limited information is available with which to characterize uncertainty, probabilistic characterizations can give the appearance of more knowledge than is really present. Alternative representations for uncertainty such as evidence theory and possibility theory merit consideration for their potential to represent uncertainty in situations where little information is available [84–92]. Finally, a significant challenge is the education of potential users of uncertainty and sensitivity analysis about (i) the importance of such analyses and their role in both large and small analyses, (ii) the need for appropriate separation of aleatory and epistemic uncertainty in the conceptual and computational implementation of analyses of complex systems [15–24], (iii) the need for a clear conceptual view of what an analysis is intended to represent and a computational design that is consistent with that view [15,124,211,212], (iv) the role that uncertainty and sensitivity analysis plays in model and

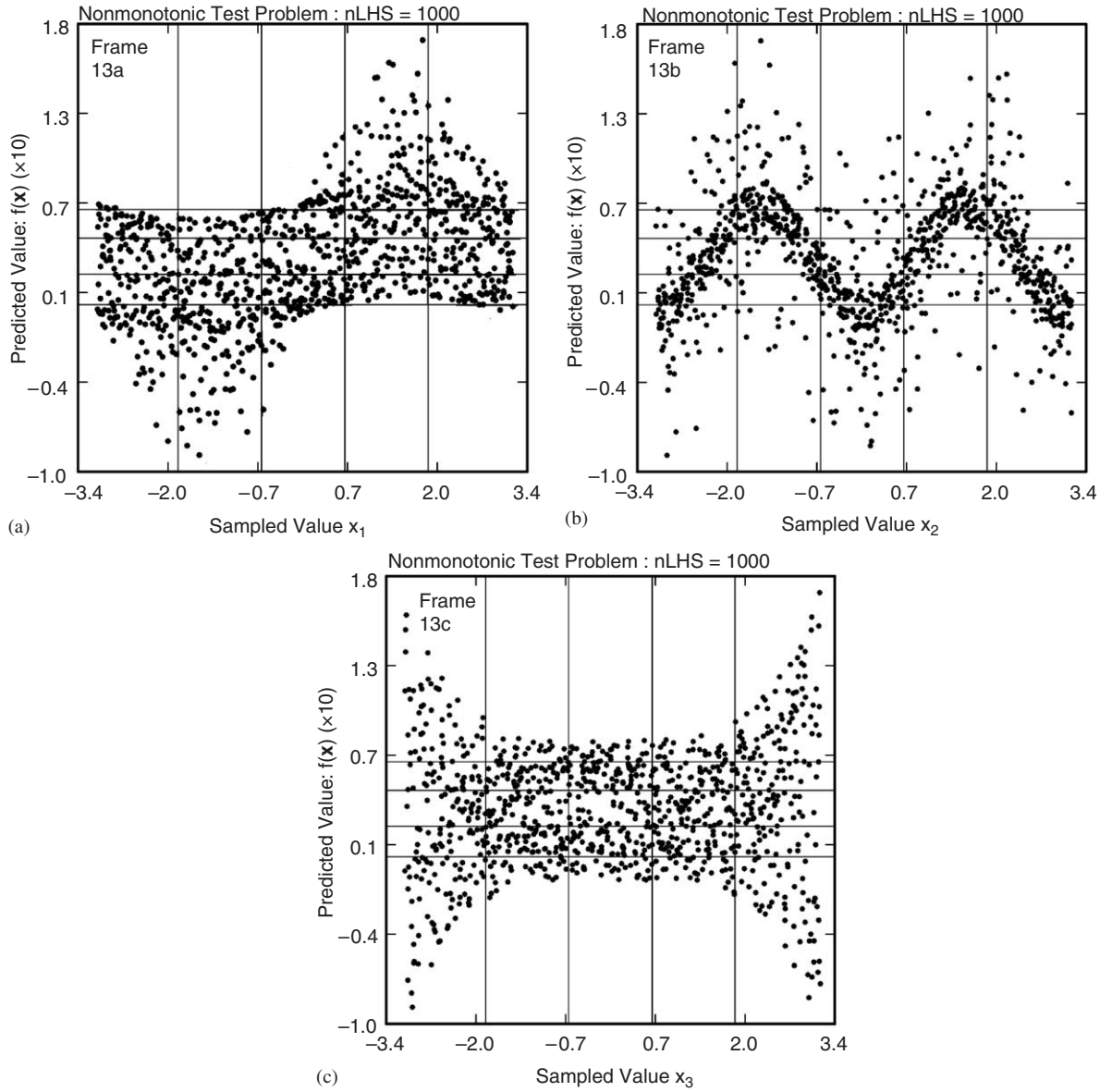

Fig. 13. Scatterplots for model in Eq. (86) with grid for SI test with $nI = nD = 5$ (adapted from Ref. [101, Fig. 9.15]).

analysis verification [5,6], and (v) the importance of avoiding deliberately conservative assumptions if meaningful uncertainty and sensitivity analysis results are to be obtained [213–217].

Acknowledgements

Work performed for Sandia National Laboratories (SNL), which is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Security Administration under contract DE-AC04-94AL-85000. Review at SNL provided by Tony Giunta and Brian Rutherford. Editorial support provided by F. Puffer, K. Best and J. Ripple of Tech Reps, a division of Ktech Corporation.

References

[1] Christie MA, Glimm J, Grove JW, Higdon DM, Sharp DH, Wood-Schultz MM. Error analysis and simulations of complex phenomena. Los Alamos Sci 2005;29:6–25.
[2] Nikolaidis E, Ghiocel DM, Singhal S, editors. Engineering design reliability handbook. Boca Raton, FL: CRC Press; 2004.
[3] Sharp DH, Wood-Schultz MM. QMU and nuclear weapons certification: what's under the hood? Los Alamos Sci 2003;28:47–53.
[4] Wagner RL. Science, uncertainty and risk: the problem of complex phenomena. APS News 2003;12(1):8.
[5] Oberkampf WL, DeLand SM, Rutherford BM, Diegert KV, Alvin KF. Error and uncertainty in modeling and simulation. Reliab Eng Syst Safety 2002;75(3):333–57.
[6] Roache PJ. Verification and validation in computational science and engineering. Albuquerque, NM: Hermosa Publishers; 1998.
[7] Ayyub BM, editor. Uncertainty modeling and analysis in civil engineering. Boca Raton, FL: CRC Press; 1997.
[8] Risk Assessment Forum. Guiding principles for Monte Carlo analysis, EPA/630/R-97/001. Washington, DC: RiskAssessmentForum

Environmental Protection Agency; 1997 Available from the NTIS as PB97-188106/XAB.
[9] NCRP (National Council on Radiation Protection and Measurements). A guide for uncertainty analysis in dose and risk assessments related to environmental contamination, NCRP Commentary No. 14. Bethesda, MD: National Council on Radiation Protection and Measurements; 1996.
[10] NRC (National Research Council). Science and judgment in risk assessment. Washington, DC: National Academy Press; 1994.
[11] NRC (National Research Council). Issues in risk assessment. Washington, DC: National Academy Press; 1993.
[12] US EPA (US Environmental Protection Agency). An SAB Report: multi-media risk assessment for radon, review of uncertainty analysis of risks associated with exposure to radon, EPA-SAB-RAC-93-014. Washington, DC: US Environmental Protection Agency; 1993.
[13] IAEA (International Atomic Energy Agency). Evaluating the reliability of predictions made using environmental transfer models, safety series no. 100. Vienna: International Atomic Energy Agency; 1989.
[14] Beck MB. Water-quality modeling: a review of the analysis of uncertainty. Water Resour Res 1987;23(8):1393–442.
[15] Helton JC. Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. J Stat Comput Simul 1997;57(1–4):3–76.
[16] Helton JC, Burmaster DE, editors. Guest editorial: treatment of aleatory and epistemic uncertainty in performance assessments for complex systems. Reliab Eng Syst Safety 1996;54(2–3):91–4.
[17] Paté-Cornell ME. Uncertainties in Risk Anal: six levels of treatment. Reliab Eng Syst Safety 1996;54(2–3):95–111.
[18] Winkler RL. Uncertainty in probabilistic risk assessment. Reliab Eng Syst Safety 1996;54(2–3):127–32.
[19] Hoffman FO, Hammonds JS. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 1994;14(5):707–12.
[20] Helton JC. Treatment of uncertainty in performance assessments for complex systems. Risk Anal 1994;14(4):483–511.
[21] Apostolakis G. The concept of probability in safety assessments of technological systems. Science 1990;250(4986):1359–64.
[22] Haan CT. Parametric uncertainty in hydrologic modeling. Trans ASAE 1989;32(1):137–46.
[23] Parry GW, Winter PW. Characterization and evaluation of uncertainty in probabilistic Risk Anal. Nucl Safety 1981;22(1): 28–42 [DIRS 159059].
[24] Kaplan S, Garrick BJ. On the quantitative definition of risk. Risk Anal 1981;1(1):11–27.
[25] Cacuci DG. Sensitivity and uncertainty analysis, vol. 1: theory. Boca Raton, FL: Chapman & Hall/CRC Press; 2003.
[26] Griewank A. Evaluating derivatives: principles and techniques of algorithmic differentiation. Philadelphia, PA: Society for Applied and Industrial Mathematics; 2000.
[27] Berz M, Bischof C, Corliss G, Griewank A. Computational differentiation: techniques, applications, and tools. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1996.
[28] Cacuci DG, Schlesinger ME. On the application of the adjoint method of sensitivity analysis to problems in the atmospheric sciences. Atmósfera 1994;7(1):47–59.
[29] Turányi T. Sensitivity analysis of complex kinetic systems, tools and applications. J Math Chem 1990;5(3):203–48.
[30] Rabitz H, Kramer M, Dacol D. Sensitivity analysis in chemical kinetics. In: Rabinovitch BS, Schurr JM, Strauss HL, editors. Annual review of physical chemistry, vol. 34. Palo Alto, CA: Annual Reviews Inc.; 1983. p. 419–61.
[31] Lewins J, Becker M, editors. Sensitivity and uncertainty analysis of reactor performance parameters. Advances in nuclear science and technology, vol. 14. New York, NY: Plenum Press; 1982.
[32] Frank PM. Introduction to system sensitivity theory. New York, NY: Academic Press; 1978.
[33] Tomovic R, Vukobratovic M. General sensitivity theory. New York, NY: Elsevier; 1972.
[34] Myers RH, Montgomery DC, Vining GG, Borror CM, Kowalski SM. Response surface methodology: a retrospective and literature review. J Quality Technol 2004;36(1):53–77.
[35] Myers RH. Response surface methodology—current status and future directions. J Quality Technol 1999;31(1):30–44.
[36] Andres TH. Sampling methods and sensitivity analysis for large parameter sets. J Stat Comput Simul 1997;57(1–4):77–110.
[37] Kleijnen JPC. Sensitivity analysis and related analyses: a review of some statistical techniques. J Stat Comput Simul 1997;57(1–4): 111–42.
[38] Kleijnen JPC. Sensitivity analysis of simulation experiments: regression analysis and statistical design. Math Comput Simul 1992;34(3–4):297–315.
[39] Myers RH, Khuri AI, Carter J, Walter H. Response surface methodology: 1966–1988. Technometrics 1989;31(2):137–57.
[40] Sacks J, Welch WJ, Mitchel TJ, Wynn HP. Design and analysis of computer experiments. Stat Sci 1989;4(4):409–35.
[41] Morton RH. Response surface methodology. Math Sci 1983; 8:31–52.
[42] Mead R, Pike DJ. A review of response surface methodology from a biometric viewpoint. Biometrics 1975;31:803–51.
[43] Myers RH. Response surface methodology. Boston, MA: Allyn & Bacon; 1971.
[44] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Safety 2003;81(1):23–69.
[45] Helton JC, Davis FJ. Illustration of sampling-based methods for uncertainty and sensitivity analysis. Risk Anal 2002;22(3):591–622.
[46] Helton JC, Davis FJ. Sampling-based methods. In: Saltelli A, Shan K, Scott EM, editors. Sensitivity analysis. New York, NY: Wiley; 2000. p. 101–53.
[47] Kleijnen JPC, Helton JC. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1: review and comparison of techniques. Reliab Eng Syst Safety 1999;65(2): 147–85.
[48] Blower SM, Dowlatabadi H. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Int Stat Rev 1994;62(2):229–43.
[49] Saltelli A, Andres TH, Homma T. Sensitivity analysis of model output. An investigation of new techniques. Comput Stat Data Anal 1993;15(2):445–60.
[50] Iman RL. Uncertainty and sensitivity analysis for computer modeling applications. In: Cruse TA, editor. Reliability technology—1992, the winter annual meeting of the American society of mechanical engineers, Anaheim, California, November 8–13, vol. 28. New York, NY: American Society of Mechanical Engineers, Aerospace Division; 1992. p. 153–68.
[51] Saltelli A, Marivoet J. Non-parametric statistics in sensitivity analysis for model output. A comparison of selected techniques. Reliab Eng Syst Safety 1990;28(2):229–53.
[52] Iman RL, Helton JC, Campbell JE. An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment. J Quality Technol 1981;13(3):174–83.
[53] Iman RL, Helton JC, Campbell JE. An approach to sensitivity analysis of computer models, Part 2. Ranking of input variables, response surface validation, distribution effect and technique synopsis. J Quality Technol 1981;13(4):232–40.
[54] Iman RL, Conover WJ. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. Commun Stat: Theory Methods A 1980;9(17):1749–842.
[55] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979;21(2):239–45.

[56] Li G, Rosenthal C, Rabitz H. High-dimensional model representations. J Phys Chem 2001;105(33):7765–77.
[57] Rabitz H, Alis OF. General foundations of high-dimensional model representations. J Math Chem 1999;25(2–3):197–233.
[58] Saltelli A, Tarantola S, Chan KP-S. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 1999;41(1):39–56.
[59] Sobol' IM. Sensitivity estimates for nonlinear mathematical models. Math Modeling Comput Exp 1993;1(4):407–14.
[60] Cukier RI, Levine HB, Shuler KE. Nonlinear sensitivity analysis of multiparameter model systems. J Comput Phys 1978;26(1): 1–42.
[61] Ionescu-Bujor M, Cacuci DG. A comparative review of sensitivity and uncertainty analysis of large-scale systems–I: deterministic methods. Nucl Sci Eng 2004;147(3):189–2003.
[62] Cacuci DG, Ionescu-Bujor M. A comparative review of sensitivity and uncertainty analysis of large-scale systems–II: statistical methods. Nucl Sci Eng 2004;147(3):204–17.
[63] Frey HC, Patil SR. Identification and review of sensitivity analysis methods. Risk Anal 2002;22(3):553–78.
[64] Saltelli A, Chan K, Scott EM, editors. Sensitivity analysis. New York, NY: Wiley; 2000.
[65] Hamby DM. A review of techniques for parameter sensitivity analysis of environmental models. Environ Monitor Assess 1994; 32(2):135–54.
[66] Helton JC. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliab Eng Syst Safety 1993;42(2–3):327–67.
[67] Ronen Y. Uncertainty analysis. Boca Raton, FL: CRC Press, Inc.; 1988.
[68] Iman RL, Helton JC. An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Anal 1988;8(1): 71–90.
[69] Blower SM, Gershengorn HB, Grant RM. A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science 2000; 287(5453):650–4.
[70] Cohen C, Artois M, Pontier D. A discrete-event computer model of feline herpes virus within cat populations. Prevent Vet Med 2000;45(3–4):163–81.
[71] Hofer E. Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. Comput Phys Commun 1999;117(1–2):21–34.
[72] Caswell H, Brault S, Read AJ, Smith TD. Harbor porpoise and fisheries: an uncertainty analysis of incidental mortality. Ecol Appl 1998;8(4):1226–38.
[73] Sanchez MA, Blower SM. Uncertainty and sensitivity analysis of the basic reproductive rate: tuberculosis as an example. Am J Epidemiol 1997;145(12):1127–37.
[74] Chan MS. The consequences of uncertainty for the prediction of the effects of schistosomiasis control programmes. Epidemiol Infect 1996;117(3):537–50.
[75] Gwo JP, Toran LE, Morris MD, Wilson GV. Subsurface stormflow modeling with sensitivity analysis using a latin-hypercube sampling technique. Ground Water 1996;34(5):811–8.
[76] Helton JC, Anderson DR, Baker BL, Bean JE, Berglund JW, Beyeler W, et al. Uncertainty and sensitivity analysis results obtained in the 1992 performance assessment for the waste isolation pilot plant. Reliab Eng Syst Safety 1996;51(1):53–100.
[77] Kolev NI, Hofer E. Uncertainty and sensitivity analysis of a postexperiment simulation of nonexplosive melt-water interaction. Exp Thermal Fluid Sci 1996;13(2):98–116.
[78] Whiting WB, Tong T-M, Reed ME. Effect of uncertainties in thermodynamic data and model parameters on calculated process performance. Ind Eng Chem Res 1993;32(7):1367–71.
[79] Ma JZ, Ackerman E, Yang J-J. Parameter sensitivity of a model of viral epidemics simulated with Monte Carlo techniques. I. Illness attack rates. Int J Biomed Comput 1993;32(3–4):237–53.
[80] Ma JZ, Ackerman E. Parameter sensitivity of a model of viral epidemics simulated with Monte Carlo techniques. II. Durations and peaks. Int J Biomed Comput 1993;32(3–4):255–68.
[81] Breshears DD, Kirchner TB, Whicker FW. Contaminant transport through agroecosystems: assessing relative importance of environmental, physiological, and management factors. Ecol Appl 1992; 2(3):285–97.
[82] Breeding RJ, Helton JC, Gorham ED, Harper FT. Summary description of the methods used in the probabilistic risk assessments for NUREG-1150. Nucl Eng Des 1992;135(1):1–27.
[83] MacDonald RC, Campbell JE. Valuation of supplemental and enhanced oil recovery projects with risk analysis. J Petroleum Technol 1986;38(1):57–69.
[84] Helton JC, Johnson JD, Oberkampf WL. An exploration of alternative approaches to the representation of uncertainty in model predictions. Reliab Eng Syst Safety 2004;85(1–3):39–71.
[85] Klir GJ. Generalized information theory: aims, results, and open problems. Reliab Eng Syst Safety 2004;85(1–3):21–38.
[86] Ross TJ. Fuzzy logic with engineering applications. 2nd ed. New York, NY: Wiley; 2004.
[87] Halpern JY. Reasoning about uncertainty. Cambridge, MA: MIT Press; 2003.
[88] Ross TJ, Booker JM, Parkinson WJ, editors. Fuzzy logic and probability applications: bridging the gap. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2002.
[89] Jaulin L, Kieffer M, Didrit O, Walter E. Applied interval analysis. New York, NY: Springer; 2001.
[90] Wolkenhauer O. Data engineering: fuzzy mathematics in systems theory and data analysis. New York, NY: Wiley; 2001.
[91] Klir GJ, Wierman MJ. Uncertainty-based information. New York, NY: Physica-Verlag; 1999.
[92] Yager RR, Kacprzyk J, Fedrizzi M, editors. Advances in the dempster-shafer theory of evidence. New York, NY: Wiley; 1994.
[93] Cooke RM, Goossens LHJ. Expert judgement elicitation for risk assessment of critical infrastructures. J Risk Res 2004;7(6):643–56.
[94] Ayyub BM. Elicitation of expert opinions for uncertainty and risks. Boca Raton, FL: CRC Press; 2001.
[95] McKay M, Meyer M. Critique of and limitations on the use of expert judgements in accident consequence uncertainty analysis. Radiat Protect Dosim 2000;90(3):325–30.
[96] Budnitz RJ, Apostolakis G, Boore DM, Cluff LS, Coppersmith KJ, Cornell CA, et al. Use of technical expert panels: applications to probabilistic seismic hazard analysis. Risk Anal 1998;18(4):463–9.
[97] Thorne MC, Williams MMR. A review of expert judgement techniques with reference to nuclear safety. Progr Nucl Safety 1992;27(2–3):83–254.
[98] Cooke RM. Experts in uncertainty: opinion and subjective probability in science. Oxford; NY: Oxford University Press; 1991.
[99] Meyer MA, Booker JM. Eliciting and analyzing expert judgment: a practical guide. New York, NY: Academic Press; 1991.
[100] Hora SC, Iman RL. Expert opinion in risk analysis: the NUREG-1150 methodology. Nucl Sci Eng 1989;102(4):323–31.
[101] Helton JC, Davis FJ. Sampling-based methods for uncertainty and sensitivity analysis, SAND99-2240. Albuquerque, NM: Sandia National Laboratories; 2000.
[102] Vaughn P, Bean JE, Helton JC, Lord ME, MacKinnon RJ, Schreiber JD. Representation of two-phase flow in the vicinity of the repository in the 1996 performance assessment for the waste isolation pilot plant. Reliab Eng Syst Safety 2000;69(1–3): 205–26.
[103] Helton JC, Bean JE, Economy K, Garner JW, MacKinnon RJ, Miller J, et al. Uncertainty and sensitivity analysis for two-phase flow in the vicinity of the repository in the 1996 performance assessment for the waste isolation pilot plant: undisturbed conditions. Reliab Eng Syst Safety 2000;69(1–3):227–61.
[104] Helton JC, Bean JE, Economy K, Garner JW, MacKinnon RJ, Miller J, et al. Uncertainty and sensitivity analysis for two-phase flow in the vicinity of the repository in the 1996 performance

assessment for the waste isolation pilot plant: disturbed conditions. Reliab Eng Syst Safety 2000;69(1–3):263–304.
[105] US DOE (US Department of Energy). Title 40 CFR Part 191 compliance certification application for the waste isolation pilot plant, DOE/CAO-1996–2184. US Department of Energy, Carlsbad Area Office,; 1996 [DIRS 100975].
[106] Helton JC, Bean JE, Berglund JW, Davis FJ, Economy K, Garner JW, et al. Uncertainty and sensitivity analysis results obtained in the 1996 performance assessment for the waste isolation pilot plant, SAND98-0365. Albuquerque, NM: Sandia National Laboratories; 1998.
[107] Helton JC, Marietta MG. Special issue: the 1996 performance assessment for the waste isolation pilot plant. Reliab Eng Syst Safety 2000;69(1–3):1–451.
[108] Helton JC, Martell M-A, Tierney MS. Characterization of subjective uncertainty in the 1996 performance assessment for the waste isolation pilot plant. Reliab Eng Syst Safety 2000;69(1–3): 191–204.
[109] Goossens LHJ, Harper FT, Kraan BCP, Metivier H. Expert judgement for a probabilistic accident consequence uncertainty analysis. Radiat Protect Dosim 2000;90(3):295–301.
[110] Goossens LHJ, Harper FT. Joint EC/USNRC expert judgement driven radiological protection uncertainty analysis. J Radiol Protect 1998;18(4):249–64.
[111] Siu NO, Kelly DL. Bayesian parameter estimation in probabilistic risk assessment. Reliab Eng Syst Safety 1998;62(1–2):89–116.
[112] Evans JS, Gray GM, Sielken Jr RL, Smith AE, Valdez-Flores C, Graham JD. Use of probabilistic expert judgement in uncertainty analysis of carcinogenic potency. Regul Toxicol Pharmacol 1994;20(1, pt. 1):15–36.
[113] Chhibber S, Apostolakis G, Okrent D. A taxonomy of issues related to the use of expert judgments in probabilistic safety studies. Reliab Eng Syst Safety 1992;38(1–2):27–45.
[114] Kaplan S. Expert information versus expert opinions: another approach to the problem of eliciting combining using expert knowledge in PRA. Reliab Eng Syst Safety 1992;35(1):61–72.
[115] Otway H, Winterfeldt DV. Expert judgement in risk analysis and management: process, context, and pitfalls. Risk Anal 1992;12(1): 83–93.
[116] Bonano EJ, Apostolakis GE. Theoretical foundations and practical issues for using expert judgments in uncertainty analysis of high-level radioactive waste disposal. Radioactive Waste Manage Nucl Fuel Cycle 1991;16(2):137–59.
[117] Keeney RL, Winterfeldt DV. Eliciting probabilities from experts in complex technical problems. IEEE Trans Eng Manage 1991;38(3): 191–201.
[118] Svenson O. on expert judgments in safety analyses in the process industries. Reliab Eng Syst Safety 1989;25(3):219–56.
[119] Mosleh A, Bier VM, Apostolakis G. A critique of current practice for the use of expert opinions in probabilistic risk assessment. Reliab Eng Syst Safety 1988;20(1):63–85.
[120] Breeding RJ, Helton JC, Murfin WB, Smith LN, Johnson JD, Jow H-N, et al. The NUREG-1150 probabilistic risk assessment for the surry nuclear power station. Nucl Eng Des 1992;135(1):29–59.
[121] Payne Jr AC, Breeding RJ, Helton JC, Smith LN, Johnson JD, Jow H-N, et al. The NUREG-1150 probabilistic risk assessment for the peach bottom atomic power station. Nucl Eng Des 1992;135(1): 61–94.
[122] Gregory JJ, Breeding RJ, Helton JC, Murfin WB, Higgins SJ, Shiver AW. The NUREG-1150 probabilistic risk assessment for the sequoyah nuclear plant. Nucl Eng Des 1992;135(1):92–115.
[123] Brown TD, Breeding RJ, Helton JC, Jow H-N, Higgins SJ, Shiver AW. The NUREG-1150 probabilistic risk assessment for the grand gulf nuclear station. Nucl Eng Des 1992;135(1):117–37.
[124] Helton JC, Breeding RJ. Calculation of reactor accident safety goals. Reliab Eng Syst Safety 1993;39(2):129–58.
[125] Morris MD. Three technometrics experimental design classics. Technometrics 2000;42(1):26–7.
[126] Evans M, Swartz T. Approximating integrals via Monte Carlo and deterministic methods. Oxford, NY: Oxford University Press; 2000.
[127] Hurtado JE, Barbat AH. Monte Carlo techniques in computational stochastic mechanics. Arch Comput Methods Eng 1998;5(1):3–330.
[128] Nicola VF, Shahabuddin P, Nakayama MK. Techniques for fast simulation of models of highly dependable systems. IEEE Trans Reliab 2001;50(3):246–64.
[129] Owen A, Zhou Y. Safe and effective importance sampling. J Am Stat Assoc 2000;95(449):135–43.
[130] Heidelberger P. Fast simulation of rare events in queueing and reliability models. ACM Trans Modeling Comput Simul 1995; 5(1):43–85.
[131] Shahabuddin P. Importance sampling for the simulation of highly reliable Markovian systems. Manage Sci 1994;40(3):333–52.
[132] Goyal A, Shahabuddin P, Heidelberger P, Nicola VF, Glynn PW. A unified framework for simulating Markovian models of highly dependable systems. IEEE Trans Comput 1992;41(1):36–51.
[133] Melchers RE. Search-based importance sampling. Struct Safety 1990;9(2):117–28.
[134] Glynn PW, Iglehart DL. Importance sampling for stochastic simulations. Manage Sci 1989;35(11):1367–92.
[135] Iman RL, Conover WJ. A distribution-free approach to inducing rank correlation among input variables. Commun Stat: Simul Comput B 1982;11(3):311–34.
[136] Iman RL, Davenport JM. Rank correlation plots for use with correlated input variables. Commun Stat: Simul Comput B 1982;11(3):335–60.
[137] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, SAND2001-0417. Albuquerque, NM: Sandia National Laboratories; 2002.
[138] Helton JC, Davis FJ, Johnson JD. A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. Reliab Eng Syst Safety 2005;89(3):305–30.
[139] Kleinjen JPC. An overview of the design and analysis of simulation experiments for sensitivity analysis. Europ J Operat Res 2005; 164(2):287–300.
[140] CRWMS M&O (Civilian Radioactive Waste Management System Management and Operating Contractor). Total system performance assessment (TSPA) model for site recommendation, MDL-WIS-PA-000002 REV 00. Las Vegas, Nevada: CRWMS M&O; 2000.
[141] CRWMS M&O (Civilian Radioactive Waste Management System Management and Operating Contractor). Total system performance assessment for the site recommendation, TDR-WIS-PA-000001 REV 00. Las Vegas, Nevada: CRWMS M&O; 2000.
[142] U.S. DOE (U.S. Department of Energy). Viability assessment of a repository at Yucca mountain, DOE/RW-0508. Washington, DC: US Department of Energy, Office of Civilian Radioactive Waste Management; 1998.
[143] Ibrekk H, Morgan MG. Graphical communication of uncertain quantities to nontechnical people. Risk Anal 1987;7(4):519–29.
[144] Tufte ER. The visual display of quantitative information. Cheshire, CN: Graphics Press; 1983.
[145] Berglund JW, Garner JW, Helton JC, Johnson JD, Smith LN. Direct releases to the surface and associated complementary cumulative distribution functions in the 1996 performance assessment for the waste isolation pilot plant: cuttings, cavings and spallings. Reliab Eng Syst Safety 2000;69(1–3):305–30.
[146] Cooke RM, van Noortwijk JM. Graphical methods. In: Saltelli A, Chan K, Scott EM, editors. Sensitivity analysis. New York, NY: Wiley; 2000. p. 245–64.
[147] Myers RH. Classical and modern regression with applications. 2nd ed. Boston, MA: Duxbury Press; 1989.
[148] Draper NR, Smith H. Applied regression analysis. 2nd ed. New York, NY: Wiley; 1981.
[149] Daniel C, Wood FS, Gorman JW. Fitting equations to data: computer analysis of multifactor data. 2nd ed. New York, NY: Wiley; 1980.
[150] Seber GA. Linear regression analysis. New York, NY: Wiley; 1977.

[151] Neter J, Wasserman W. Applied linear statistical models: regression, analysis of variance, and experimental designs. Homewood: Richard D. Irwin; 1974.
[152] Iman RL, Shortencarier MJ, Johnson JD. A FORTRAN 77 program and user's guide for the calculation of partial correlation and standardized regression coefficients, NUREG/CR-4122, SAND85-0044. Albuquerque: Sandia National Laboratories; 1985.
[153] Iman RL, Conover WJ. The use of the rank transform in regression. Technometrics 1979;21(4):499–509.
[154] Scheffé H. The analysis of variance. New York, NY: Wiley; 1959.
[155] Conover WJ. Practical nonparametric statistics. 2nd ed. New York, NY: Wiley; 1980.
[156] Barnard GA. Contribution to the discussion of professor Bartlett's paper. J R Stat Soc Ser B 1963;26:294.
[157] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in FORTRAN: the art of scientific computing. 2nd ed. Cambridge, NY: Cambridge University Press; 1992.
[158] Kleijnen JPC, Helton JC. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 2: robustness of techniques. Reliab Eng Syst Safety 1999;65(2):187–97.
[159] Kleijnen JPC, Helton JC. Statistical analyses of scatterplots to identify important factors in large-scale simulations, SAND98-2202. Albuquerque: Sandia National Laboratories; 1999.
[160] Wagner HM. Global sensitivity analysis. Oper Res 1995;43:948–69.
[161] Granger C, Lin J-L. Using the mutual information coefficient to identify lags in nonlinear models. J Time Ser Anal 1994;15(4): 371–84.
[162] Mishra S, Knowlton RG. Testing for input–output dependence in performance assessment models. In: Proceedings of the 10th international high-level radioactive waste management conference (IHLRWM), March 30–April 2, 2003. Las Vegas, Nevada. La Grange Park, IL: American Nuclear Society; p. 882–7.
[163] Bonnlander BV, Weigend AS. Selecting input variables using mutual information and nonparametric density estimation. In: Proceedings of the international symposium on artificial neural networks (ISANN'94), Tainan, Taiwan, 1994. p. 42–50.
[164] Moddemeijer R. On the estimation of entropy and mutual information of continuous distributions. Signal Process 1989; 16(3):233–48.
[165] Simonoff JS. Smoothing methods in statistics. New York, NY: Springer; 1996.
[166] Hastie TJ, Tibshirani RJ. Generalized additive models. London: Chapman & Hall; 1990.
[167] Chambers JM, Hastie TJ. Statistical models in S. Pacific Grove, CA: Wadsworth & Brooks; 1992.
[168] Storlie CB, Helton JC. Multiple Predictor smoothing methods for sensitivity analysis, SAND2005-???? Albuquerque, NM: Sandia National Laboratories; to appear.
[169] Ruppert D, Ward MP, Carroll RJ. Semiparametric regression. Cambridge, MA: Cambridge University Press; 2003.
[170] Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Belmont, CA: Wadsworth Intl; 1984.
[171] Mishra S, Deeds NE, RamaRao BS. Application of classification trees in the sensitivity analysis of probabilistic model results. Reliab Eng Syst Safety 2003;79(2):123–9.
[172] Hora SC, Helton JC. A distribution-free test for the relationship between model input and output when using Latin hypercube sampling. Reliab Eng Syst Safety 2003;79(3):333–9.
[173] Peacock JA. Two-dimensional goodness-of-fit testing in astronomy. Monthly Notices R Astronom Soc 1983;202(2):615–27.
[174] Fasano G, Franceschini A. A multidimensional version of the Kolmogorov–Smirnov test. Monthly Notices R Astronom Soc 1987;225(1):155–70.
[175] Garvey JE, Marschall EA, Wright RA. From star charts to stoneflies: detecting relationships in continuous bivariate data. Ecology 1998;79(2):442–7.
[176] Ripley BD. Spatial point pattern analysis in ecology. In: Legendre P, Legendre L, editors. Developments in numerical ecology. NATO ASI series, series G: ecological sciences, vol. 14. Berlin; NY: Springer; 1987. p. 407–30.
[177] Zeng G, Dubes RC. A comparison of tests for randomness. Pattern Recogn 1985;18(2):191–8.
[178] Diggle PJ. Statistical analysis of spatial point patterns. New York, NY: Academic Press; 1983.
[179] Diggle PJ, Cox TF. Some distance-based tests of independence for sparsely sampled multivariate spatial point patterns. Int Stat Rev 1983;51(1):11–23.
[180] Byth K. On robust distance-based intensity estimators. Biometrics 1982;38(1):127–35.
[181] Byth K, Ripley BD. On sampling spatial patterns by distance methods. Biometrics 1980;36(2):279–84.
[182] Diggle PJ. On parameter estimation and goodness-of-fit testing for spatial point patterns. Biometrics 1979;35(1):87–101.
[183] Diggle PJ. Statistical methods for spatial point patterns in ecology. In: Cormack RM, Ord JK, editors. Spatial and temporal analysis in ecology. Fairfield, MD: International Co-operative Publishing House; 1979. p. 95–150.
[184] Ripley BD. Tests of randomness for spatial point patterns. J R Stat Soc 1979;41(3):368–74.
[185] Besag J, Diggle PJ. Simple Monte Carlo tests for spatial pattern. Appl Stat 1977;26(3):327–33.
[186] Diggle PJ, Besag J, Gleaves JT. Statistical analysis of spatial point patterns by means of distance methods. Biometrics 1976;32:659–67.
[187] Cox TF, Lewis T. A conditioned distance ratio method for analyzing spatial patterns. Biometrika 1976;63(3):483–91.
[188] Holgate P. The use of distance methods for the analysis of spatial distribution of points. In: Lewis PAW, editor. Stochastic point processes: statistical analysis, theory, and applications. New York: Wiley-Interscience; 1972. p. 122–35.
[189] Holgate P. Tests of randomness based on distance methods. Biometrika 1965;52(3–4):345–53.
[190] Clark PJ, Evans FC. Distance to nearest neighbour as a measure of spatial relationships in populations. Ecology 1954;35:23–30.
[191] Hopkins B, Shellam JG. A new method for determining the type of distribution of plant individuals. Ann Botany 1954;18(70): 213–27.
[192] Cressie NAC. Statistics for spatial data. New York, NY: Wiley; 1993.
[193] Iman RL, Conover WJ. A measure of top-down correlation. Technometrics 1987;29(3):351–7.
[194] Campolongo F, Saltelli A, Sorensen T, Tarantola S. Hitchhiker's guide to sensitivity analysis. In: Saltelli A, Chan K, Scott M, editors. Sensitivity analysis. New York: Wiley; 2000. p. 15–47.
[195] Saltelli A, Tarantola S, Campolongo F, Ratto M. Sensitivity analysis in practice. New York, NY: Wiley; 2004.
[196] Cukier RI, Fortuin CM, Shuler KE, Petschek AG, Schiably JH. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, I. Theory. J Chem Phys 1973;59(8):3873–8.
[197] Schaibly JH, Shuler KE. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, II. Applications. J Chem Phys 1973;59(8):3879–88.
[198] McRae GJ, Tilden JW, Seinfeld JH. Global sensitivity analysis—a computational implementation of the Fourier amplitude sensitivity test (FAST). Comput Chem Eng 1981;6(1):15–25.
[199] Saltelli A, Sobol' IM. About the use of rank transformation in sensitivity analysis of model output. Reliab Eng Syst Safety 1995;50(3):225–39.
[200] Homma T, Saltelli A. Importance measures in global sensitivity analysis of nonlinear models. Reliab Eng Syst Safety 1996; 52(1):1–17.
[201] Archer GEB, Saltelli A, Sobol' IM. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. J Stat Comput Simul 1997;58(2):99–120.
[202] Saltelli A, Bolado R. An alternative way to compute Fourier amplitude sensitivity test (FAST). Comput Stat Data Anal 1998; 26(4):267–79.

[203] Saltelli A, Tarantola S, Campolongo F. Sensitivity analysis as an ingredient of modeling. Stat Sci 2000;15(4):377–95.
[204] Chan K, Saltelli A, Tarantola S. Winding stairs: a sampling tool to compute sensitivity indices. Stat Comput 2000;10(3):187–96.
[205] Rabitz H, Alis OF, Shorter J, Shim K. Efficient input–output model representations. Comput Phys Commun 1999;117(1–2):11–20.
[206] Jansen MJW, Rossing WAH, Daamen RA. Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. In: Grasman J, van Straten G, editors. Predictability and nonlinear modeling in natural sciences and economics. Boston: Kluwer Academic Publishers; 1994. p. 334–43.
[207] Jansen MJW. Analysis of variance designs for model output. Comput Phys Commun 1999;117(1–2):35–43.
[208] McKay MD. Evaluating prediction uncertainty, LA-12915-MS; NUREG/CR-6311. Los Alamos, NM: Los Alamos National Laboratory; 1995.
[209] McKay MD. Nonparametric variance-based methods of assessing uncertainty importance. Reliab Eng Syst Safety 1997;57(3):267–79.
[210] McKay MD, Morrison JD, Upton SC. Evaluating prediction uncertainty in simulation models. Comput Phys Commun 1999; 117(1–2):44–51.
[211] Helton JC. Mathematical and numerical approaches in performance assessment for radioactive waste disposal: dealing with uncertainty. In: Scott EM, editor. Modelling radioactivity in the environment. New York, NY: Elsevier Science; 2003. p. 353–90.
[212] Helton JC, Anderson DR, Basabilvazo G, Jow H-N, Marietta MG. Conceptual structure of the 1996 performance assessment for the waste isolation pilot plant. Reliab Eng Syst Safety 2000;69(1–3): 151–65.
[213] Diaz NJ. Realism and conservatism, remarks by chairman diaz at the 2003 nuclear safety research conference. NRC News. No. S-03-023, Washington, DC: US Nuclear Regulatory Commission; October 20, 2003.
[214] Paté-Cornell E. Risk and uncertainty analysis in government safety decisions. Risk Anal 2002;22(3):633–46.
[215] Caruso MA, Cheok MC, Cunningham MA, Holahan GM, King TL, Parry GW, et al. An approach for using risk assessment in risk-informed decisions on plant-specific changes to the licensing basis. Reliab Eng Syst Safety 1999;63:231–42.
[216] Sielken Jr RL, Bretzlaff RS, Stevenson DE. Challenges to default assumptions stimulate comprehensive realism as a new tier in quantitative cancer risk assessment. Regul Toxicol Pharmacol 1995;21:270–80.
[217] Nichols AL, Zeckhauser RJ. The perils of prudence: how conservative risk assessments distort regulation. Regul Toxicol Pharmacol 1988;8:61–75.