SANDIA REPORT

SAND99–2240
Unlimited Release
Printed July 2000

Sampling-Based Methods for Uncertainty and Sensitivity Analysis

J. C. Helton and F. J. Davis

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

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

Sandia National Laboratories

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

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

Prices available from (703) 605-6000
Web site: http://www.ntis.gov/ordering.htm

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

NTIS price codes
Printed copy: A06
Microfiche copy: A01

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

SAND99-2240
Unlimited Release
Printed July 2000

Sampling-Based Methods for Uncertainty and Sensitivity Analysis

J.C. Helton and F.J. Davis
P.O. Box 5800
Sandia National Laboratories
Albuquerque, New Mexico 87185-0779

Abstract

Sampling-based methods for uncertainty and sensitivity analysis are reviewed. Topics considered include (i) separation of stochastic (i.e., aleatory) and subjective (i.e., epistemic) uncertainty, (ii) construction of distributions to characterize subjective uncertainty, (iii) sampling procedures (i.e., random sampling, importance sampling, Latin hypercube sampling), (iv) propagation of uncertainty through models, (v) display of uncertainty in model predictions, and (vi) sensitivity analysis procedures (i.e., examination of scatterplots, regression analysis, stepwise regression analysis, correlation and partial correlation, rank transformations, identification of nonmonotonic and nonrandom patterns). Procedures are illustrated with (i) a model for two-phase fluid flow, (ii) a sequence of simple test functions, and (iii) a performance assessment for a radioactive waste disposal facility.

Contents

Figures

Tables

1.0 Introduction

Sampling-based methods for uncertainty and sensitivity analysis involve the generation and exploration of a mapping from uncertain analysis inputs to analysis results (Iman et al. 1981a, b, Iman 1992, Helton 1993b). Conceptually, the analysis or model under consideration can be represented by a vector function

$$\mathbf{y} = [y_1, y_2, \ldots, y_{nY}], \tag{1.1}$$

and the associated input can be represented by a vector

$$\mathbf{x} = [x_1, x_2, \ldots, x_{nX}], \tag{1.2}$$

where nX and nY are the dimensions of $\mathbf{x}$ and $\mathbf{y}$, respectively, and each value of $\mathbf{x}$ produces a corresponding value $\mathbf{y}(\mathbf{x})$. Most real analyses are quite complex, with the result that the dimensions of $\mathbf{x}$ and $\mathbf{y}$ can be large.

If the value for $\mathbf{x}$ was unambiguously known, then $\mathbf{y}(\mathbf{x})$ could be determined and presented as the unique outcome of the analysis. However, there is uncertainty with respect to the appropriate value to use for $\mathbf{x}$ in most analyses, with the result that there is also uncertainty with respect to the value of $\mathbf{y}(\mathbf{x})$. The uncertainty in $\mathbf{x}$ and its associated effect on $\mathbf{y}(\mathbf{x})$ lead to two closely related questions: (i) "What is the uncertainty in $\mathbf{y}(\mathbf{x})$ given the uncertainty in $\mathbf{x}$?", and (ii) "How important are the individual elements of $\mathbf{x}$ with respect to the uncertainty in $\mathbf{y}(\mathbf{x})$?" Attempts to answer these two questions are typically referred to as uncertainty analysis and sensitivity analysis, respectively.

An assessment of the uncertainty in $\mathbf{y}$ derives from a corresponding assessment of the uncertainty in $\mathbf{x}$. In particular, $\mathbf{y}$ is assumed to have been developed so that appropriate analysis results are obtained if the appropriate value for $\mathbf{x}$ is used in the evaluation of $\mathbf{y}$. Unfortunately, it is impossible to unambiguously specify the appropriate value of $\mathbf{x}$ in most analyses; rather, there are many possible values for $\mathbf{x}$ of varying levels of plausibility. Such uncertainty is often given the designation subjective or epistemic and is characterized by assigning a distribution

$$D_1, D_2, \ldots, D_{nX} \tag{1.3}$$

to each element x_j of $\mathbf{x}$. Correlations and other restrictions involving the x_j are also possible. These distributions and any associated conditions characterize a degree of belief as to where the appropriate value of each variable x_j is located for use in evaluation of $\mathbf{y}$ and in turn lead to distributions for the individual elements of $\mathbf{y}$. Given that the distributions in Eq. (1.3) characterize a degree of belief with respect to where the appropriate input to use in the analysis is located, the resultant distributions for the elements of $\mathbf{y}$ characterize a corresponding degree of belief with respect to where the appropriate values of the outcomes of the analysis are located.

Sampling-based methods for uncertainty and sensitivity analysis are based on a sample

$$\mathbf{x}_k = [x_{k1}, x_{k2}, \ldots, x_{k,nX}], \quad k = 1, 2, \ldots, nS, \tag{1.4}$$

of size nS from the possible values for $\mathbf{x}$ as characterized by the distributions in Eq. (1.3) and on the corresponding evaluations

$$\mathbf{y}(\mathbf{x}_k) = [y_1(\mathbf{x}_k), y_2(\mathbf{x}_k), \ldots, y_{nY}(\mathbf{x}_k)], \quad k = 1, 2, \ldots, nS, \tag{1.5}$$

of $\mathbf{y}$. The pairs

$$[\mathbf{x}_k, \mathbf{y}(\mathbf{x}_k)], \quad k = 1, 2, \ldots, nS, \tag{1.6}$$

form a mapping from the uncertain analysis inputs (i.e., the $\mathbf{x}_k$'s) to the corresponding uncertain analysis results (i.e., the $\mathbf{y}(\mathbf{x}_k)$'s). When an appropriate probabilistic procedure has been used to generate the sample in Eq. (1.4) from the distributions in Eq. (1.3), the resultant distributions for the elements of $\mathbf{y}$ characterize the uncertainty in the results of the analysis (i.e., constitute the outcomes of an uncertainty analysis). Further, examination of scatterplots, regression analysis, partial correlation analysis and other procedures for investigating the mapping in Eq. (1.6) provide a way to determine the effects of the elements of $\mathbf{x}$ on the elements of $\mathbf{y}$ (i.e., constitute procedures for sensitivity analysis).

When viewed at a high level, performance of a sampling-based uncertainty and sensitivity analysis involves five components: (i) definition of the distributions in Eq. (1.3) that characterize uncertainty, (ii) generation of the sample in Eq. (1.4) from the distributions in Eq. (1.3), (iii) evaluation of $\mathbf{y}$ for the individual elements of the sample in Eq. (1.4) to produce the model evaluations in Eq. (1.5), (iv) generation of displays of the uncertainty in $\mathbf{y}$ from the analysis outcomes in Eq. (1.5), and (v) exploration of the mapping in Eq. (1.6) to

determine the effects of the elements of **x** on the elements of **y**. The preceding components of a sampling-based uncertainty and sensitivity analysis are discussed (Sects. 4 - 8). Further, the classification of uncertainty and the potential effects that this classification has on sampling-based analyses are discussed (Sect. 2), and an example problem is introduced for use in illustrating the ideas and techniques under consideration (Sect. 3). In addition, the ideas and techniques described in this presentation are illustrated with a sequence of relatively simple test problems (Sect. 9) and also with an analysis involving both stochastic (i.e., aleatory) and subjective (i.e., epistemic) uncertainty (Sect. 10). Finally, the presentation ends with a summary discussion (Sect. 11).

2.0 Classification of Uncertainty

The need for an appropriate treatment of uncertainty in complex analyses is recognized by most, if not all, analysts (e.g., Cullen and Frey 1999, Risk Assessment Forum 1997, Thompson and Graham 1996, Helton and Burmaster 1996, Paté-Cornell 1996, Hoffman and Hammonds 1994, Apostolakis 1990). Yet, the treatment of uncertainty in large analyses often causes confusion because uncertainty and its probabilistic characterization can arise from two distinct sources.

First, there is the uncertainty that arises because the system under study can behave in many different ways. For example, the number of possible sequences of weather conditions that could occur at a fixed location over some specified time interval in the future is quite large; similarly, the number of potential sequences of operating conditions that could occur at an industrial facility over the course of one year is also quite large. This type of uncertainty is often referred to as stochastic or aleatory uncertainty and is a property of the system under consideration (Helton 1994, 1997).

Second, there is the uncertainty that arises from an inability to specify the exact value of a quantity that is assumed to have a constant value within a particular analysis. For example, a system component might be assumed to have a uniquely determined failure strength, with the exact value of this failure strength being imprecisely known. As another example, some process might be assumed to occur at a particular rate, with the exact value of this rate being imprecisely known. This type of uncertainty is often referred to as subjective or epistemic uncertainty and is a property of the analysts carrying out the study (Helton 1994, 1997).

Probability is typically used to characterize both stochastic and subjective uncertainty. This dual usage of probability has the potential to result in considerable confusion when care is not taken to specify which interpretation of uncertainty is intended. In this presentation, the assumption is made that the goal of uncertainty and sensitivity analysis is to investigate the effects of subjective uncertainty. Thus, the distributions in Eq. (1.3) are characterizing subjective uncertainty.

As described in the next section, an example that derives from a performance assessment (PA) for the Waste Isolation Pilot Plant (WIPP) will be used to illustrate sampling-based methods for uncertainty and sensitivity analysis (U.S. DOE 1996, Helton et al. 1998a). This example uses a model for two-phase (i.e., gas and brine) fluid flow and involves only subjective uncertainty. Then, in Sect. 10, a second example will be introduced that considers a complementary cumulative distribution function (CCDF) specified in the U.S. Environmental Protection Agency's (EPA's) standard for the geologic disposal of radioactive waste (U.S. EPA 1985, 1993, 1996) and involves both stochastic and subjective uncertainty.

As already indicated, it is important to maintain a clear distinction between the use of probability to characterize stochastic uncertainty and the use of probability to characterize subjective uncertainty. The concept of a probability space provides a convenient way to maintain this distinction. A probability space $(\mathcal{S}, \mathcal{A}, p)$ is the formal structure on which the mathematical development of probability is based and consists of three components: (i) a set $\mathcal{S}$ that contains everything that could occur in the particular universe under consideration, (ii) a suitably restricted collection $\mathcal{A}$ of subsets of $\mathcal{S}$ for which probabilities are defined, and (iii) a function p that defines the probabilities of the elements of $\mathcal{A}$ (p. 116, Feller 1971). Thus, an analysis that involves both stochastic and subjective uncertainty has two probability spaces associated with it: a probability space $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ for stochastic uncertainty and probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ for subjective uncertainty, where the subscripts *st* and *su* designate stochastic and subjective, respectively. The distributions in Eq. (1.3) and any associated restrictions are defining a probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ for subjective uncertainty.

Accessible discussions on the origins of the use of probability to characterize subjective and stochastic uncertainty are given by Hacking 1975 and Bernstein 1996.

3.0 Example Analysis Problem

Analysis procedures are easier to understand and assess when they are illustrated by real examples. For this reason, a nontrivial example from a PA carried out in support of the 1996 Compliance Certification Application (CCA) for the WIPP will be used to illustrate the procedures under consideration (U.S. DOE 1996, Helton et al. 1998a). The WIPP is under development near Carlsbad, NM, by the U.S. Department of Energy (DOE) for the geologic (i.e., deep underground) disposal of transuranic (TRU) waste (Rechard 1999, NRC 1996). Waste disposal will take place in excavated chambers located in a bedded salt formation (Fig. 1, Helton et al. 1998b).

A number of mathematical models are involved in assessing the potential behavior of the WIPP, its surrounding environment, and the radionuclides emplaced there (see Sect. 10.1 for a summary of these models). Most of these models involve the numerical solution of systems of partial differential equations used to represent material deformation, fluid flow and radionuclide transport. The model used to represent two-phase (i.e., gas and brine) fluid flow in the vicinity of the repository will be used for illustration, with this model implemented by the BRAGFLO program.

The model for two-phase fluid flow is based on the following system of nonlinear partial differential equations:

Gas Conservation

$$\nabla \bullet \left[\frac{\alpha \rho_g \mathbf{K}_g k_{rg}}{\mu_g} \left(\nabla p_g + \rho_g g \nabla h \right) \right] + \alpha q_{wg} + \alpha q_{rg} = \alpha \frac{\partial \left(\phi \rho_g S_g \right)}{\partial t} \tag{3.1}$$

Brine Conservation

$$\nabla \bullet \left[\frac{\alpha \rho_b \mathbf{K}_b k_{rb}}{\mu_b} \left(\nabla p_b + \rho_b g \nabla h \right) \right] + \alpha q_{wb} + \alpha q_{rb} = \alpha \frac{\partial \left(\phi \rho_b S_b \right)}{\partial t} \tag{3.2}$$

Saturation Constraint

$$S_g + S_b = 1 \tag{3.3}$$

Capillary Pressure Constraint

$$p_C = p_g - p_b = f(S_b) \tag{3.4}$$

Gas Density

ρ_g determined by Redlich-Kwong-Soave equation of state (see Eqs. (4.2.27), (4.2.28), Helton et al. 1998a)

Brine Density

$$\rho_b = \rho_0 \exp\left[\beta_b (p_b - p_{b0})\right] \tag{3.5}$$

Formation Porosity

$$\phi = \phi_0 \exp\left[\beta_f (p_b - p_{b0})\right] \tag{3.6}$$

where

- g = acceleration of gravity (m/s^2)
- h = vertical distance from a reference location (m)
- $\mathbf{K}_l$ = permeability tensor (m^2) for fluid l ($l = g$ ~ gas, $l = b$ ~ brine)
- k_{rl} = relative permeability (dimensionless) to fluid l
- p_C = capillary pressure (Pa)
- p_l = pressure of fluid l (Pa)
- q_{rl} = rate of production (or consumption, if negative) of fluid l due to chemical reaction (kg/m^3/s)
- q_{wl} = rate of injection (or removal, if negative) of fluid l (kg/m^3/s)
- S_l = saturation of fluid l (dimensionless)
- t = time (s)
- α = geometry factor (m in 1996 WIPP PA)
- ρ_l = density of fluid l (kg/m^3)
- μ_l = viscosity of fluid l (Pa s)
- ϕ = porosity (dimensionless)
- ϕ_0 = reference (i.e., initial) porosity (dimensionless)
- p_{b0} = reference (i.e., initial) brine pressure (Pa), constant in Eq. (3.5) and spatially variable in Eq. (3.6)
- ρ_0 = reference (i.e., initial) brine density (kg/m^3)
- β_f = pore compressibility (Pa^{-1})
- β_b = brine compressibility (Pa^{-1})

and f is defined by the model for capillary pressure in use (see the right hand sides of Eqs. (4.2.9), (4.2.15)

and (4.2.18) in Helton et al. 1998a). The conservation equations are valid in one (i.e., $\nabla = [\partial/\partial x]$), two (i.e., $\nabla = [\partial/\partial x \ \partial/\partial y]$) and three (i.e., $\nabla = [\partial/\partial x \ \partial/\partial y \ \partial/\partial z]$) dimensions. In the present example, the preceding system of equations is used to model two-phase fluid flow in a two-dimensional region (Fig. 3.1).

In general, the individual terms in Eqs. (3.1) - (3.6) are functions of location and time (e.g., $p_g(x, y, t)$, $\rho_g(x, y, t)$, $k_{rg}(x, y, t)$, ...) and often other variables as well (i.e., elements of the vector $\mathbf{x}$ in Eq. (1.2)). A full description of how the individual terms in these equations are defined is beyond the scope of this presentation and is available elsewhere (Bean et al. 1996; Sect. 4.2, Helton et al. 1998a). The system of partial differential equations in Eqs. (3.1) - (3.6) is too complex to permit a closed form solution. In the present analysis, these equations were solved with finite difference procedures implemented by the BRAGFLO program on the computational grid in Fig. 3.1 (WIPP PA 1996, Bean et al. 1996).

Two analysis problems involving the solution of Eqs. (3.1) - (3.6) will be considered. The first problem involves undisturbed conditions (i.e., E0 conditions in the terminology of the 1996 WIPP PA). In this problem, the behavior of the repository is modeled under the assumption that it experiences no human disruptions after its final decommissioning and closure. The second problem involves a drilling intrusion that occurs 1000 yr after the closure of the repository, passes through a waste panel, and does not penetrate an area of pressurized brine (i.e., a brine pocket) in the Castile Formation (Fm) beneath the repository (i.e., an E2 intrusion in the terminology of the 1996 WIPP PA). The differences between the two problems are implemented through the specification of the properties of the regions labeled 1A, 1B and 1C in Fig. 3.1. Due to regulatory requirements (U.S. EPA 1985, 1993, 1996), the modeled period extends from slightly before closure of the repository ($t = -5$ yr), through closure of the repository ($t = 0$ yr), and out to $t = 10{,}000$ yr. The 1996 WIPP PA also considered drilling intrusions that passed through the repository and penetrated pressurized brine in the Castile Fm (i.e., E1 intrusions) but these calculations are not used for illustration in this presentation.

A number of factors contribute to the presence of subjective uncertainty in the formulation and solution of the model embodied in Eqs. (3.1) - (3.6): (i) a geologic system that can never be fully observed and characterized is under consideration, (ii) the waste to be emplaced at the WIPP is not fully characterized, (iii) the mechanical and chemical evolution of the waste panels cannot be predicted with certainty, (iv) many of the inputs to the analysis are spatially and possibly temporally averaged values for quantities (e.g., permeabilities) that vary in space and possibly in time, and (v) a very long time period (i.e., 10,000 yr) is under consideration. For these and other reasons, considerable uncertainty exists with respect to the appropriate values to use for many of the quantities that enter into the formulation of the model in Eqs. (3.1) - (3.6).

To assess the effects of such uncertainty, the 1996 WIPP PA identified 31 uncertain inputs to the BRAGFLO program required in the formulation of the model in Eqs. (3.1) - (3.6) (Table 3.1). The exact manner in which these inputs were used in the definition of the coefficients in Eqs. (3.1) - (3.6) is described in Table 5.2.1 of Helton et al. 1998a.

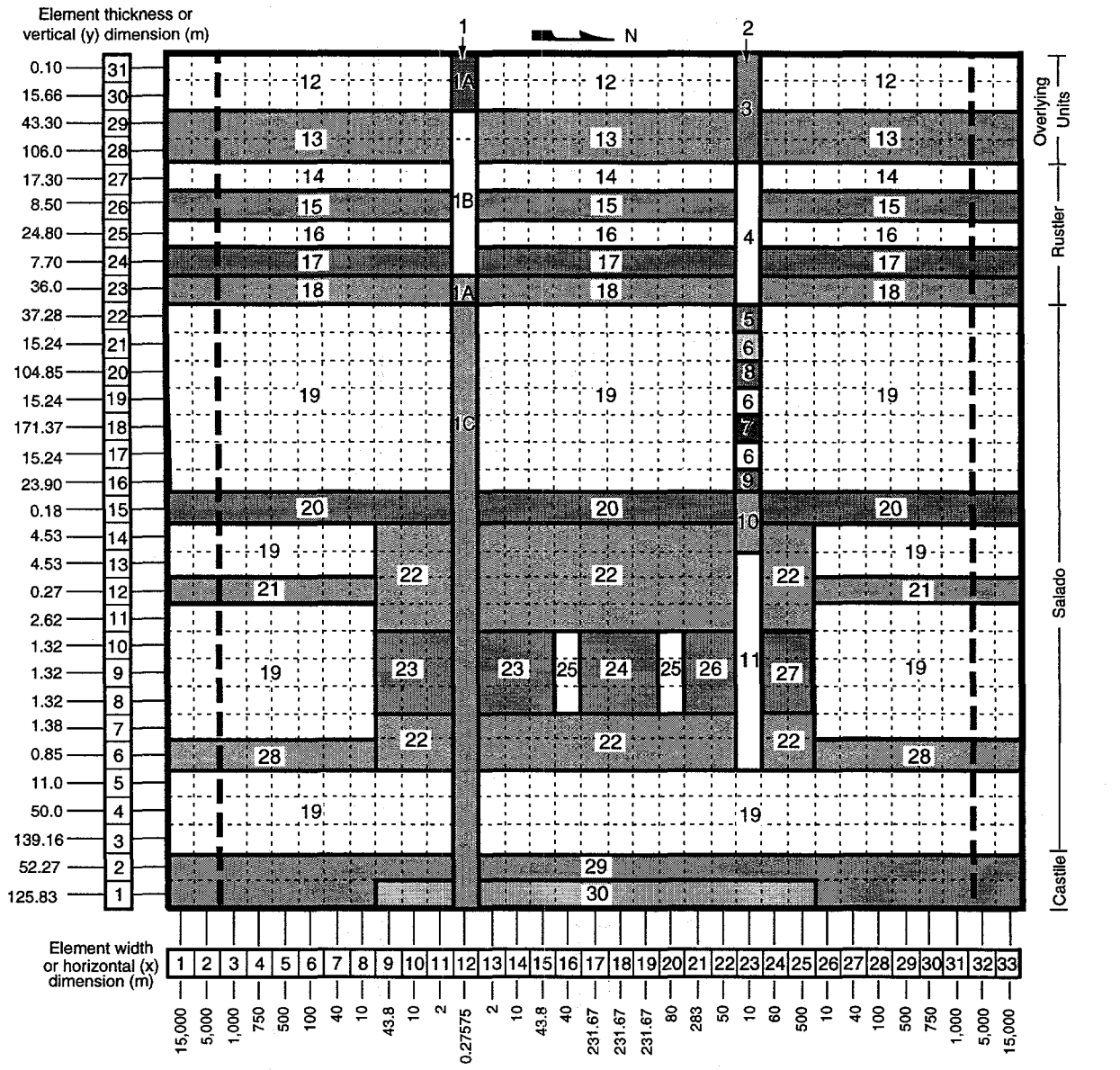
The analyses under consideration were structured to require a single value for each of the variables in Table 3.1. However, the exact values to use for these variables were felt to be poorly known. Therefore, ranges of possible values for these variables were developed, and distributions were assigned to these ranges to characterize a degree of belief with respect to the location of the appropriate values to use in the 1996 WIPP PA. Thus, the distributions indicated in Table 3.1 are characterizing subjective uncertainty.

Put another way, the distributions and associated correlations in Table 3.1 are defining a probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ for subjective uncertainty. In this space, the elements $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$ are vectors of the form

$$\mathbf{x}_{su} = [ANHBCEXP, ANHBCVGP, \ldots, WRGSSAT] \quad (3.7)$$

and correspond to the vector $\mathbf{x}$ in Eq. (1.2). Similarly, $\mathbf{y}$ in Eq. (1.1) corresponds to the totality of the results generated in the solution of Eqs. (3.1) - (3.6). Actually, there are two $\mathbf{y}$ values in this example: one for solution of the equations for undisturbed (i.e., E0) conditions and one for solution of the equations for disturbed conditions (i.e., an E2 intrusion at 1000 yr).

All 31 variables in Table 3.1 are used in the formulation of Eqs. (3.1) - (3.6) for the E2 intrusion. However, *BHPRM* relates only to the E1 and E2 intrusions and so was not used in the formulation of Eqs. (3.1) - (3.6) for E0 conditions. Further, the variables affecting the brine pocket (i.e., *BPCOMP*, *BPINTPRS*, *BPPRM*, *BPVOL*) are effectively removed from the calculation of any results associated with the repository for E0 and E2 conditions due to the absence of a connection between the brine pocket and the repository (Fig. 3.1).



1. Borehole (first 200 years)
   1A. Borehole concrete plug
   1B. Upper unrestricted borehole
   1C. Lower unrestricted borehole
2. Shaft
3. Earth fill
4. Rustler compacted clay column
5. Asphalt
6. Concrete
7. Crushed salt (compacted salt column)
   Boundary of accessible environment
8. Upper Salado compacted clay column
9. Lower Salado compacted clay column
10. Lower clay component
11. Concrete monolith
12. Units above the Dewey Lake
13. Dewey Lake
14. Forty-niner
15. Magenta
16. Tamarisk
17. Culebra
18. Unnamed lower member
19. Impure halite
20. MB138
21. Anhydrite layer a and b
22. Disturbed rock zone
23. Waste panel
24. Rest of repository
25. Panel closures
26. Operation region
27. Experimental area
28. MB139
29. Castile
30. Brine reservoir

TRI-6342-5288-0

Fig. 3.1. Computational grid used in BRAGFLO to represent two-phase flow in 1996 WIPP PA subsequent to a drilling intrusion. Same formulation is used in the absence of a drilling intrusion except that regions 1A, 1B and 1C have the same properties as the regions to either side.

Table 3.1. Uncertain Variables Used as Input to BRAGFLO in the 1996 WIPP PA (see Table 5.2.1, Helton et al. 1998a and App. PAR, U.S. DOE 1996, for additional information)

ANHBCEXP—Brooks-Corey pore distribution parameter for anhydrite (dimensionless). Distribution: Student's with 5 degrees of freedom. Range: 0.491 to 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} to 2.75×10^{-10} Pa^{-1}. Mean, Median: 8.26×10^{-11} Pa^{-1}. Correlation: –0.99 rank correlation (Iman and Conover 1982) with *ANHPRM*.

ANHPRM—Logarithm of anhydrite permeability (m^2). Distribution: Student's with 5 degrees of freedom. Range: –21.0 to –17.1 (i.e., permeability range is 1×10^{-21} to $1 \times 10^{-17.1}$ 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} to 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} to 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} to 1×10^{-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 $1 \times 10^{-11.3}$ to 1×10^{-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 to 1.70×10^7 Pa. Mean, mode: 1.36×10^7 Pa, 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 $1 \times 10^{-14.7}$ to $1 \times 10^{-9.80}$ m^2). Mean, mode: –12.1, –11.8. Correlation: –0.75 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} to 1.92×10^{-10} PA^{-1}. Mean, median: 9.75×10^{-11} Pa^{-1}, 9.75×10^{-11} Pa^{-1}. Correlation: –0.99 rank correlation with *HALPRM*.

HALPOR—Halite porosity (dimensionless). Distribution: Piecewise uniform. Range: 1.0×10^{-3} to 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 1×10^{-21} m^2). Mean, median: –22.5, –22.5. Correlation: –0.99 rank correlation with *HALCOMP*.

Table 3.1. Uncertain Variables Used as Input to BRAGFLO in the 1996 WIPP PA (see Table 5.2.1, Helton et al. 1998a and App. PAR, U.S. DOE 1996, for additional information) (continued)

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 Marker Bed (MB) 139 (Pa). Distribution: Uniform. Range: 1.104×10^7 to 1.389×10^7 Pa. Mean, median: 1.247×10^7 Pa, 1.247×10^7 Pa.

SHBCEXP—Brooks-Corey pore distribution parameter for shaft (dimensionless). Distribution: Piecewise uniform. Range: 0.11 to 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 to −18 (i.e., permeability range is 1×10^{-21} to 1×10^{-18} m^2). Mean, mode: −19.7, −20.0.

SHPRMCLY—Logarithm of permeability (m^2) for clay components of shaft. Distribution: Triangular. Range: −21 to −17.3 (i.e., permeability range is 1×10^{-21} to $1 \times 10^{-17.3}$ m^2). Mean, mode: −18.9, −18.3.

SHPRMCON—Same as *SHPRMASP* but for concrete component of shaft seal for 0 to 400 yr. Distribution: Triangular. Range: −17.0 to −14.0 (i.e., permeability range is 1×10^{-17} to 1×10^{-14} m^2). Mean, mode: −15.3, −15.0.

SHPRMDRZ—Logarithm of permeability (m^2) of DRZ surrounding shaft. Distribution: Triangular. Range: −17.0 to −14.0 (i.e., permeability range is 1×10^{-17} to 1×10^{-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 to 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, 10000 yr]. *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 to 0.4. Mean, median: 0.2, 0.2.

SHRGSSAT—Residual gas saturation in shaft (dimensionless). Distribution: Uniform. Range: 0 to 0.4. Mean, median: 0.2, 0.2.

WASTWICK—Increase in brine saturation of waste due to capillary forces (dimensionless). Distribution: Uniform. Range: 0 to 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 to 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 to 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 to 1.27×10^{-9} mol/kg•s. Mean, median: 6.34×10^{-10} mol/kg•s, 6.34×10^{-10} mol/kg•s.

Table 3.1. Uncertain Variables Used as Input to BRAGFLO in the 1996 WIPP PA (see Table 5.2.1, Helton et al. 1998a and App. PAR, U.S. DOE 1996, for additional information) (continued)

WGRMICI—Microbial degradation rate for cellulose under inundated conditions (mol/kg•s). Distribution: Uniform. Range: 3.17×10^{-10} to 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 to 0.552. Mean, median: 0.276, 0.276.

WRGSSAT—Residual gas saturation in waste (dimensionless). Distribution: Uniform. Range: 0 to 0.15. Mean, median: 0.075, 0.075.

4.0 Definition of Distributions for Subjective Uncertainty

The definition of the distributions in Eq. (1.3) used to characterize subjective uncertainty is, in many ways, the most important single part of a sampling-based uncertainty and sensitivity analysis because these distributions determine both the uncertainty in **y** and the relative importance of the individual elements of **x** that give rise to this uncertainty. However, the determination of such distributions is not the primary focus of this presentation and thus will be treated rather briefly.

It is important for everyone involved in the definition of these distributions to understand the type of information that is being quantified. In particular, the purpose of these distributions is to characterize a degree of belief with respect to where the appropriate value of each element of $\mathbf{x}_{su}$ is located for use in the analysis. In concept, the analysis structure has been developed to the point that a single value for each element of $\mathbf{x}_{su}$ is required, but the precise values for these elements, and hence for $\mathbf{x}_{su}$, are not known.

A common error is to define the D_j so that they characterize spatial, temporal or experimental variability. If the analysis uses a quantity that is held constant over an extended period of time or over an extended area, then the corresponding distribution D_j should not be defined to characterize temporal or spatial variability. Rather, given that the model uses a spatially or temporally averaged input, the distribution D_j should characterize the uncertainty in this averaged quantity rather than the variability that is averaged over. Similarly, experimental variability is not the same as the uncertainty in an analysis input derived from variable experimental outcomes.

Due to its importance and pervasiveness, the characterization of subjective uncertainty has been widely studied (e.g., Berger 1985; Cook and Unwin 1986; Mosleh et al. 1988; Hora and Iman 1989; Keeney and von Winterfeldt 1991; Bonano et al. 1990; Bonano and Apostolakis 1991; Cooke 1991; Meyer and Booker 1991; Ortiz et al. 1991; NRC 1992; Thorne 1993). Perhaps the largest example of an analysis to use a formal expert review process to assess the uncertainty in its inputs is the U.S. Nuclear Regulatory Commission's reassessment of the risks from commercial nuclear power stations (U.S. NRC 1990-1991; Harper et al. 1990, 1991, 1992; Breeding et al. 1992). Another large example is an assessment of seismic risks in the eastern United States (EPRI 1989).

Although formal statistical procedures might be useful in the construction of the distributions D_j, $j = 1, 2, \ldots, nX$, in Eq. (1.3) in some situations, in most cases such distributions are probably best developed by specifying selected quantile values without making an attempt to specify a particular distribution type and its associated parameters (e.g., normal, log normal, beta, ...) (Sect. 3.1, Helton 1993b). For example, the construction procedure might start by specifying minimum, median and maximum values for the variable under consideration (i.e., the points $(x_{0.00}, 0.00)$, $(x_{0.5}, 0.5)$ and $(x_{1.00}, 1.00)$ on the cumulative distribution function (CDF) in Fig. 4.1). Then, resolution could be added by specifying additional quantile values (e.g., the points $(x_{0.10}, 0.10)$, $(x_{0.25}, 0.25)$, $(x_{0.75}, 0.75)$ and $(x_{0.90}, 0.90)$ in Fig. 4.1). The process can be continued until it is felt that the distribution is providing an adequate characterization of the uncertainty in the variable under consideration. Hopefully, the expert, or experts, whose knowledge is being quantified by this distribution should be able to provide a documentable rationale for the selection of specific quantile values. The expert is more likely to be able to justify the selection of specific quantile values than the choice of specific parameters to define a beta distribution or some other formal distribution.

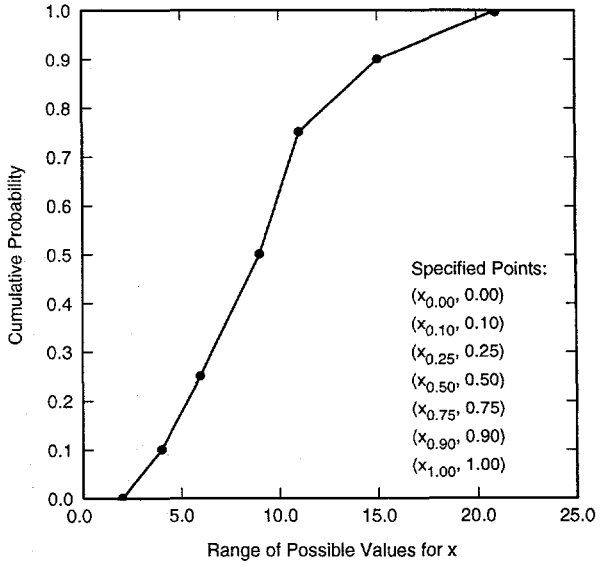


Fig. 4.1. Construction of CDF from specified quantile values.

When several experts are used to develop a distribution for a variable, one possibility is first to have each expert independently develop a distribution as indicated in Fig. 4.1. Then, these distributions can be vertically averaged to produce a new distribution based on the distributions supplied by the individual experts (Fig. 4.2). This is easiest to do if each expert's distribution is assigned equal weight (i.e., the divisor in the averaging process is nE, where nE is the number of experts). In practice, the assigning of different weights to different experts is very difficult.

As previously indicated, this presentation uses an example from the 1996 WIPP PA. The variables that comprise the elements of $\mathbf{x}_{su}$ in this example are listed in Table 3.1. The distributions assigned to these variables were defined by appropriate members of the experimental programs that were being carried out at Sandia National Laboratories to support the development of the WIPP, with these distributions intended to characterize a degree of belief with respect to where the appropriate values of these variables are located for use in the 1996 WIPP PA. The distributions assigned to *WMICDFLG* and *WSOLAM3C* are illustrated in Fig. 4.3, with *WMICDFLG* having a discrete distribution and *WSOLAM3C* having a piecewise uniform distribution (i.e., the type of distribution that results when quantiles are defined as indicated in Fig. 4.1 and then connected by straight lines).

The care and effort used in the definition of the distributions in Eq. (1.3) are dependent on both the purpose of an analysis and the amount of time and resources available for its implementation. If the analysis is primarily exploratory in nature or if limited time and resources are available, then rather crude specifications for these distributions might be used (e.g., uniform and loguniform for variables with uncertainty ranges less than and greater than one order of magnitude, respectively). As long as the ranges are not unreasonably small or large, such an approach can lead to considerable insights into the behavior of a system and the variables that influence this behavior. However, more robust insights would require greater effort in the definition of the distributions. An efficient approach is to carry out an initial screening analysis with uniform and loguniform distributions to identify the most important variables and then to characterize more carefully the uncertainty in these variables for use in a second analysis. This iterative approach allows resources to be concentrated on characterizing the uncertainty in the most important variables. If a variable has little effect on the outcome of an analysis, then the accuracy with which its uncertainty is characterized is not very important to the outcome of the analysis.

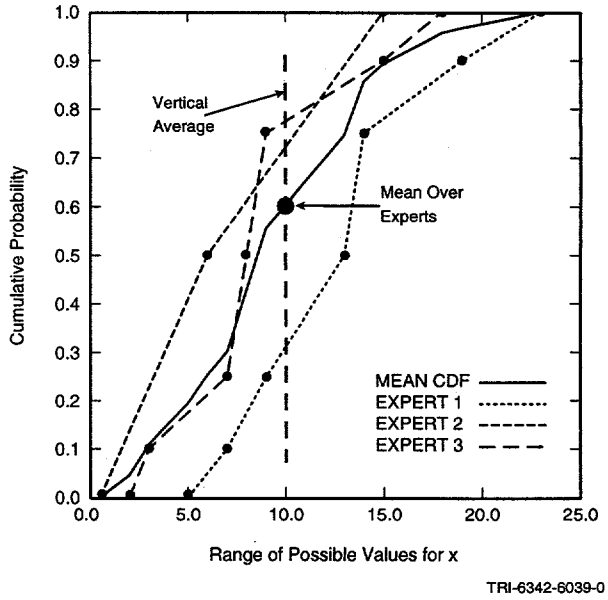


Fig. 4.2. Construction of mean CDF by averaging of CDFs defined by individual experts, with equal weight (i.e., $1/nE = 1/3$, where $nE = 3$ is the number of experts) given to each expert.

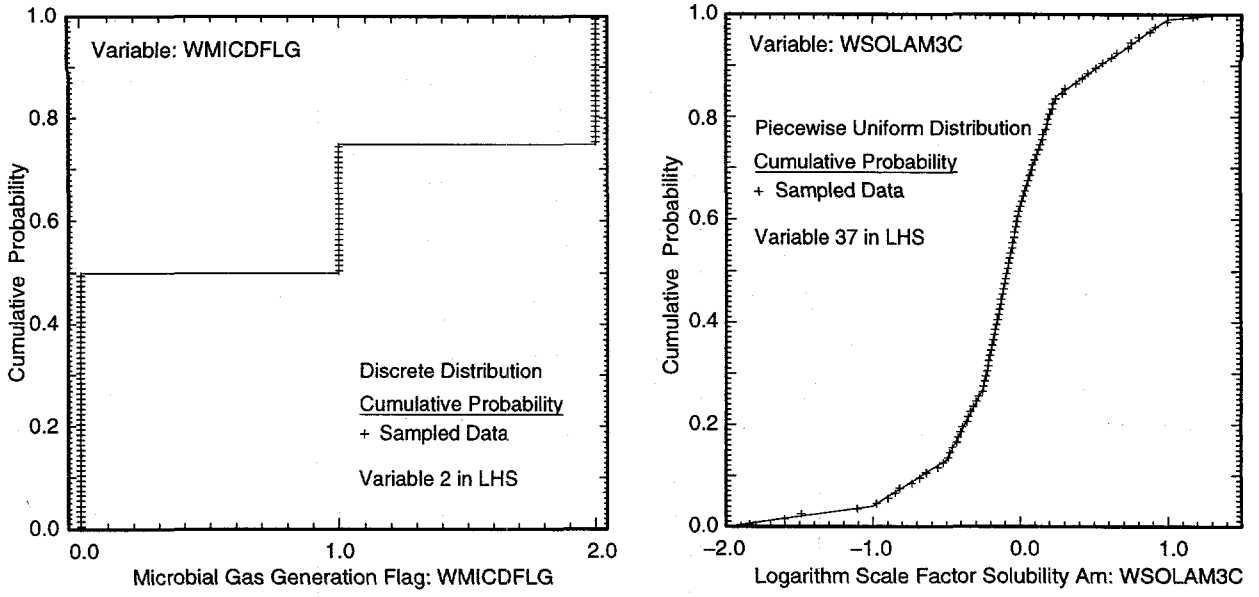


Fig. 4.3. Examples of uncertain variables, their associated distributions, and sampled values obtained with a Latin hypercube sample (Sect. 5.6) of size 100 (see App. PAR, U.S. DOE 1996, and App., Helton et al. 1998a for distributions of the $nX = 31$ variables in $\mathbf{x}_{su}$).

5.0 Sampling Procedures

Sampling-based methods for uncertainty and sensitivity analysis obviously require sampling procedures. Three sampling procedures are discussed in this section: Random sampling (Sect. 5.1), importance sampling (Sect. 5.2), and Latin hypercube sampling (Sect. 5.3). Random and Latin hypercube sampling are compared with a simple function (Sect. 5.4). A correlation control procedure for use in conjunction with random and Latin hypercube sampling is then discussed (Sect. 5.5). Finally, the use of Latin hypercube sampling to generate a sample from the probability space ($\mathcal{S}_{su}$, $\mathcal{A}_{su}$, p_{su}) for subjective uncertainty introduced in Sect. 3 is described in Sect. 5.6.

5.1 Random Sampling

For notational convenience, assume that the variables under consideration are represented by

$$\mathbf{x} = [x_1, x_2, ..., x_{nX}] \tag{5.1}$$

and that the corresponding probability space is ($\mathcal{S}$, $\mathcal{A}$, p). In random sampling, sometimes also called simple random sampling, the observations

$$\mathbf{x}_k = [x_{k1}, x_{k2}, ..., x_{k,nX}], \; k = 1, 2, ..., nR, \tag{5.2}$$

where nR is the sample size, are selected according to the joint probability distribution for the elements of $\mathbf{x}$ as defined by ($\mathcal{S}$, $\mathcal{A}$, p). In practice, ($\mathcal{S}$, $\mathcal{A}$, p) is defined by specifying a distribution D_j for each element x_j of $\mathbf{x}$ as indicated in Eq. (1.3). Points from different regions of the sample space $\mathcal{S}$ occur in direct relationship to the probability of occurrence of these regions. Further, each sample element is selected independently of all other sample elements. As illustrated in Fig. 5.1 for $x_1 = U$, $x_2 = V$, $nX = 2$ and $nR = 5$, the numbers $RU(1)$, $RU(2)$, ..., $RU(5)$ are sampled from a uniform distribution on [0, 1] and in turn lead to a sample $U(1)$, $U(2)$, ..., $U(5)$ from U based on the CDF for U. Similarly, the numbers $RV(1)$, $RV(2)$, ..., $RV(5)$ lead to a sample $V(1)$, $V(2)$, ..., $V(5)$ from V. The pairs

$$\mathbf{x}_k = [U(k), V(k)], \; k = 1, 2, ..., nR = 5, \tag{5.3}$$

then constitute a random sample from $\mathbf{x} = [U, V]$, where U has a normal distribution on [−1, 1] and V has a triangular distribution on [0, 4] in this example.

Random samples are generated in an analogous manner when $\mathbf{x}$ has a dimensionality greater than 2 (e.g., $nX = 100$). Specifically, if the elements of $\mathbf{x}$ are represented by U, V, ..., W and a random sample of size nR is to be generated, then random numbers $RU(1)$, $RU(2)$, ..., $RU(nR)$ are sampled uniformly from [0,1] and used to obtain corresponding values $U(1)$, $U(2)$, ..., $U(nR)$ for U; random numbers $RV(1)$, $RV(2)$, ..., $RV(nR)$ are sampled uniformly from [0,1] and used to obtain corresponding values $V(1)$, $V(2)$, ..., $V(nR)$ for V, and so on, with the process continuing through all elements of $\mathbf{x}$ and ending with the selection of random numbers $RW(1)$, $RW(2)$, ..., $RW(nR)$ from [0,1] and the generation of the corresponding values $W(1)$, $W(2)$, ..., $W(nR)$ for W. The vectors

$$\mathbf{x}_k = [U(k), V(k), ..., W(k)], \; k = 1, 2, ..., nR, \tag{5.4}$$

then constitute a random sample from $\mathbf{x} = [U,V, ...,W]$.

The preceding sampling procedure depends on the generation of random samples from a uniform distribution on [0, 1] (i.e., uniform random variates). The generation of such samples is widely discussed (e.g., Press et al. 1992, Barry 1996, Fishman 1996, L'Ecuyer 1998), and the capability to do so is taken for granted in this presentation.

5.2 Importance Sampling

In random sampling, there is no assurance that points will be sampled from any given sub-region of the sample space $\mathcal{S}$. Also, it is possible for an inefficient sampling of $\mathcal{S}$ to occur due to several sampled values falling very close together. The preceding problems can be partially ameliorated by using importance sampling. With this technique, $\mathcal{S}$ is exhaustively divided into a number of nonoverlapping subregions (i.e., strata) $\mathcal{S}_i$, $i = 1, 2, ..., nS$. Then, nS_i values for $\mathbf{x}$ are randomly sampled from $\mathcal{S}_i$, with the random sampling carried out in consistency with the definition of ($\mathcal{S}$, $\mathcal{A}$, p) and the restriction of $\mathbf{x}$ to $\mathcal{S}_i$. The resultant vectors

$$\mathbf{x}_k = [x_{k1}, x_{k2}, ..., x_{k,nX}], \; k = 1, 2, ..., \sum_{i=1}^{nS} nS_i \,, \tag{5.5}$$

then constitute an importance-based sample from $\mathcal{S}$ (i.e., a sample obtained by importance sampling). Typically, only one value is sampled from each $\mathcal{S}_i$, with the result that the sample has the form

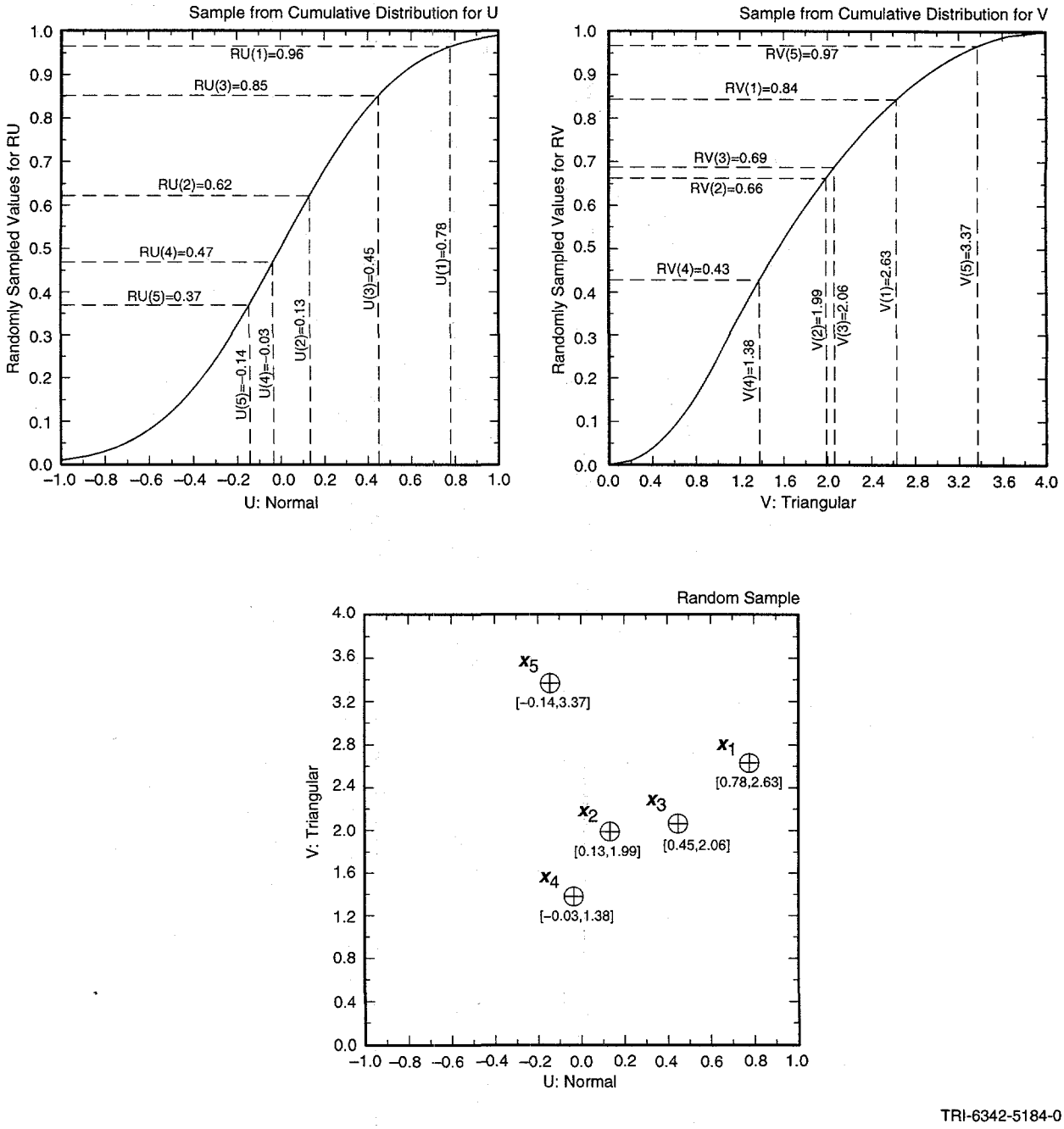


Fig. 5.1. Example of random sampling to generate a sample of size $nR = 5$ from $\mathbf{x} = [U, V]$, with U normal on $[-1, 1]$ (mean = 0, 0.01 quantile = −1, 0.99 quantile = 1) and V triangular on $[0, 4]$ (mode = 1).

$$\mathbf{x}_k = [x_{k1}, x_{k2}, ..., x_{k,nX}], \quad k = 1, 2, ..., nS. \tag{5.6}$$

The name importance sampling derives from the fact that the S_i are in part defined on the basis of how important the $\mathbf{x}$'s contained in each set are to the final outcome of the analysis. Often, importance sampling is used to ensure the inclusion in an analysis of $\mathbf{x}$'s that have high consequences but low probabilities (i.e., the probabilities $p(S_i)$ are small for the S_i that contain such $\mathbf{x}$'s). When importance sampling is used, the probabilities $p(S_i)$ and number of observations nS_i taken from each S_i must be folded back into the analysis before results can be meaningfully presented.

Several examples of importance sampling for $\mathbf{x} = [U, V]$ are given in Fig. 5.2. The two top frames are for strata of equal probability (i.e., all $p(S_i)$ are equal). For two uniform distributions, this results in all strata having the same area (upper left frame). For two nonuniform distributions, different strata can have different areas even though they have the same probability (upper right frame). The two lower frames are for strata of unequal probability. In this case, the variable distributions and the strata probabilities interact to determine the area of the strata. However, it is important to recognize that specifying variable distributions, number of strata and strata probabilities does not uniquely define an importance sampling procedure; rather, there are many ways in which the strata S_i can be defined that are consistent for the preceding constraints. In particular, appropriate definition of strata will depend on specific properties of individual analyses. Similar ideas also hold for more than two variables, in which case the strata become volumes in a space with the same dimension as $\mathbf{x}$.

5.3 Latin Hypercube Sampling

Importance sampling operates to ensure the full coverage of specified regions in the sample space. This idea is carried farther in Latin hypercube sampling (McKay et al. 1979) to ensure the full coverage of the range of each variable. Specifically, the range of each variable (i.e., the x_j) is divided into *nLHS* intervals of equal probability and one value is selected at random from each interval. The *nLHS* values thus obtained for x_1 are paired at random and without replacement with the *nLHS* values obtained for x_2. These *nLHS* pairs are combined in a random manner without replacement with the *nLHS* values of x_3 to form *nLHS* triples. This process is continued until a set of *nLHS* *nX*-tuples is formed. These *nX*-tuples are of the form

$$\mathbf{x}_k = [x_{k1}, x_{k2}, ..., x_{k,nX}], \quad k = 1, ..., nLHS, \tag{5.7}$$

and constitute the Latin hypercube sample (LHS). The individual x_j must be independent for the preceding construction procedure to work; a method for generating Latin hypercube and random samples from correlated variables has been developed by Iman and Conover (1982) and is discussed in Sect. 5.5. Latin hypercube sampling is an extension of quota sampling (Steinberg 1963) and can be viewed as an *n*-dimensional randomized generalization of Latin square sampling (pp. 206-209, Raj 1968).

The generation of an LHS of size $nLHS = 5$ from $\mathbf{x} = [U, V]$ is illustrated in Fig. 5.3. Initially, the ranges of U and V are subdivided into five intervals of equal probability, with this subdivision represented by the lines that originate at 0.2, 0.4, 0.6 and 0.8 on the ordinates of the two upper frames in Fig. 5.3, extend horizontally to the CDFs, and then drop vertically to the abscissas to produce the 5 indicated intervals. Random values $U(1)$, $U(2)$, ..., $U(5)$ and $V(1)$, $V(2)$, ..., $V(5)$ are then sampled from these intervals. The sampling of these random values is implemented by (i) sampling $RU(1)$ and $RV(1)$ from a uniform distribution on [0, 0.2], $RU(2)$ and $RV(2)$ from a uniform distribution on [0.2, 0.4], and so on, and (ii) then using the CDFs to identify (i.e., sample) the corresponding U and V values, with this identification represented by the dashed lines that originate on the ordinates of the two upper frames in Fig. 5.3, extend horizontally to the CDFs, and then drop vertically to the abscissas to produce $U(1)$, $U(2)$, ..., $U(5)$ and $V(1)$, $V(2)$, ..., $V(5)$. The generation of the LHS is then completed by randomly pairing (without replacement) the resulting values for U and V. As this pairing is not unique, many possible LHSs can result. Two such *LHS*s are shown in the lower two frames in Fig. 5.3, with one LHS resulting from the pairings [$U(1)$, $V(5)$], [$U(2)$, $V(1)$], [$U(3)$, $V(2)$], [$U(4)$, $V(3)$], [$U(5)$, $V(4)$] (lower left frame) and the other LHS resulting from the pairings [$U(1)$, $V(3)$], [$U(2)$, $V(2)$], [$U(3)$, $V(4)$], [$U(4)$, $V(5)$], [$U(5)$, $V(1)$] (lower right frame).

The generation of an LHS for $nX > 2$ proceeds in a manner similar to that shown in Fig. 5.3 for $nX = 2$. The sampling of the individual variables for $nX > 2$ takes place in the same manner as shown in Fig. 5.3. However, the nX variables define an nX-dimensional solid rather than a 2-dimensional rectangle in the plane. Thus, the two lower frames in Fig. 5.3 would involve a partitioning of an nX-dimensional solid rather than a rectangle.

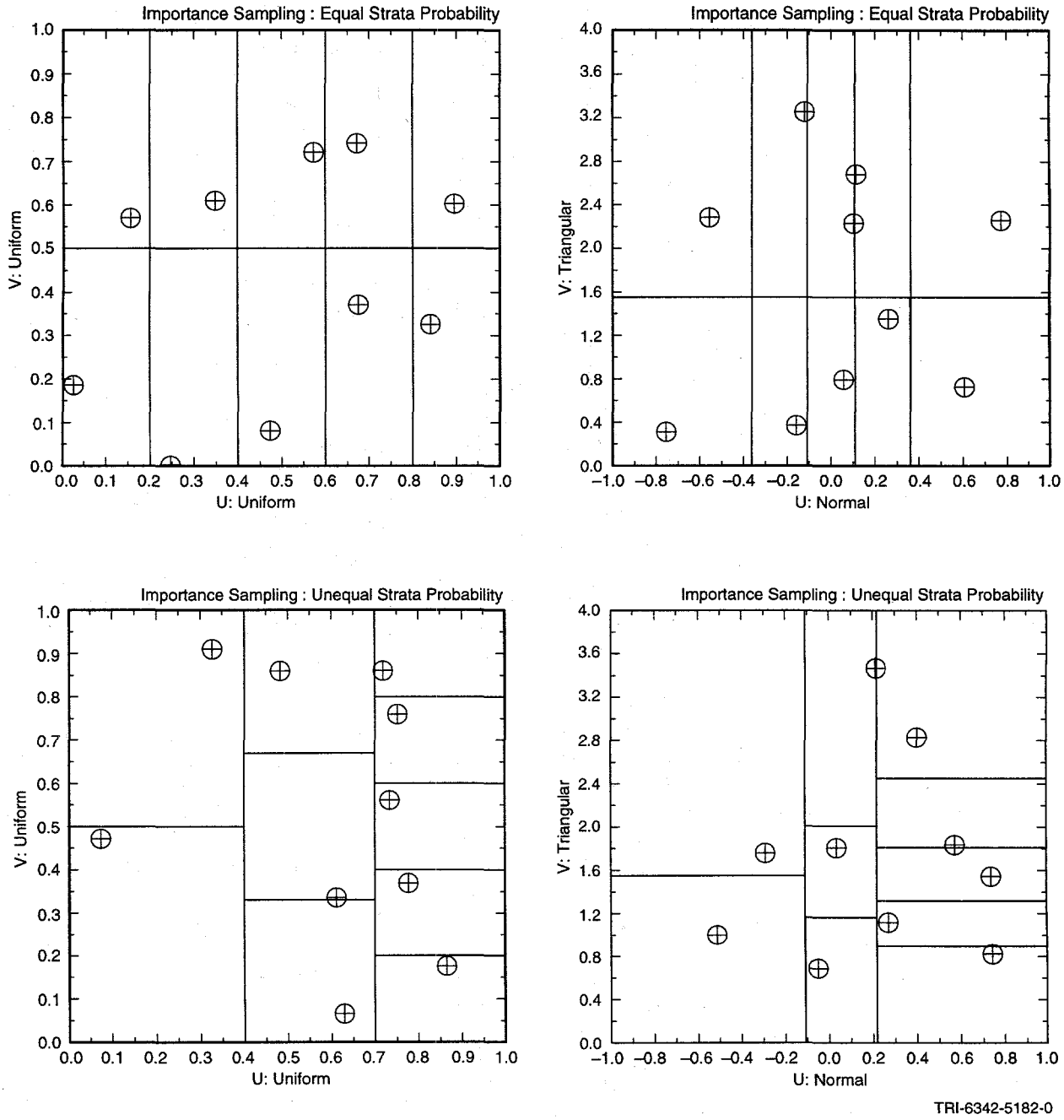


Fig. 5.2. Examples of importance sampling with ten strata (i.e., $nS = 10$), one random sample per strata (i.e., $nS_i = 1$), equal strata probability (i.e., $p(S_i) = 1/10$, upper frames), unequal strata probability (i.e., $p(S_i) = 0.2, 0.2, 0.1, 0.1, 0.1, 0.06, 0.06, 0.06, 0.06, 0.06$, lower frames), U and V uniform on [0, 1] (left frames) and U normal on [−1, 1] (mean = 0, 0.01 quantile = −1, 0.99 quantile = 1) and V triangular on [0, 4] (mode = 1) (right frames).

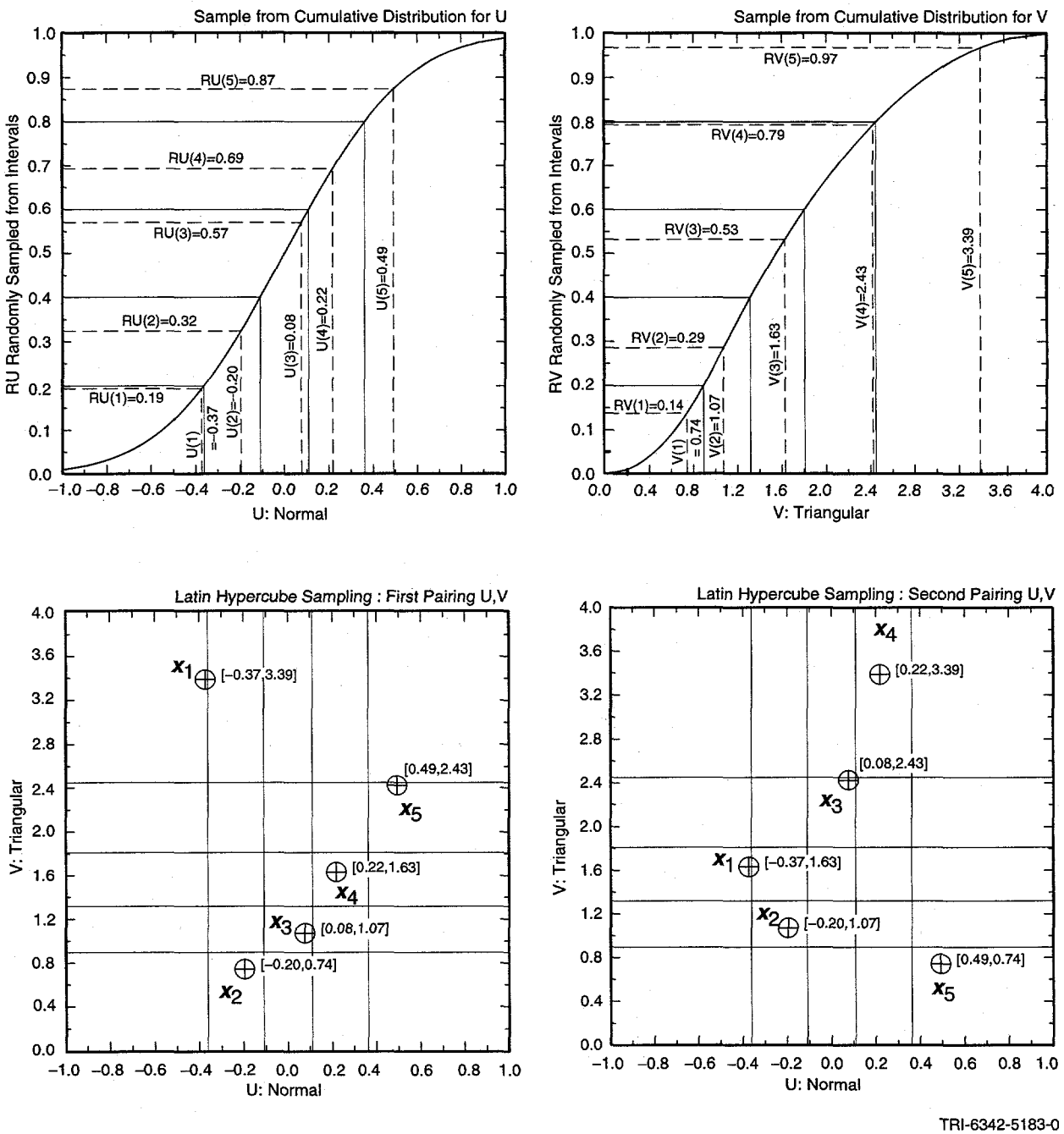


Fig. 5.3. Example of Latin hypercube sampling to generate a sample of size $nLHS = 5$ from $\mathbf{x} = [U, V]$ with U normal on [−1, 1] (mean = 0, 0.01 quantile = −1, 0.99 quantile = 1) and V triangular on [1, 4] (mode = 1).

5.4 Comparison of Random and Latin Hypercube Sampling

Random sampling is the preferred technique when sufficiently large samples are possible because it is easy to implement, easy to explain, and provides unbiased estimates for means, variances and distribution functions. The possible problems with random sampling derive from the rather vague phrase "sufficiently large" in the preceding sentence. When the underlying models are expensive to evaluate (e.g., many hours of CPU time per evaluation) or estimates of extreme quantiles are needed (e.g., the 0.999999 quantile), the required sample size to achieve a specific purpose may be too large to be computationally practicable. In the 1996 WIPP PA, random sampling was used for the estimation of complementary cumulative distribution functions (CCDFs) for radionuclide releases to the accessible environment (i.e., for integration over $(\mathcal{S}_{st}, \mathcal{J}_{st}, p_{st})$; see Sect. 10) because it was possible to develop a computational strategy that allowed the use of a sample of size nS = 10,000 to estimate an exceedance probability of 0.001.

When random sampling is not computationally feasible for the estimation of extreme quantiles, importance sampling is often employed. However, the use of importance sampling on nontrivial problems is not easy due to the difficulty of defining the necessary strata and also of calculating the probabilities of these strata. For example, the fault and event tree techniques used in probabilistic risk assessments for nuclear power stations and other complex engineered facilities can be viewed as algorithms for defining importance sampling procedures. The bottom line is that the definition and implementation of an importance sampling procedure is not easy. Further, without extensive *a priori* knowledge, the strata may end up being defined more finely than is necessary, with the result that the importance sampling procedure ends up requiring more calculations than the use of random sampling to calculate the same outcomes. For example, the number of strata in the importance sampling procedure used to estimate CCDFs in the 1991 and 1992 WIPP PAs (Helton and Iuzzolino 1993) greatly exceeds the size of the random samples used in the 1996 WIPP PA to estimate CCDFs. The unequal strata probabilities also make the outcomes of analyses based on importance sampling inconvenient for use in sensitivity analyses (e.g., how does one interpret a scatterplot or a regression analysis derived from results obtained from an importance sampling procedure?).

Latin hypercube sampling is used when large samples are not computationally practicable and the estimation of very high quantiles is not required. The preceding is typically the case in uncertainty and sensitivity studies to assess the effects of subjective uncertainty. First, the models under consideration are often computationally demanding, with the result that the number of calculations that can be performed to support the analysis is necessarily limited. For example, the totality of the model calculations (i.e., BRAGFLO, NUTS, PANEL, GRASP_INV, SECOFL2D, SECOTP2D, CUTTINGS_S, BRAGFLO_DBR; see Sect. 10) in the 1996 WIPP PA was too extensive to permit the generation of thousands of CCDFs in an uncertainty/sensitivity study to assess the effects of subjective uncertainty on compliance with environmental regulations (i.e., 40 CFR 191.13; see Sect. 10). Second, the estimation of very high quantiles is generally not required in an analysis to assess the effects of subjective uncertainty. Typically, a 0.90 or 0.95 quantile is adequate to establish where the available information indicates a particular analysis outcome is likely to be located; in particular, a 0.99, 0.999 or 0.9999 quantile is usually not needed in assessing the effects of subjective uncertainty.

Desirable features of Latin hypercube sampling include unbiased estimates for means and distribution functions and dense stratification across the range of each sampled variable (McKay et al. 1979). In particular, uncertainty and sensitivity analysis results obtained with Latin hypercube sampling have been observed to be quite robust even when relatively small samples (i.e., $nLHS$ = 50 to 200) are used (Iman and Helton 1988, 1991; Helton et al. 1995a).

For perspective, Latin hypercube and random sampling are illustrated in Fig. 5.4 for two different distribution pairs. To facilitate comparisons, the grid that underlies the LHSs is also shown for the random samples, although it plays no role in the actual generation of these samples. The desirability of Latin hypercube sampling derives from the full coverage of the range of the sampled variables; specifically, each equal probability interval for U and also each equal probability interval for V has exactly one value sampled from it. In contrast, random sampling makes less efficient use of the sampled points, with the possibility existing that significant parts of a variable's range will be omitted (e.g., only one value below the 0.5 quantile for U in the lower left frame and no values for U below the 0.19 quantile nor above the 0.85 quantile in the lower right frame) and that other parts will be overemphasized

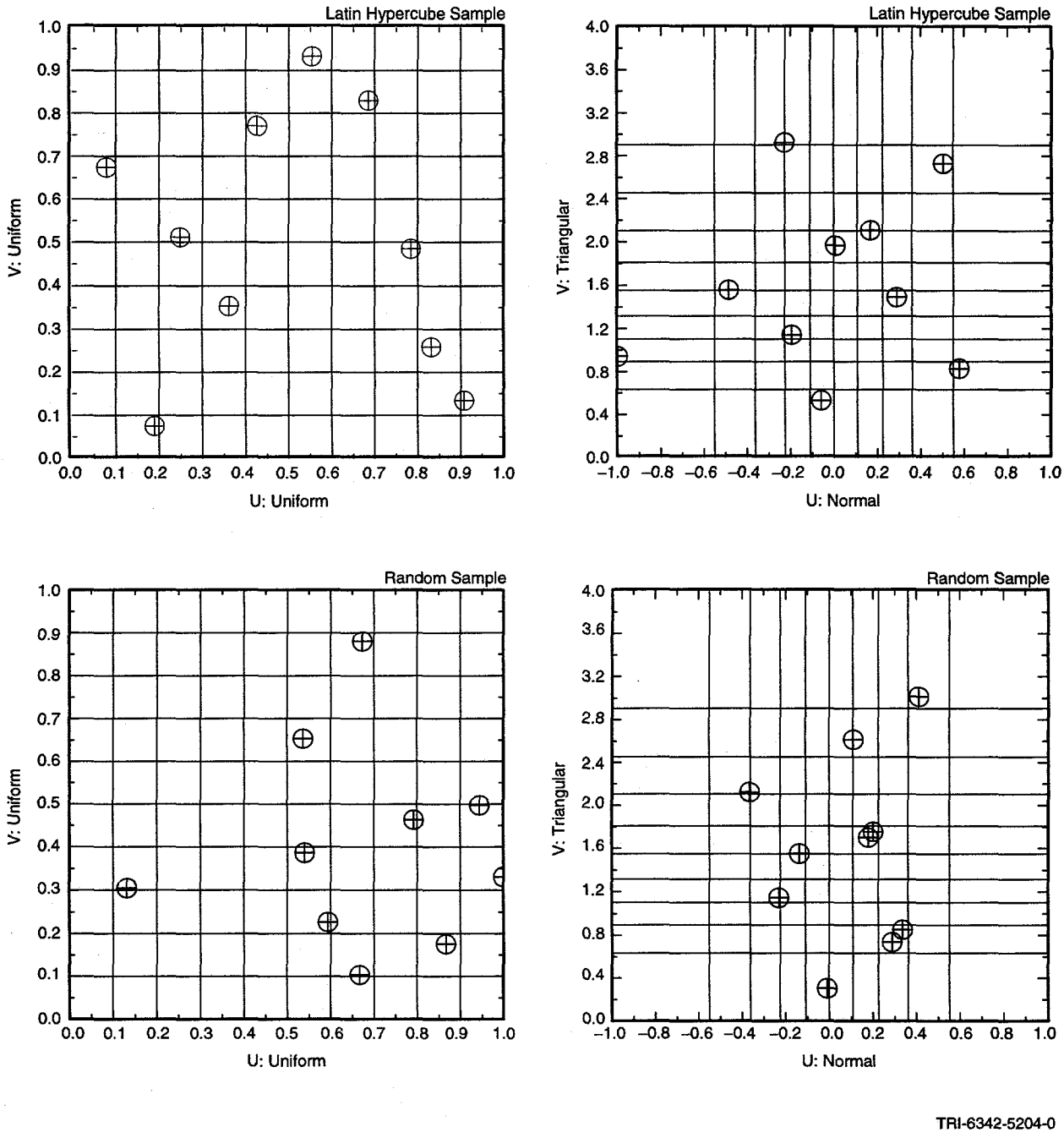


Fig. 5.4. Examples of Latin hypercube and random sampling to generate a sample of size 10 from variables U and V with (1) U and V uniform on [0, 1] (left frames), and (2) U normal on [−1, 1] (mean = 0, 0.01 quantile = −1, 0.99 quantile = 1) and V triangular on [0, 4] (mode = 1) (right frames).

(e.g., 5 out of 10 values for U fall between the 0.5 and 0.7 quantiles for U in the lower left frame, and two pairs of sampled points fall close together in the lower right frame). The enforced stratification in Latin hypercube sampling prevents such inefficient samplings while still providing unbiased estimates for means and distribution functions.

The outcome of the enforced stratification associated with Latin hypercube sampling is that estimates of means and distribution functions tend to be more stable when generated by Latin hypercube sampling than by random sampling. Here, stability refers to the amount of variation between results obtained with different samples generated by the particular sampling technique under consideration. This stability can be illustrated by comparison of estimates of the CDF for the simple function

$$f(U, V) = U + V + UV \tag{5.8}$$

obtained with Latin hypercube and random sampling under the assumption that U and V are uniformly distributed on [0, 2]. In particular, each sampling technique is used to generate 100 samples of size 10 and also 100 samples of size 100 from U and V. Each sample gives rise to an estimated CDF for f (Fig. 5.5). The goal is to compare the variability between the estimates obtained with Latin hypercube and random sampling.

Presenting plots similar to those in Fig. 5.5 for 100 CDFs at a time is not very informative because the CDFs tend to turn into a solid black mass. A more informative presentation is to summarize the distributions of CDFs with mean and percentile curves. The location of the percentile curves then provides an indication of how stable the estimates of the CDFs are. In particular, limited separation between low and high percentiles (e.g., the 10th and 90th) indicates that the sampling procedure is providing stable estimates of the CDF (i.e., there is little variability in the estimated CDF from one sample to the next); in contrast, a large spread between low and high percentiles indicates that the sampling procedure is not providing stable estimates of the CDF (i.e., there is substantial variability in the estimated CDF from one sample to the next). The previously indicated 100 samples of size 10 and 100 are summarized in this manner in Fig. 5.6. Further, the analysis was replicated three times to give three estimates of the 10th percentile, three estimates of the 50th percentile, and so on.

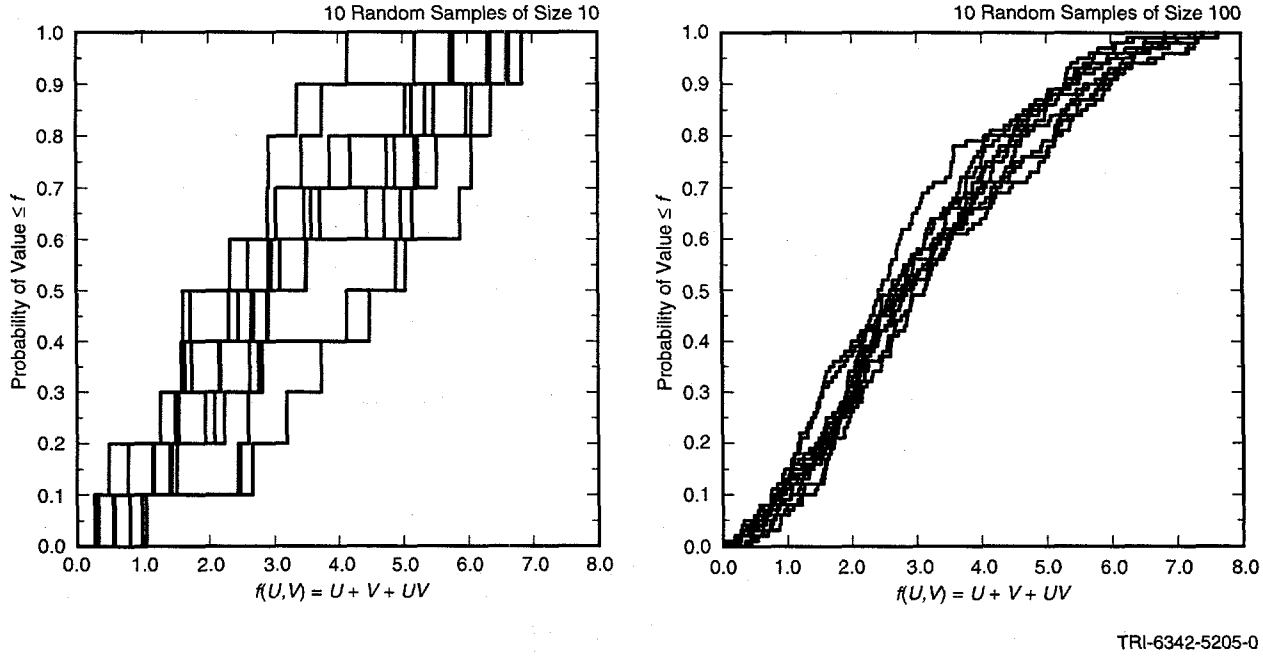


Fig. 5.5. Example CDFs for $f(U, V) = U + V + UV$ estimated with random samples of size 10 and 100 under the assumption that U and V are uniformly distributed on [0, 2].

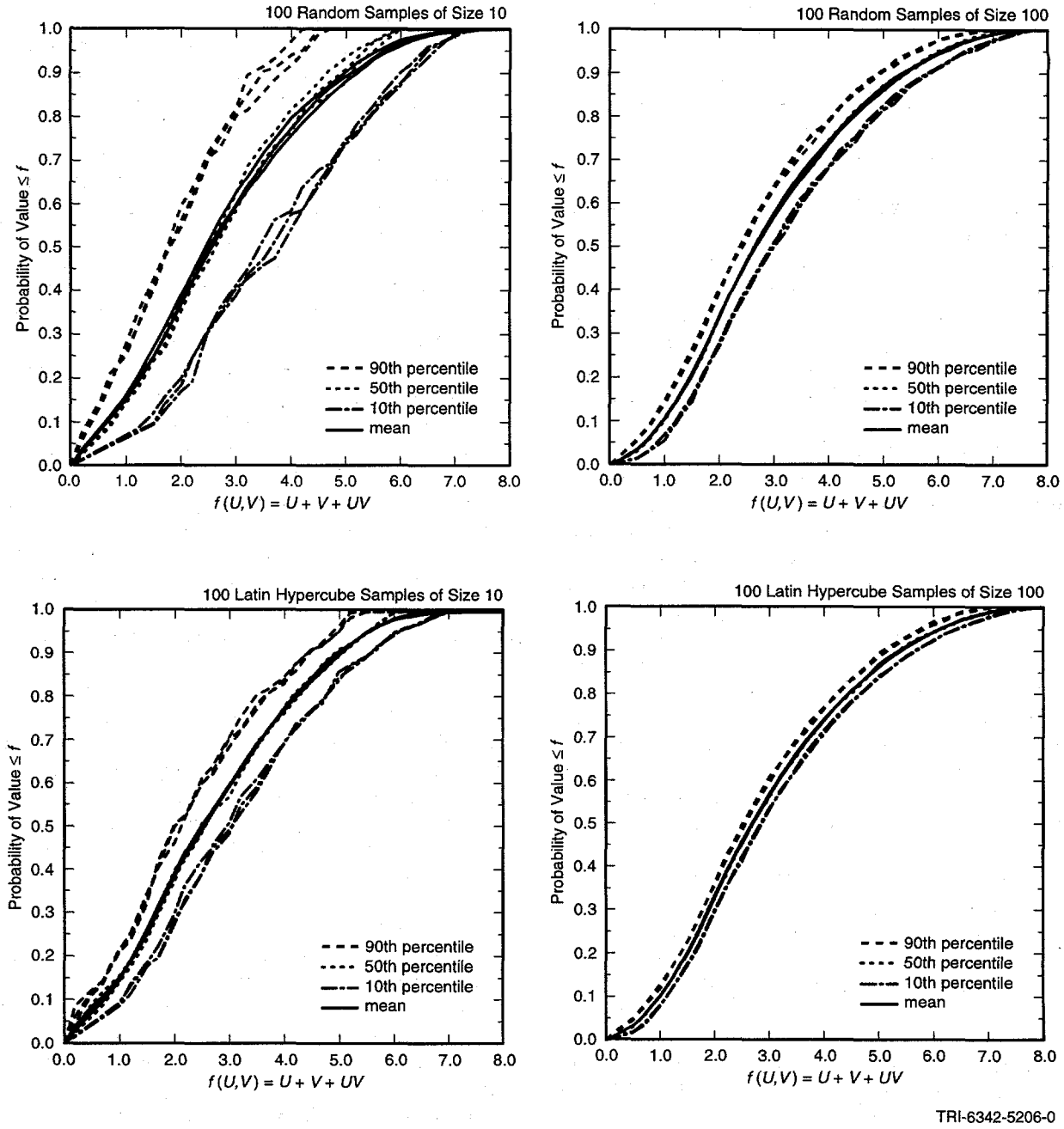


Fig. 5.6. Summary of distribution of CDFs for $f(U, V) = U + V + UV$ estimated with 3 replications of 100 Latin hypercube samples and 100 random samples of size 10 and 100 under the assumption that U and V are uniformly distributed on [0, 2].

As examination of Fig. 5.6 shows, Latin hypercube sampling is producing CDF estimates that are more stable than those produced by random sampling (i.e., the spread between the 10th and 90th percentile curves is tighter for Latin hypercube sampling than for random sampling). The stability of the mean and percentile estimates across the three replicates indicates that the observed stability is real rather than a chance occurrence associated with a particular set of 100 Latin hypercube or random samples.

From the perspective of uncertainty and sensitivity analysis, the full stratification over the range of each sampled variable is a particularly desirable property of Latin hypercube sampling. In a large study, there are potentially hundreds of predicted variables that will be examined at some point in associated uncertainty and sensitivity analyses. Further, it is likely that almost every sampled variable will be important with respect to at least one of these predicted variables. With Latin hypercube sampling, every variable gets equal treatment (i.e., full stratification) within the sample; should a variable be important with respect to a particular output variable, it has been sampled in a way that will permit this importance to be identified. In contrast, it is very difficult to design an importance sampling procedure that provides acceptable results for a large number of sampled and predicted variables. In some sense, Latin hypercube sampling can be viewed as a compromise importance sampling procedure when *a priori* knowledge of the relationships between the sampled and predicted variables is not available. When random sampling is used with a small sample size in an analysis that involves a large number of sampled and predicted variables, the possibility exists that the chance structure of the sample will result in a poor representation of the relationships between some of the sampled and predicted variables. Such poor relationships can also occur for Latin hypercube sampling when several sampled variables affect a given predicted variable, but are less likely to occur than is the case with random sampling.

Formal results involving Latin hypercube sampling and other sampling procedures are available in a number of publications (e.g., Owen 1992, Stein 1987, Iman and Conover 1982, McKay et al. 1979).

5.5 Correlation Control

Control of correlation within a sample can be very important. If two or more variables are correlated, then it is necessary that the appropriate correlation structure be incorporated into the sample if meaningful results are to be obtained in subsequent uncertainty/sensitivity studies. On the other hand, it is equally important that variables do not appear to be correlated when they are really independent.

It is often difficult to induce a desired correlation structure on a sample. Indeed, multivariate distributions can be incompatible with correlation patterns that are proposed for them. Thus, it is possible to encounter analysis situations where the proposed variable distributions and the suggested correlations between the variables are inconsistent; that is, it is not possible to have both the desired variable distributions and the requested correlations between the variables.

In response to this situation, Iman and Conover (1982) proposed a method for controlling the correlation structure in random and Latin hypercube samples that is based on rank correlation (i.e., on rank-transformed variables) rather than sample correlation (i.e., on the original untransformed data). With their technique, it is possible to induce any desired rank-correlation structure onto the sample. This technique has a number of desirable properties: (i) It is distribution free. That is, it may be used with equal facility on all types of distribution functions. (ii) It is simple. No unusual mathematical techniques are required to implement the method. (iii) It can be applied to any sampling scheme for which correlated input variables can logically be considered, while preserving the intent of the sampling scheme. That is, the same numbers originally selected as input values are retained; only their pairing is affected to achieve the desired rank correlations. This means that in Latin hypercube sampling the integrity of the intervals is maintained. If some other structure is used for selection of values, that same structure is retained. (iv) The marginal distributions remain intact.

For many, if not most, uncertainty/sensitivity analysis problems, rank-correlation is probably a more natural measure of congruent variable behavior than is the more traditional sample correlation. What is known in most situations is some idea of the extent to which variables tend to move up or down together; more detailed assessments of variable linkage are usually not available. It is precisely this level of knowledge that rank correlation captures.

The following discussion provides an overview of the Iman/Conover procedure for inducing a desired rank correlation structure on either a random or a Latin hypercube sample and is adapted from Sect. 3.2 of Helton 1993b. The procedure begins with a sample of size m from the n input variables under consideration. This sample can be represented by the $m \times n$ matrix

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{5.9}$$

where x_{ij} is the value for variable j in sample element i. Thus, the rows of **X** correspond to sample elements, and the columns of **X** contain the sampled values for individual variables.

The procedure is based on rearranging the values in the individual columns of **X** so that a desired rank correlation structure results between the individual variables. For convenience, let the desired correlation structure be represented by the $n \times n$ matrix

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \tag{5.10}$$

where c_{kl} is the desired rank correlation between variables x_k and x_l.

Although the procedure is based on rearranging the values in the individual columns of **X** to obtain a new matrix **X*** that has a rank correlation structure close to that described by **C**, it is not possible to work directly with **X**. Rather, it is necessary to define a new matrix

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{m1} & s_{m2} & \cdots & s_{mn} \end{bmatrix} \tag{5.11}$$

that has the same dimensions as **X**, but is otherwise independent of **X**. Each column of **S** contains a random permutation of the m van der Waerden scores (Conover 1980) $\Phi^{-1}(i/m + 1)$, $i = 1, 2, \ldots, m$, where Φ^{-1} is the inverse of the standard normal distribution. The matrix **S** is then rearranged to obtain the correlation structure defined by **C**. This rearrangement is based on the Cholesky factorization (Golub and van Loan 1983) of **C**. That is, a lower triangular matrix P is constructed such that

$$\mathbf{C} = \mathbf{P}\mathbf{P}^T. \tag{5.12}$$

This construction is possible because **C** is a symmetric, positive-definite matrix (Golub and van Loan 1983, p. 88).

If the correlation matrix associated with **S** is the $n \times n$ identity matrix (i.e., if the correlations between the values in different columns of **S** are zero), then the correlation matrix for

$$\mathbf{S}^* = \mathbf{S}\mathbf{P}^T \tag{5.13}$$

is **C** (Anderson 1984, p. 25). At this point, the success of the procedure depends on the following two conditions: (1) that the correlation matrix associated with **S** be close to the $n \times n$ identity matrix; and (2) that the correlation matrix for **S*** be approximately equal to the rank correlation matrix for **S***. If these two conditions hold, then the desired matrix **X*** can be obtained by simply rearranging the values in the individual columns of **X** in the same rank order as the values in the individual columns of **S***. This is the first time that the variable values contained in **X** enter into the correlation process. When **X*** is constructed in this manner, it will have the same rank correlation matrix as **S***. Thus, the rank correlation matrix for **X*** will approximate **C** to the same extent that the rank correlation matrix for **S*** does.

The condition that the correlation matrix associated with **S** be close to the identity matrix is now considered. For convenience, the correlation matrix for **S** will be represented by **E**. Unfortunately, **E** will not always be the identity matrix. However, it is possible to make a correction for this. The starting point for this correction is the Cholesky factorization for **E**:

$$\mathbf{E} = \mathbf{Q}\mathbf{Q}^T. \tag{5.14}$$

This factorization exists because **E** is a symmetric, positive-definite matrix. The matrix **S*** defined by

$$\mathbf{S}^* = \mathbf{S}(\mathbf{Q}^{-1})^T\mathbf{P}^T \tag{5.15}$$

has **C** as its correlation matrix. In essence, multiplication of **S** by $(\mathbf{Q}^{-1})^T$ transforms **S** into a matrix whose associated correlation matrix is the $n \times n$ identity matrix; then, multiplication by $\mathbf{P}^T$ produces a matrix whose associated correlation matrix is **C**. As it is not possible to be sure that **E** will be an identity matrix, the matrix **S*** used in the procedure to produce correlated input should be defined in the corrected form shown in Eq. (5.15) rather than in the uncorrected form shown in Eq. (5.13).

The condition that the correlation matrix for **S*** be approximately equal to the rank correlation matrix for **S*** depends on the choice of the scores used in the definition of **S**. On the basis of empirical investigations, Iman and Conover (1982) found that van der Waerden scores provided an effective means of defining **S**, and these scores are incorporated into the rank correlation procedure in the widely used LHS program (Iman and Shortencarier 1984). Other possibilities for defining these scores exist, but have not been extensively investigated. The user should examine the rank correlation matrix associated with **S*** to ensure that it is close to the target correlation matrix **C**. If this is not the case, the construction procedure used to obtain **S*** can be repeated until a suitable approximation to **C** is obtained. Results given in Iman and Conover 1982 indicate that the use of van der Waerden scores leads to rank correlation matrices for **S*** that are close to the target matrix **C**.

Additional information on the Iman/Conover (i.e., restricted pairing) technique to induce a desired rank-correlation structure is given in the original article. Further, the technique is implemented in the widely used LHS program (Iman and Shortencarier 1984). The results of various rank-correlation assumptions are illustrated in Iman and Davenport (1980, 1982).

5.6 Latin Hypercube Sampling in the 1996 WIPP PA

As discussed in Sect. 3, this presentation uses an example from the 1996 WIPP PA. In this analysis, the LHS program (Iman and Shortencarier 1984) was used to produce three independently generated LHSs of size $nLHS = 100$ each from the 31 variables in Table 3.1, for a total of 300 sample elements. Each individual replicate is an LHS of the form

$$\mathbf{x}_{su,k} = [x_{k1}, x_{k2}, \ldots, x_{k,nX}],\ k = 1, 2, \ldots, nLHS = 100, \tag{5.16}$$

with $nX = 31$. The three replicated samples were generated to provide a way to observe the stability of results obtained with Latin hypercube sampling. For notational convenience, the replicates are designated by R1, R2 and R3 for replicates 1, 2 and 3, respectively.

The restricted pairing technique described in Sect. 5.5 was used to induce requested correlations and also to ensure that uncorrelated variables had correlations close to zero. The variable pairs (*ANHCOMP*, *ANHPRM*), (*HALCOMP*, *HALPRM*) and (*BPCOMP*, *BPPRM*) were assigned rank correlations of –0.99, –0.99 and –0.75, respectively (Table 3.1). Further, all other variable pairs were assigned rank correlations of zero. The restricted pairing technique was quite successful in producing these correlations (Table 5.1). Specifically, the correlated variables have correlations that are close to their specified values and uncorrelated variables have correlations that are close to zero.

Table 5.1. Example Rank Correlations in Replicate 1

WGRCOR	1.0000							
WMICDFLG	0.0198	1.0000						
HALCOMP	0.0011	0.0235	1.0000					
HALPRM	–0.0068	–0.0212	–0.9879	1.0000				
ANHCOMP	0.0080	0.0336	–0.0123	–0.0025	1.0000			
ANHPRM	0.0049	–0.0183	0.0037	0.0113	–0.9827	1.0000		
BPCOMP	0.0242	0.1071	–0.0121	0.0057	–0.0184	0.0078	1.0000	
BPPRM	–0.0514	–0.0342	0.0035	0.0097	0.0283	–0.0202	–0.7401	1.0000
	WGRCOR	*WMICDFLG*	*HALCOMP*	*HALPRM*	*ANHCOMP*	*ANHPRM*	*BPCOMP*	*BPPRM*

6.0 Evaluation of Model

Once the sample in Eq. (1.4) has been generated, the corresponding model evaluations in Eq. (1.5) must be carried out. The nature of these evaluations is model specific and outside the scope of this presentation. However, this brief section is included to emphasize that these evaluations are something that must be done as part of a sampling-based uncertainty and sensitivity analysis. If the model under consideration is expensive to evaluate, then this will probably be the most computationally demanding part of the analysis and may significantly influence the sample size selected for use and possibly other aspects of the analysis. For example, the model introduced in Sect. 3 for use as an example requires approximately 4 to 5 hours of CPU time on a VAX Alpha per evaluation (i.e., for each sample element) and produces a large quantity of temporally and spatially variable results. Thus, for this example, carrying out and then saving the necessary model evaluations involved a significant expenditure of human and computational resources. In contrast, this part of the analysis can be relatively undemanding for models that are less complex and computationally intensive.

7.0 Uncertainty Analysis

After the sample in Eq. (1.4) has been generated and the corresponding model evaluations in Eq. (1.5) have been carried out, the primary computational portions of the uncertainty analysis component of a sampling-based analysis have been completed. What remains to be done is to display the uncertainty information contained in the mapping between analysis inputs and analysis results in Eq. (1.6). Two cases are considered: results represented by single numbers (Sect. 7.1), and results represented by functions (Sect. 7.2). Finally, example analysis outcomes illustrating the stability of uncertainty analysis results obtained with Latin hypercube sampling are presented (Sect. 7.3).

7.1 Scalar Results

When a single real-valued result is under consideration, the vector-valued function $\mathbf{y}(\mathbf{x}_k)$ in Eqs. (1.5) and (1.6) becomes the scalar-valued function

$$y_k = y(\mathbf{x}_k),\ k = 1, 2, \ldots, nS. \tag{7.1}$$

One possibility is to summarize the uncertainty in y with a mean and a variance. If random or Latin hypercube sampling was used to generate the results in Eq. (7.1), then estimates $\hat{E}(y)$ and $\hat{V}(y)$ for the expected value and variance of y are given by

$$\hat{E}(y) = \sum_{k=1}^{nS} y_k / nS \tag{7.2}$$

and

$$\hat{V}(y) = \sum_{k=1}^{nS} [y_k - \hat{E}(y)]^2 / (nS - 1)\,. \tag{7.3}$$

If importance sampling was used in the generation of the results in Eq. (7.1), then the probabilities of the individual strata in the importance sampling procedure would have to be used in the determination of $\hat{E}(y)$ and $\hat{V}(y)$.

Although the estimation of means and variances is a possibility for summarizing the uncertainties in scalar-valued results, these quantities do not provide very good summaries of subjective uncertainty for at least two reasons. First, information is always lost in the calculation of means and variances. Specifically, there is more information in the nS numbers in Eq. (7.1) and their associated weights (i.e., the reciprocal of the sample size for random and Latin hypercube sampling and the strata probabilities for importance sampling) than there is in the two numbers in Eqs. (7.2) and (7.3). Second, means and variances are not very natural quantities for summarizing subjective uncertainty. Specifically, the quantiles associated with a distribution summarizing subjective uncertainty convey more meaningful information about where the quantity under consideration is believed to be located.

Distribution functions provide a more effective summary of the information associated with the mapping in Eq. (7.1) than means and variances. In particular, this mapping can be summarized with either a CDF or a CCDF, with the CCDF simply being one minus the CDF (Fig. 7.1). The presence of the included and excluded points in Fig. 7.1 results from the use of a finite number of y values and the inequalities in the definitions of CDFs and CCDFs. Technically, the vertical lines should not be present in the CDF and CCDF in Fig. 7.1 but these lines are typically included to make plots of CDFs and CCDFs easier to read. For the same reason, the distinction between included and excluded points is typically omitted. When random or Latin hypercube sampling is used, the step heights in the definitions of CDFs and CCDFs are the reciprocal of the sample size nS (i.e., $1/nS$ and thus 1/10 in Fig. 7.1); when importance sampling is used, the step heights correspond to the strata probabilities. An example with real data is given in Fig. 7.2.

The value of CDFs and CCDFs is that they provide a display of all the information associated with the mapping in Eq. (7.1). In particular, they allow an easy extraction of the probabilities of having values in different subsets of the range of y. Although CDFs and CCDFs are equivalent in their information content, CCDFs are often used for display purposes when large samples are in use and it is important to display the effects of low probability but high consequence analysis outcomes (i.e., unlikely but large y values); further, CCDFs answer the question "How likely is it to be this bad or worse?", which is typically the question of interest in risk assessments. Given that the distributions assigned to the elements of $\mathbf{x}$ are characterizing subjective uncertainty, then the resultant probabilities

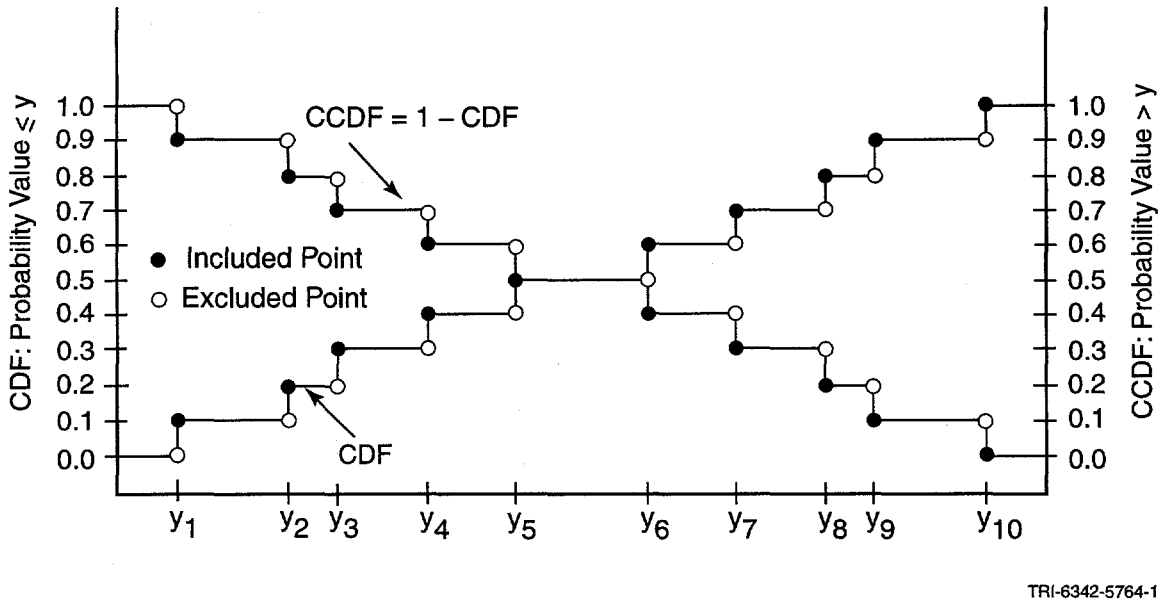


Fig. 7.1. Example of construction of CDFs and CCDFs for a sample of size $nS = 10$ (i.e., $y_k = y(\mathbf{x}_k)$, $k = 1, 2, \ldots, nS = 10$ in Eq. (7.1)).

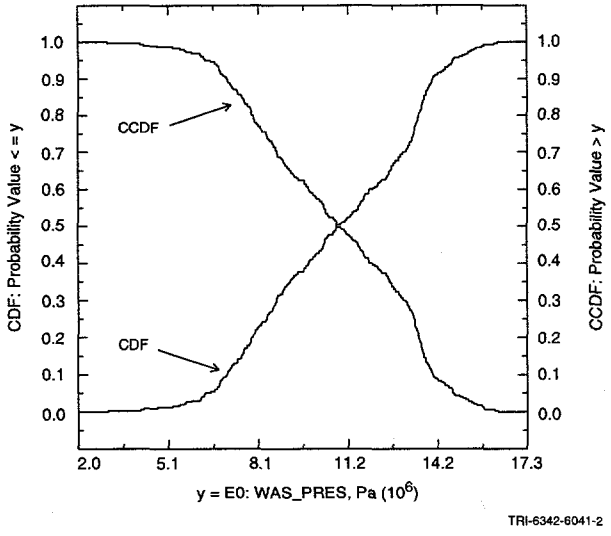


Fig. 7.2. Example of estimated CDF and CCDF for repository pressure at 10,000 yr under undisturbed conditions (i.e., $y = E0{:}WAS\_PRES$) obtained from the 300 LHS elements that result from pooling replicates R1, R2 and R3 (see Sect. 5.6).

extracted from CDFs and CCDFs are also characterizing subjective uncertainty and are thus providing quantitative measures of where the value of y is believed to be located.

Many individuals prefer density functions rather than CDFs or CCDFs for the display of distributions. Density functions have the advantage that they make it easy to identify the mode of a distribution but do not allow an easy extraction of the probabilities associated with various subranges of the dependent variable. Further, unless smoothing procedures are used, the best that can be obtained from the results in Eq. (7.1) is a histogram that approximates the shape of the density function, with the potential that the shape of this histogram will be significantly influenced by the resolution at which the y_k's are binned (Silverman 1986). As recommended by Ibrekk and Morgan (1987), an alternative display is to plot the CDF, the mean, and the associated density function on the same plot frame (Fig. 7.3).

One disadvantage associated with CDFs, CCDFs and density functions is that displays using these distributional summaries can become quite cluttered when results for a number of different analysis outcomes are presented in a single plot frame (e.g., a plot involving CDFs, CCDFs or density functions for 10 different

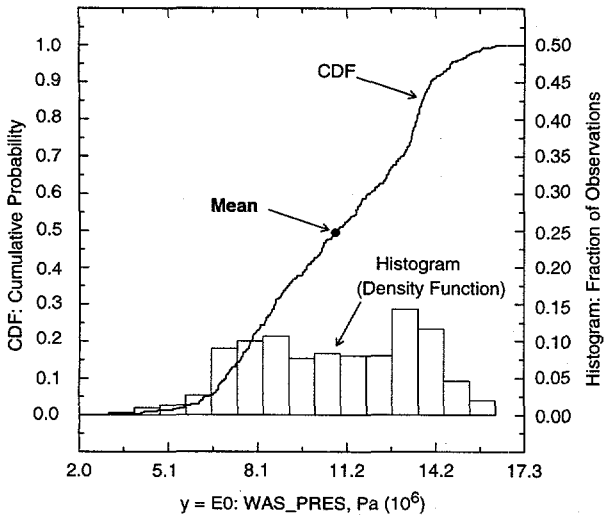


Fig. 7.3. Uncertainty display including estimated distribution function, density function, and mean for repository pressure at 10,000 yr under undisturbed conditions (i.e., y = *E0:WAS_PRES*).

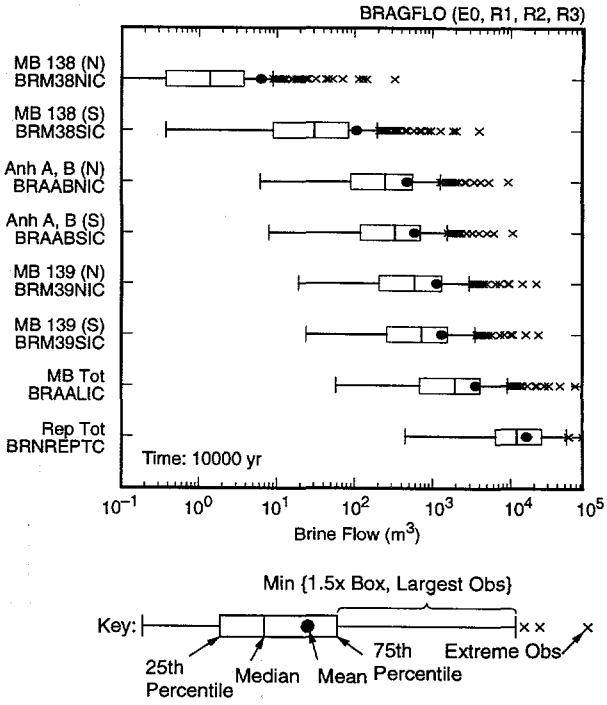


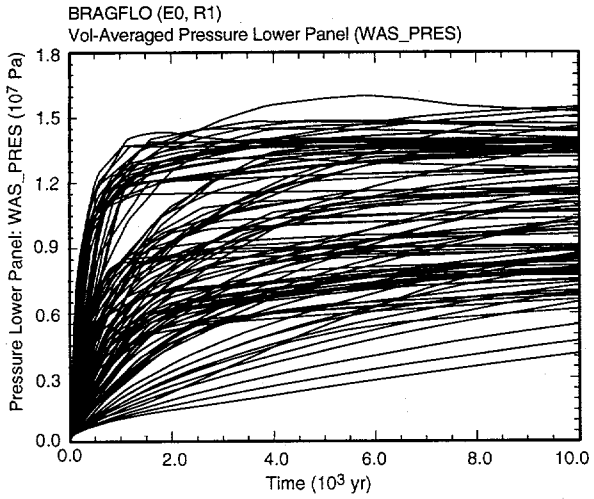
Fig. 7.4. Examples of box plots for cumulative brine flow over 10,000 yr into various regions in disturbed rock zone surrounding repository (*E0:BRM38NIC*, *E0:BRM38SIC*, *E0:BRAABNIC*, *E0:BRAABSIC*, *E0:BRM39NIC*, *E0:BRM39SIC* and *E0:BRAALIC*) and into repository (*E0:BRNREPTC*) under undisturbed conditions in the 1996 WIPP PA (Fig. 7.2.2, Helton et al. 1998a).

analysis outcomes can be hard to read due to the tendency of the individual curves to repeatedly cross each other). Box plots provide an alternative, less congested display of multiple distributions (Fig. 7.4). In such plots, the endpoints of the boxes are formed by the lower and upper quartiles of the data, that is $x_{0.25}$ and $x_{0.75}$. The vertical line within the box represents the median, $x_{0.50}$. The mean is identified by the large dot. The bar on the right of the box extends to the minimum of $x_{0.75} + 1.5(x_{0.75} - x_{0.25})$ and the maximum value. In a similar manner, the bar on the left of the box extends to the maximum of $x_{0.25} - 1.5(x_{0.75} - x_{0.25})$ and the minimum value. The observations falling outside of these bars are shown with crosses. In symmetric distributions, these values would be considered outliers. Box plots contain the same information as a distribution function, but in a somewhat reduced form. Further, their flattened shape makes it convenient to place many distributions on a single plot and also to compare different distributions.

7.2 Functions

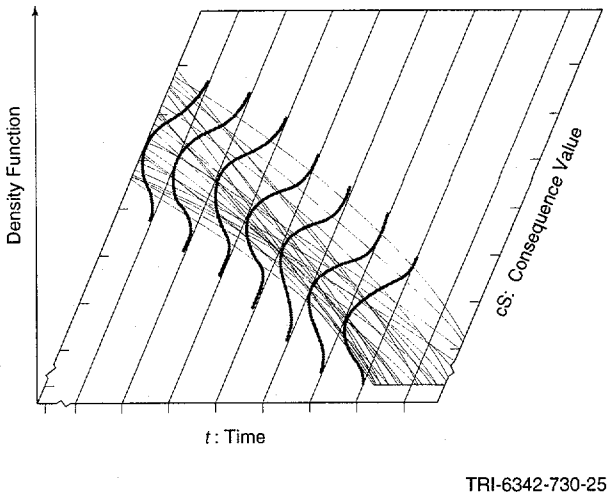
In many analyses, outcomes of interest are functions of one or more variables. In the example used in this presentation, many results are functions of time (Fig. 7.5). Thus, time is the independent variable (i.e., function argument). However, there is also subjective uncertainty in the variables required in the estimation of these functions, with this uncertainty leading to multiple possible functions as illustrated in Fig. 7.5. The estimated distribution presented in Fig. 7.5 was obtained from the LHS in Eq. (5.16) associated with replicate R1 (i.e., each curve in Fig. 7.5 was calculated conditional on the occurrence of one of the sample elements $\mathbf{x}_{su,k}$ in Eq. (5.16) for replicate R1).

The family of curves in Fig. 7.5 is an approximation obtained with an LHS of size 100 to the actual distribution associated with the probability space ($\mathcal{S}_{su}$, $\overline{\mathcal{S}}_{su}$, p_{su}) for subjective uncertainty. Although such families provide an impression of the shape of the associated distributions, they do not directly provide probabilistic information. In concept, these distributions can be summarized by presenting density functions for the values on the ordinate for a sequence of values on the abscissa (Fig. 7.6). In practice, only a



TRI-6342-5721-0a

Fig. 7.5. Repository pressure under undisturbed conditions (i.e., y = *E0:WAS_PRES*) for 100 LHS elements in replicate R1.



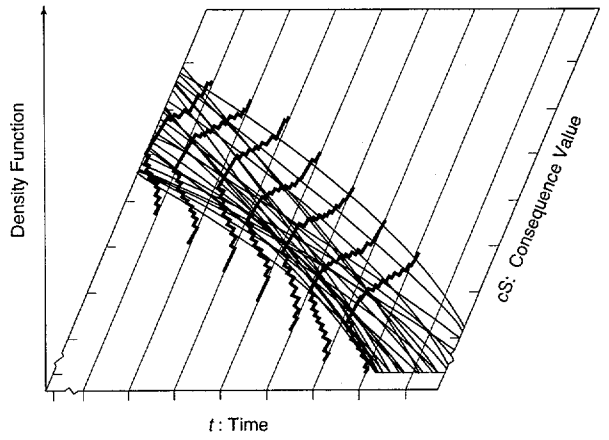
TRI-6342-730-25

Fig. 7.6. Density functions characterizing subjective uncertainty in consequence values for individual times.

finite number of curves will be available as shown in Fig. 7.5, with the result that the density functions indicated in Fig. 7.6 will have to be approximated from these curves (Fig. 7.7).

Although the representations in Figs. 7.6 and 7.7 are intuitively appealing (especially for individuals who like to use density functions to represent distributions), these representations do not seem to work very well in practice. In particular, they are difficult to construct (e.g., construction problems arise if specific consequence values have nonzero probabilities) and also difficult to extract information from (e.g., to determine specific quantile values). Some of these problems could be alleviated by plotting CDFs in the third dimension, but the resultant plots are still difficult to read.

An alternative and often effective representation is to determine mean values and quantiles conditional on individual values on the abscissa and then to plot these means and quantiles above the values for which they were determined (Fig. 7.8). Conceptually, a vertical line is drawn through the curves above a given value on the abscissa (Fig. 7.8a). The locations where this line passes through the individual curves identifies the corresponding consequence values, with the number of consequence values equal to the sample size in use. These values can be used to produce a mean value and also selected quantile values (Fig. 7.8a). If desired, the definition of the mean and quantile values can be



TRI-6342-730-26

Fig. 7.7. Histograms constructed from a sample $\mathbf{x}_{su,k}$, k = 1, 2, ..., *nLHS*, from $\mathcal{S}_{su}$ that characterize subjective uncertainty in consequence values for individual times.

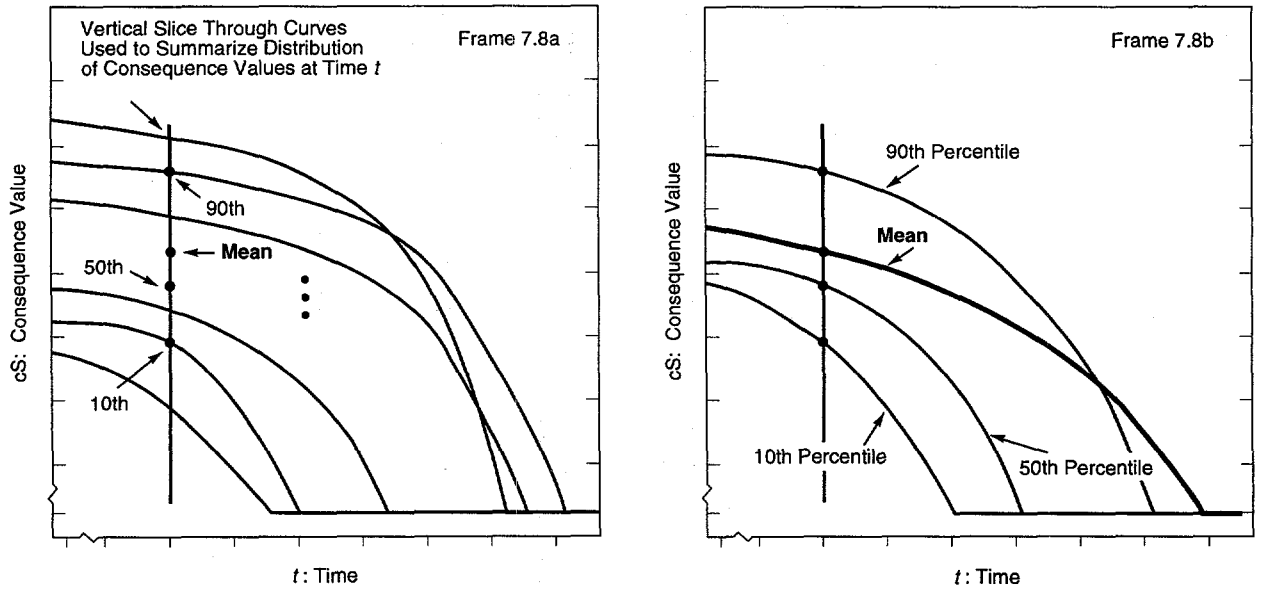
represented formally by integrals over $\mathcal{S}_{su}$ (Helton 1996), with the sampling procedure being used to provide approximations to these integrals. Once the mean and quantile values have been determined, they can be plotted above the corresponding values on the abscissa and then connected to form continuous curves (Fig. 7.8b). With this summary procedure, the quantile values are defined conditional on individual times on the abscissa; as a result, the quantile curves (Fig. 7.8b) should not be viewed as being quantiles for the distribution of curves (i.e., it is inappropriate to assume that

there is a probability of 0.9 that a randomly selected element of $\mathcal{S}_{su}$ will produce a curve that falls below the 0.9 quantile curve indicated in Fig. 7.8b).

Repository pressure has been used as an example of an uncertain function (Figs. 7.5). Estimates of the corresponding mean and quantile curves are given in Fig. 7.9, with these estimates obtained as indicated in conjunction with Fig. 7.8 from the 300 curves that result from pooling the outcomes associated with all three replicates (see Eq. (5.16)). Results such as those given in Fig. 7.9 provide a more quantitative summary of the distribution of curves in Fig. 7.5 than the intuitive impression that is obtained by visually examining the distributions themselves.

7.3 Stability of Results

As indicated in Fig. 5.6, Latin hypercube sampling tends to produce more stable results than random sampling. The reason the LHS in Eq. (5.16) was replicated 3 times (i.e., $nR = 3$) was to provide a measure of the stability of the results obtained in the 1996 WIPP PA. For the pressure results in Fig. 7.5, the results were quite stable from sample to sample (Fig. 7.10). Indeed, the results obtained with the individual replicates were quite stable across the large number of predicted outcomes examined in the analysis, with no instance occurring where different replicates would have lead to different conclusions with respect to system behavior.



TRI-6342-730-27a,b

Fig. 7.8. Mean and quantile curves constructed from a sample $\mathbf{x}_{su,k}$, $k = 1, 2, \ldots, nLHS$, from $\mathcal{S}_{su}$ that characterize uncertainty in consequence values for individual times.

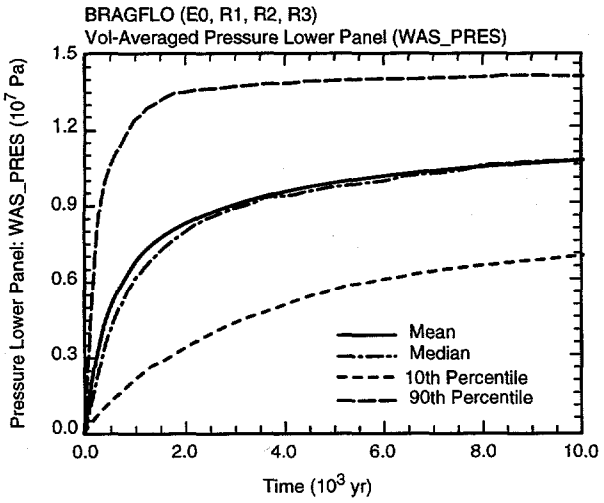


Fig. 7.9. Mean and quantile curves for pressure in lower waste panel under undisturbed conditions (i.e., y = *E0:WAS_PRES*) obtained from the 300 observations that result from pooling replicates R1, R2 and R3.

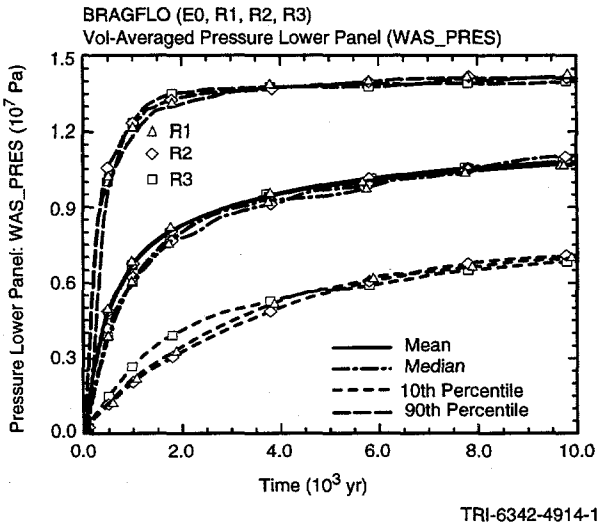


Fig. 7.10. Mean and quantile curves for three replicated LHSs for pressure in lower waste panel under undisturbed conditions (i.e., y = *E0:WAS_PRES*).

This determination leads to the following matrix equation that defines the coefficient vector **b** for which the sum $S(\mathbf{b})$ in Eq. (8.6) is a minimum:

$$\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}. \tag{8.7}$$

For the analysis to produce a unique value for **b**, the matrix $\mathbf{X}^T\mathbf{X}$ must be invertible. Then, **b** is given by

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}^T\mathbf{y}. \tag{8.8}$$

The matrix $\mathbf{X}^T\mathbf{X}$ will always be invertible when the columns of **X** are linearly independent. This is usually the case in a sampling-based study in which the number of sample elements (i.e., m) exceeds the number of independent variables (i.e., n).

The following identity holds for the least squares regression model and plays an important role in assessing the adequacy of such models:

$$\sum_{k=1}^{m}(y_k-\bar{y})^2 = \sum_{k=1}^{m}(\hat{y}_k-\bar{y})^2 + \sum_{k=1}^{m}(\hat{y}_k-y_k)^2, \tag{8.9}$$

where $\hat{y}_k$ denotes the estimate of y_k obtained from the regression model and $\bar{y}$ is the mean of the y_k (Sect. 3.4, Myers 1990). For notational convenience, the preceding equality is often written as

$$SS_{tot} = SS_{reg} + SS_{res}, \tag{8.10}$$

where

$$SS_{tot} = \sum_{k=1}^{m}(y_k-\bar{y})^2,\ SS_{reg} = \sum_{k=1}^{m}(\hat{y}_k-\bar{y})^2,$$

$$SS_{res} = \sum_{k=1}^{m}(\hat{y}_k-y_k)^2.$$

The three preceding summations are called the total sum of squares, regression sum of squares, and residual sum of squares, respectively.

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

$$R^2 = SS_{reg} / SS_{tot} \tag{8.11}$$

provides a measure of the extent to which the regression model can match the observed data. Specifically, when the variation about the regression model is small (i.e., when SS_{res} is a 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 the y_k. Conversely, an R^2 value close to zero indicates that the regression model is not very successful in accounting for the uncertainty in the y_k. Another name for R^2 is the coefficient of multiple determination.

An important situation occurs when the rows of the matrix **X** (i.e., the variable values at which the model is evaluated) are selected so that $\mathbf{X}^T\mathbf{X}$ is a diagonal matrix. In this case, the columns of **X** are said to be orthogonal, and the estimated regression coefficients are given by

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}^T\mathbf{y}$$

$$= \begin{bmatrix} d_0 & 0 & \cdots & 0 \\ 0 & d_1 & & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & & d_n \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{m1} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & & x_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \tag{8.12}$$

and so each element b_j of **b** is given by

$$b_j = \sum_{k=1}^{m} x_{kj}y_k / d_j = \sum_{k=1}^{m} x_{kj}y_k \Big/ \sum_{k=1}^{m} x_{kj}^2 \ . \tag{8.13}$$

Thus, the estimate of the regression coefficient b_j for the variable x_j depends only on the values for x_j in the design matrix **X** (i.e., x_{1j}, ..., x_{mj}). This is true regardless of the number of variables included in the regression. As long as the design is orthogonal, the addition or deletion of variables from the model will not change the regression coefficients for the remaining variables. Further, when the design matrix **X** is orthogonal, the R^2 value for the regression can be expressed as

$$R^2 = SS_{reg} / SS_{tot} = R_1^2 + R_2^2 + \cdots + R_n^2, \tag{8.14}$$

where R_j^2 is the R^2 value that results from regressing y on only x_j (p. 99, Draper and Smith 1981). Thus, R_j^2 is equal to the contribution of x_j to R^2 when the design matrix **X** is orthogonal.

The regression model in Eq. (8.3) can be algebraically reformulated as

$$(\hat{y}-\bar{y})/\hat{s} = \sum_{j=1}^{n}\left(b_j\hat{s}_j/\hat{s}\right)\left(x_j-\bar{x}_j\right)/\hat{s}_j, \tag{8.15}$$

where

$$\bar{y} = \sum_{k=1}^{m} y_k / m, \quad \hat{s} = \left[\sum_{k=1}^{m} (y_k - \bar{y})^2 / (m-1) \right]^{1/2},$$

$$\bar{x}_j = \sum_{k=1}^{m} x_{kj} / m, \quad \hat{s}_j = \left[\sum_{k=1}^{m} (x_{kj} - \bar{x}_j)^2 / (m-1) \right]^{1/2}.$$

The coefficients $b_j \hat{s}_j / \hat{s}$ appearing in Eq. (8.15) are called standardized regression coefficients (SRCs). When the x_j are independent, the absolute value of the SRCs can be used to provide a measure of variable importance. Specifically, the coefficients provide a measure of importance based on the effect of moving each variable away from its expected value by a fixed fraction of its standard deviation while retaining all other variables at their expected values. Calculating SRCs is equivalent to performing the regression analysis with the input and output variables normalized to mean zero and standard deviation one.

An example regression analysis is now given. The output variable (i.e., y) is pressure (Pa) in the repository at 10,000 yr under undisturbed (i.e., E0) conditions (i.e., the pressure values above 10,000 yr in Fig. 7.5). To keep the example at a convenient size, 3 independent variables (i.e., x_j) will be considered (Table 3.1): pointer variable for microbial degradation of cellulose (*WMICDFLG*), halite porosity (*HALPOR*), and corrosion rate for steel (*WGRCOR*). The following regression model is obtained using the preceding three variables and the pooled LHS indicated in conjunction with Eq. (5.16) (i.e., $n = 3$ and $m = 300$):

$$y = 5.72 \times 10^6 + 2.46 \times 10^6 \bullet WMICDFLG + 1.55 \times 10^8 \bullet HALPOR + 1.52 \times 10^{20} \bullet WGRCOR. \quad (8.16)$$

The coefficients in the preceding model show the effect of a one unit change in an input variable (i.e., an x_j) on the output variable (i.e., y). The sign of a regression coefficient indicates whether y tends to increase (a positive regression coefficient) or tends to decrease (a negative regression coefficient) as the corresponding input variable increases. Thus, y tends to increase as each of *WMICDFLG*, *HALPOR* and *WGRCOR* increases.

It is hard to assess variable importance from the regression coefficients in Eq. (8.16) because of the effects of units and distribution assumptions. In particular, the regression coefficient for *WGRCOR* is much larger than the regression coefficients for *WMICDFLG* and *HALPOR*, which does not necessarily imply that *WGRCOR* has greater influence on the uncertainty in y than *WMICDFLG* or *HALPOR*. Variable importance is more clearly shown by the following reformation of Eq. (8.16) with SRCs:

$$y = 0.722\ WMICDFLG + 0.468\ HALPOR + 0.246\ WGRCOR, \quad (8.17)$$

where y, *WMICDFLG*, *HALPOR* and *WGRCOR* have been standardized to mean zero standard deviation one as indicated in Eq. (8.15). The SRCs in Eq. (8.17) provide a better characterization of variable importance than the unstandardized coefficients in Eq. (8.16). For perturbations equal to a fixed fraction of their standard deviation, the impact of *WMICDFLG* is approximately 50% larger than the impact of *HALPOR* (i.e., (0.722 – 0.468)/0.468 = 0.54) and almost 200% larger than the impact of *WGRCOR* (i.e., (0.722 – 0.246)/0.246 = 1.96). Both regression models have an R^2 value of 0.79 and thus can account for approximately 79% of the uncertainty in y. Standardized regression coefficients are a popular way of ranking variable importance in sampling-based sensitivity analysis and many examples of their use exist (e.g., Chan 1996, Helton et al. 1996, Hamby 1995, Ma et al. 1993, Ma and Ackerman 1993, Whiting et al. 1993).

8.3 Statistical Tests in Regression Analysis

Determination of the regression coefficients b_0, b_1, b_2, ..., b_n that constitute the elements of the vector **b** in Eq. (8.8) involves no statistics. Rather, as already indicated, this determination is based entirely on procedures involving minimization of functions and algebraic manipulations. If desired, formal statistical procedures can be used to indicate if these coefficients appear to be different from zero. However, such procedures are based on assumptions that are not satisfied in sampling-based sensitivity studies of deterministic models (i.e., models for which a given input always produces the same result), and thus the outcome of using formal statistical procedures to make assessments about the significance of individual coefficients or other entities in sampling-based sensitivity studies should be regarded simply as one form of guidance as to whether or not a model prediction appears to be affected by a particular model input.

In the usual construction of tests for the significance of regression coefficients, the relationship between the dependent and independent variables is assumed to be of the form

$$y = \beta_0 + \sum_{j=1}^{n} \beta_j x_j + \varepsilon, \tag{8.18}$$

where ε is normally distributed with mean 0 and standard deviation σ and characterizes the variation in y that is observed when y is repeatedly evaluated for $\mathbf{x} = [x_1, x_2, \ldots, x_n]$. Further, σ is assumed to be the same for all values of $\mathbf{x}$. It is the distributional assumptions involving ε that allows the construction of statistical tests for the coefficients $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$. These assumptions are not satisfied in sampling-based sensitivity studies with deterministic models because a given $\mathbf{x}$ always produces the same value for y.

Given the preceding assumptions involving ε, the relationship in Eq. (8.9) can be used in the development of tests to indicate if various of the β_j in Eq. (8.18) appear to be different from zero. For notational convenience, let

$$SS_{reg}(\beta_1, \beta_2, \ldots, \beta_n | \beta_0) = \sum_{k=1}^{m} (\hat{y}_k - \bar{y})^2 \tag{8.19}$$

when the vector $\mathbf{b}$ in Eq. (8.8), and hence the associated regression model, contains estimates for $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$. The preceding quantity is called the regression sum of squares and constitutes the part of the total sum of squares (i.e., the left-hand side of Eq. (8.9)) that can be explained by the regression model. More generally, if $\beta_1, \beta_2, \ldots, \beta_n$ are partitioned into vectors β_1 and β_2 where β_1 contains p_1 of the coefficients $\beta_1, \beta_2, \ldots, \beta_n$ and β_2 contains the remaining $p_2 = n - p_1$ coefficients, then

$$SS_{reg}(\beta_1, \beta_2, \ldots, \beta_n | \beta_0) = SS_{reg}(\beta_1 | \beta_2, \beta_0) + SS_{reg}(\beta_2 | \beta_0), \tag{8.20}$$

where $SS_{reg}(\beta_1|\beta_2, \beta_0)$ is the increase in the regression sum of squares that results from extending a regression model involving estimates for β_0 and the β_j's in β_2 to a regression model involving estimates for β_0 and the coefficients in β_1 and β_2.

Given the assumptions involving ε indicated in conjunction with Eq. (8.18), $SS_{reg}(\beta_1|\beta_2, \beta_0)$ can be used to test the hypothesis that $\beta_1 = \mathbf{0}$. In particular, if $\beta_1 = \mathbf{0}$ and the assumptions involving ε are satisfied, then

$$F = [SS_{reg}(\beta_1|\beta_2, \beta_0) / p_1] / \hat{s}^2 \tag{8.21}$$

can be regarded as a randomly sampled value from an F-distribution with $(p_1, m - n - 1) = (p_1, m - p_1 - p_2 - 1)$ degrees of freedom, where

$$\hat{s}^2 = \sum_{k=1}^{m} (y_k - \hat{y}_k)^2 / (m - n - 1) \tag{8.22}$$

is an approximation to σ^2 (see Sect. 3.4, Myers 1990, or any other standard text on regression analysis). The probability $prob_F(\tilde{F} > F|\eta_1, \eta_2)$ of exceeding an F-statistic value of F calculated with (η_1, η_2) degrees of freedom can be estimated by

$$prob_F(\tilde{F} > F|\eta_1, \eta_2) = I_\nu(\eta_2 / 2, \eta_1 / 2), \quad \nu = \eta_2 / (\eta_2 + \eta_1 F), \tag{8.23}$$

where $I_\nu(a, b)$ designates the incomplete beta function (p. 222, Press et al. 1992). Thus, under the assumption that $\beta_1 = \mathbf{0}$, the probability that a larger value for $SS_{reg}(\beta_1|\beta_2, \beta_0)$ would result from chance alone can be calculated and used to make an assessment as to whether or not it appears to be reasonable to reject the assumption that $\beta_1 = \mathbf{0}$, with this probability typically called the p-value or α-value for F and the corresponding vector β_1. Small p-values indicate that the observed value for F is unlikely to have occurred due to chance and thus suggest that $\beta_1 \neq \mathbf{0}$.

The statistic F in Eq. (8.21) can be used to test the hypothesis that

$$\beta = [\beta_1, \beta_2, \ldots, \beta_n] = \mathbf{0}. \tag{8.24}$$

In this case,

$$F = [SS_{reg}(\beta|\beta_0) / n] / \hat{s}^2 \tag{8.25}$$

can be regarded as a randomly sampled value from an F-distribution with $(n, m - n - 1)$ degrees of freedom. A small p-value for F suggests that $\beta \neq \mathbf{0}$.

Another important special case occurs when a single regression coefficient (i.e., β_j) is under consideration, with the result that $p_1 = 1$ and $p_2 = n - 1$ in Eq. (8.21). Then,

$$F = [SS_{reg}(\beta_j | \beta_2, \beta_0) / 1] / \hat{s}^2 \tag{8.26}$$

can be used to indicate if β_j appears to differ from zero given that estimates for β_0 and the coefficients in β_2 are

included in the regression model. Specifically, under the assumption that $\beta_j = 0$, F can be regarded as a randomly sampled value from an F-distribution with $(1, m-n-1)$ degrees of freedom. An equivalent test involving β_j can be based on the statistic

$$t = b_j / \hat{s}\sqrt{c_{jj}}, \tag{8.27}$$

where b_j is the estimated value for β_j, $\hat{s}$ is defined in Eq. (8.22), and c_{jj} is the j^{th} diagonal element of the matrix $(\mathbf{X}^T\mathbf{X})^{-1}$ in Eq. (8.8) (p. 98, Myers 1990). Under the assumption that $\beta_j = 0$, t can be regarded as a randomly sampled value from a t-distribution with $m - n - 1$ degrees of freedom. The probability $prob_t(|\tilde{t}| \geq |t| \,|\, m-n-1)$ of obtaining a value $\tilde{t}$ from the preceding distribution for which $|\tilde{t}|$ exceeds $|t|$ is given by

$$prob_t(|\tilde{t}| \geq |t| \,|\, m-n-1) = 1 - I_x[(m-n-1)/2,\ 1/2]$$
$$x = \frac{m-n-1}{m-n-1+t^2}, \tag{8.28}$$

where $I_x(a, b)$ designates the incomplete beta function (p. 222, Press et al. 1992). Thus, t as defined in Eq. (8.27) can also be used to test if an individual regression coefficient appears to be different from zero. The equality $F = t^2$ holds for F and t as defined in Eqs. (8.26) and (8.27). Further, identical significance results (i.e., p- or α-values) are produced by the use of F in conjunction with the relationship in Eq. (8.23) and the use of t in conjunction with the relationship in Eq. (8.28)

As already indicated, the distributional assumptions that lead to the p-values defined by Eqs. (8.23) and (8.28) are not satisfied in sampling-based sensitivity studies. However, these p-values still provide a useful criterion for assessing variable importance because they provide an indication of how viable the relationships between input and output variables would appear to be in a study in which the underlying distributional assumptions were satisfied.

As an illustration, results of a formal statistical analysis of the regression models in Eqs. (8.16) and (8.17) are presented in Table 8.1, with the coefficients in these models appearing in the columns labeled "Regression Coefficient" and "Standardized Regr Coef," respectively. The p-value for the regression model containing all three variables (Footnote e, Table 8.1) is less than 10^{-4}, as are the p-values for adding individual variables to the regression model (Footnote n, Table 8.1). Thus, in a study in which the necessary distributional assumptions were satisfied (see Eq. (8.18)), the implication would be that *WMICDFLG*, *HALPOR* and *WGRCOR* have significant influences (i.e., nonzero regression coefficients) on y = *E0:WAS_PRES*. The p-values for the individual variables (Footnote n, Table 8.1) are more useful from a sensitivity analysis perspective than the p-value for all three variables (Footnote e, Table 8.1) as they indicate whether or not individual variables appear to affect y. In contrast, the p-value for the variables collectively only indicates that at least one of the variables appears to affect y.

The regression analysis summarized in Eq. (8.16), Eq. (8.17) and Table 8.1 only involves the variables *WMICDFLG*, *HALPOR* and *WGRCOR*, with these variables selected for illustrative purposes on the basis of *a priori* knowledge that they had identifiable effects on y. As a result, these variables result in regression models with small p-values. In a sensitivity analysis with no *a priori* knowledge, all of the variables in Table 3.1 would have to be investigated for their effects on y. This implies the construction of a regression model with all 31 variables from Table 3.1, with the outcome of this construction summarized in Table 8.2. Actually, the regression model in Table 8.2 only involves 24 variables because (i) the variables *ANHCOMP* and *HALCOMP* were not used in the construction of the model due to the rank correlations of –0.99 assumed to exist within the variable pairs (*ANHCOMP*, *ANHPRM*) and (*HALCOMP*, *HALPRM*) (see Sect. 8.7) and (ii) the variables *BHPRM*, *BPCOMP*, *BPINTPRS*, *BPPRM* and *BPVOL* were not used because they were not involved in the calculation of the dependent variable under consideration. Of the 24 variables, six have p-values less than 0.02 and thus appear to affect y (i.e., *WMICDFLG*, *HALPOR*, *WGRCOR*, *ANHPRM*, *SHRGSSAT*, *SALPRES*). The remaining variables have larger p-values, and thus the regression analysis does not indicate an effect for these variables. However, it is important to realize that the failure of a regression analysis to identify an effect for a variable does not necessarily imply that no effect exists. In particular, the regression model is based on identifying a linear relationship and can completely miss other types of relationships (see Sects. 8.8, 8.9).

The results presented in Table 8.2 are rather unwieldy, with much of the table involving variables that appear to have no effect on y. Stepwise regression analysis provides a more informative and less cumbersome procedure for constructing and presenting regression models and will be described in Sect. 8.5.

Table 8.1. Summary of Regression Analysis for y = *E0:WAS_PRES* at 10,000 yr, x_1 = *WMICDFLG*, x_2 = *HALPOR* and x_3 = *WGRCOR*

Source	DofF[a]	SS[b]	MS[c]	F[d]	SIGNIF[e]
Regression	3	1.9009E+15	6.3365E+14	3.7643E+02	0.0000
Residual	296	4.9827E+14	1.6833E+12		
Total	299	2.3992E+15			

R-Square[f] = 0.79232 Intercept[g] = 5.7274E+06

Variable[h]	Regression[i] Coefficient	Standardized[j] Regr Coeff	Partial[k] SSQ	T-Test[l] Values	R-Square[m] Deletes	Alpha[n] Hats
WMICDFLG	2.4625E+06	7.2201E–01	1.2482E+15	2.7231E+01	2.7206E–01	0.0000E+00
HALPOR	1.5529E+08	4.6809E–01	5.2479E+14	1.7657E+01	5.7359E–01	–4.4409E–16[o]
WGRCOR	1.5210E+20	2.4649E–01	1.4559E+14	9.3000E+00	7.3164E–01	–4.4409E–16

[a] Degrees of freedom associated with regression (SS_{reg}), residual (SS_{res}) and total (SS_{tot}) sums of squares; see Eqs. (8.9), (8.10).
[b] Regression (SS_{reg}), residual (SS_{res}) and total (SS_{tot}) sums of squares.
[c] Mean sums of squares (SS_{reg}/n, $SS_{res}/(m-n-1)$, where estimates for $\beta_1, \beta_2, \ldots, \beta_n$ are obtained from m observations).
[d] F-statistic ($[SS_{reg}/n]/[SS_{res}/(m-n-1)]$); see Eq. (8.25).
[e] p- or α-value for F; see Eq. (8.23).
[f] R^2 value for regression model with estimates for $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$; see Eq. (8.11).
[g] Estimate for β_0.
[h] Variables in regression model ($x_1, x_2, \ldots, x_n$).
[i] Regression coefficients ($b_1, b_2, \ldots, b_n$); see Eq. (8.8).
[j] Standardized regression coefficients; see Eq. (8.15).
[k] Partial sum of squares for variable (i.e., x_j) in row ($SS_{reg}(\beta_j|\beta_2, \beta_0)$); see Eqs. (8.20), (8.26).
[l] t-statistic for variable in row; see Eq. (8.27).
[m] For variable (x_j) in row, R^2 value for regression model constructed with x_i, $i = 1, 2, \ldots, n$ and $i \neq j$.
[n] For variable (x_j) in row, p- or α-value for addition of x_j to regression model containing x_i, $i = 1, 2, \ldots, n$ and $i \neq j$; use of F-statistic or t-statistic produces same value; see Eqs. (8.23) and (8.26) for F-statistic and Eqs. (8.27) and (8.28) for t-statistic.
[o] Negative values result from numerical errors in the calculation of very small p-values with the STEPWISE program (Iman et al. 1980).

8.4 Correlation and Partial Correlation

The ideas of correlation and partial correlation are useful concepts that often appear in sampling-based uncertainty/sensitivity studies. For a sequence of observations (x_k, y_k), $k = 1, \ldots, m$, the (sample or Pearson) correlation r_{xy} between x and y is defined by

$$r_{xy} = \frac{\sum_{k=1}^{m} (x_k - \bar{x})(y_k - \bar{y})}{\left[\sum_{k=1}^{m} (x_k - \bar{x})^2\right]^{1/2} \left[\sum_{k=1}^{m} (y_k - \bar{y})^2\right]^{1/2}} \qquad (8.29)$$

where $\bar{x}$ and $\bar{y}$ are defined in conjunction with Eq. (8.15). The correlation coefficient (CC) r_{xy} provides a measure of the linear relationship between x and y. For the regression model defined by Eq. (8.8), the R^2 value in Eq. (8.11) is equal to the square of the correlation between y and $\hat{y}$ (*i.e.*, $R^2 = r^2_{y\hat{y}}$) (p. 91, Draper and Smith 1981).

The nature of r_{xy} is perhaps most readily understood by considering the regression

$$\hat{y} = b_0 + b_1 x . \qquad (8.30)$$

The definition of r_{xy} in Eq. (8.29) is equivalent to the definition

$$r_{xy} = \text{sign}(b_1)(R^2)^{1/2}, \qquad (8.31)$$

Table 8.2. Summary of Regression Analysis for y = *E0:WAS_PRES* at 10,000 yr and x_1, x_2, ..., x_{24} Corresponding to *ANHBCEXP*, *ANHBCVGP*, *ANRBRSAT*, *ANHPRM*, *ANRGSSAT*, *HALPOR*, *HALPRM*, *SALPRES*, *SHBCEXP*, *SHPRMASP*, *SHPRMCLY*, *SHPRMCON*, *SHPRMDRZ*, *SHPRMHAL*, *SHRBRSAT*, *SHRGSSAT*, *WASTWICK*, *WFBETCEL*, *WGRCOR*, *WGRMICH*, *WGRMICI*, *WMICDFLG*, *WRBRNSAT* and *WRGSSAT* (see Table 8.1 for description of table structure)

Source	DofF	SS	MS	F	SIGNIF
Regression	24	1.9921E+15	8.3003E+13	5.6063E+01	0.0000
Residual	275	4.0714E+14	1.4805E+12		
Total	299	2.3992E+15			

R-Square = 0.83030 Intercept = 1.2896E+07

Variable	Regression Coefficient	Standardized Regr Coeff	Partial SSQ	T-Test Values	R-Square Deletes	Alpha Hats[a]
WMICDFLG	2.4669E+06	7.2329E–01	1.2002E+15	2.8472E+01	3.3007E–01	2.7828E–08
HALPOR	1.5429E+08	4.6510E–01	5.1332E+14	1.8620E+01	6.1635E–01	2.7828E–08
WGRCOR	1.5156E+20	2.4561E–01	1.4349E+14	9.8446E+00	7.7050E–01	2.7828E–08
ANHPRM	5.5924E+05	1.2774E–01	3.8910E+13	5.1266E+00	8.1408E–01	5.8498E–07
SHRGSSAT	1.7177E+06	7.0177E–02	1.1729E+13	2.8147E+00	8.2541E–01	5.2352E–03
SALPRES	2.1946E–01	6.3855E–02	9.6907E+12	2.5584E+00	8.2626E–01	1.1051E–02
WASTWICK	4.9174E+05	5.0273E–02	6.0154E+12	2.0157E+00	8.2779E–01	4.4804E–02
HALPRM	1.6369E+05	5.0099E–02	5.9533E+12	2.0053E+00	8.2782E–01	4.5914E–02
WGRMICH	–3.3997E+14	–4.3972E–02	4.5935E+12	–1.7614E+00	8.2839E–01	7.9278E–02
SHPRMCLY	–1.5252E+05	–4.1991E–02	4.2037E+12	–1.6850E+00	8.2855E–01	9.3118E–02
ANHBCEXP	1.6070E+06	3.3745E–02	2.7011E+12	1.3507E+00	8.2918E–01	1.7790E–01
ANHBCVGP	6.0999E+04	3.2355E–02	2.4143E+12	1.2770E+00	8.2930E–01	2.0269E–01
WGRMICI	2.8456E+13	2.6713E–02	1.7003E+12	1.0717E+00	8.2959E–01	2.8481E–01
ANRBRSAT	–2.0139E+06	–1.9366E–02	8.9467E+11	–7.7736E–01	8.2993E–01	4.3761E–01
SHPRMCON	–5.2400E+04	–1.3969E–02	4.6597E+11	–5.6101E–01	8.3011E–01	5.7525E–01
SHPRNHAL	–1.2571E+05	–1.2830E–02	3.9130E+11	–5.1410E–01	8.3014E–01	6.0759E–01
SHPRMSAP	5.2519E+04	1.1577E–02	3.1972E+11	4.6471E–01	8.3017E–01	6.4251E–01
WRGSSAT	7.8406E+05	1.2014E–02	3.4357E+11	4.8173E–01	8.3016E–01	6.3038E–01
SHBCEXP	1.1063E+04	1.0169E–02	2.4150E+11	4.0388E–01	8.3020E–01	6.8662E–01
ANRGSSAT	7.4311E+05	8.8597E–03	1.8755E+11	3.5592E–01	8.3022E–01	7.2217E–01
WFBETCEL	–8.1238E+04	–8.2986E–03	1.6325E+11	–3.3206E–01	8.3023E–01	7.4010E–01
WRBRNSAT	–1.0212E+05	–5.7539E–03	7.8584E+10	–2.3039E–01	8.3027E–01	8.1796E–01
SHPRMDRZ	1.6592E+03	3.6624E–04	3.2009E+08	1.4704E–02	8.3030E–01	9.8828E–01
SHRBRSAT	–2.9090E+03	–1.8048E–04	7.6305E+07	–7.1791E–03	8.3030E–01	9.9428E–01

[a] Identical values result from lack of resolution in algorithm used in the calculation of very small p-values in the STEPWISE program (Iman et al. 1980).

where $\text{sign}(b_1) = 1$ if $b_1 \geq 0$, $\text{sign}(b_1) = -1$ if $b_1 < 0$, and R^2 is the coefficient of determination that results from regressing y on x. With respect to interpretation, the CC r_{xy} provides a measure of the linear relationship between x and y, and the regression coefficient b_1 characterizes the effect that a unit change in x will have on y.

The definition of r_{xy} in Eq. (8.29) is also equivalent to the definition

$$r_{xy} = b_1 \hat{s}_1 / \hat{s}\,, \tag{8.32}$$

where $\hat{s}_1$ and $\hat{s}$ are defined in conjunction with Eq. (8.15) with x assumed to correspond to x_1. Thus, r_{xy} is also equal to the standardized regression coefficient that results from regressing y on x. Hence, r_{xy} can be viewed as characterizing the effect that changing x by a fixed fraction of its standard deviation will have on y, with this effect being measured relative to the standard deviation of y. The CC can also be viewed as a parameter in a joint normal distribution involving x and y (Sect. 2.13, Myers 1990); however, this interpretation is not as intuitively appealing as the two preceding interpretations involving the regression model in Eq. (8.30). Further, x and y typically do not have normal distributions in sampling-based sensitivity analyses.

When more than one input variable is under consideration, partial correlation coefficients (PCCs) can be used to provide a measure of the linear relationships between the output variable y and the individual input variables. The PCC between an individual variable x_j and y is obtained from the use of a sequence of regression models. First, the following two regression models are constructed:

$$\hat{x}_j = c_0 + \sum_{\substack{p=1 \\ p \neq j}}^{n} c_p x_p \text{ and } \hat{y} = b_0 + \sum_{\substack{p=1 \\ p \neq j}}^{n} b_p x_p\,. \tag{8.33}$$

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 $p_{x_j y}$ between x_j and y is the CC between $x_j - \hat{x}_j$ and $y - \hat{y}$. Thus, the PCC provides a measure of the linear relationship between x_j and y with the linear effects of the other variables removed. The preceding provides a rather intuitive development of what a PCC is. A formal development of PCCs is provided by Iman et al. (1985).

The PCC characterizes the strength of the linear relationship between two variables after a correction has been made for the linear effects of the other variables in the analysis, and the SRC characterizes the effect on the output variable that results from perturbing an input variable 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 provides a measure of variable importance that tends to exclude the effects of other variables, the assumed distribution for the particular input variable under consideration, and the magnitude of the impact of an input variable on an output variable. In contrast, the value for an SRC is more influenced by the distribution assigned to an input variable and the impact that this variable has on an output variable.

The following relationship exists between $p_{x_j y}$ and the SRC $c_j = b_j \hat{s}_j / \hat{s}$ in Eq. (8.15):

$$p_{x_j y} = c_j [(1 - R_j^2) / (1 - R_y^2)]^{1/2}\,, \tag{8.34}$$

where R_j^2 is the R^2 value that results from regressing x_j on y and the x_i, $i = 1, 2, \ldots, n$ with $i \neq j$, and R_y^2 is the R^2 value that results from regressing y on the x_i, $i = 1, 2, \ldots, n$ (Eq. (1), Iman et al. 1995). If the x_i are orthogonal, then

$$R_y^2 = \sum_{i=1}^{n} R_i^2 = \sum_{i=1}^{n} r_{x_i y}^2 = \sum_{i=1}^{n} c_i^2\,, \tag{8.35}$$

with the first equality following from Eq. (8.14), and the second and third equalities following from Eqs. (8.31) and (8.32). Thus,

$$\begin{aligned} p_{x_j y} &= c_j \left[(1 - c_j^2) / \left(1 - \sum_{i=1}^{n} c_i^2 \right) \right]^{1/2} \\ &= r_{x_j y} \left[(1 - r_{x_j y}^2) / \left(1 - \sum_{i=1}^{n} r_{x_i y}^2 \right) \right]^{1/2}\,. \end{aligned} \tag{8.36}$$

Because of the inequality

$$b(1-b^2)^{1/2} > a(1-a^2)^{1/2} \tag{8.37}$$

for $a^2 + b^2 < 1$ and $0 \le a < b$ (see Fig. 7, Kleijnen and Helton 1999a), an ordering of variable importance based on $|p_{x_j y}|$, $|c_j|$ or $|r_{x_j y}|$ produces the same results when the x_j are orthogonal; further, the values for c_j and $r_{x_j y}$ will be the same and generally different from $p_{x_j y}$.

Many output variables are functions of time or location. A useful way to present sensitivity results for such variables is with plots of PCCs or SRCs. An example of such a presentation for the pressure curves in Fig. 7.5 is given in Fig. 8.3, which displays two sets of curves. The left set contains SRCs plotted as a function of time; the right set contains PCCs plotted in a similar manner. For both sets of curves, the dependent variables are pressures at fixed times, and each curve displays the values of SRCs or PCCs relating these pressures to a single input variable as a function of time. Many additional examples of the use of PCCs in sampling-based sensitivity analysis also exist (e.g., Helton et al. 1996, Hamby 1995, Whiting et al. 1993, Breshears et al. 1992).

Determination of CCs and PCCs involves no statistical assumptions. However, as previously discussed for regression coefficients in Sect. 8.3, statistical tests can be performed conditional on suitable assumptions. For example,

$$t = r_{xy}(m-2)^{1/2} / (1-r_{xy}^2)^{1/2} \tag{8.38}$$

can be regarded as a random sample from a t-distribution with $m - 2$ degrees of freedom when (i) r_{xy} is calculated from the observations (x_k, y_k), $k = 1, 2, \ldots, m$, and (ii) x and y are uncorrelated and have a bivariate normal distribution (p. 631, Press et al. 1992). Then, the probability of observing a stronger correlation due to chance variation is given by the relationship in Eq. (8.28). The preceding test is identical to the test involving the t-statistic described in Sect. 8.3 for the significance of b_1 in Eq. (8.30) (p. 70, Myers 1990). Further,

$$z = r_{xy}\sqrt{m} \tag{8.39}$$

is distributed approximately normally with mean 0 and standard deviation 1 when x and y are uncorrelated, x and y have enough convergent moments (i.e., the tails of their distributions die off sufficiently rapidly), and m is sufficiently large (p. 631, Press et al. 1992). Given the preceding assumptions, the probability

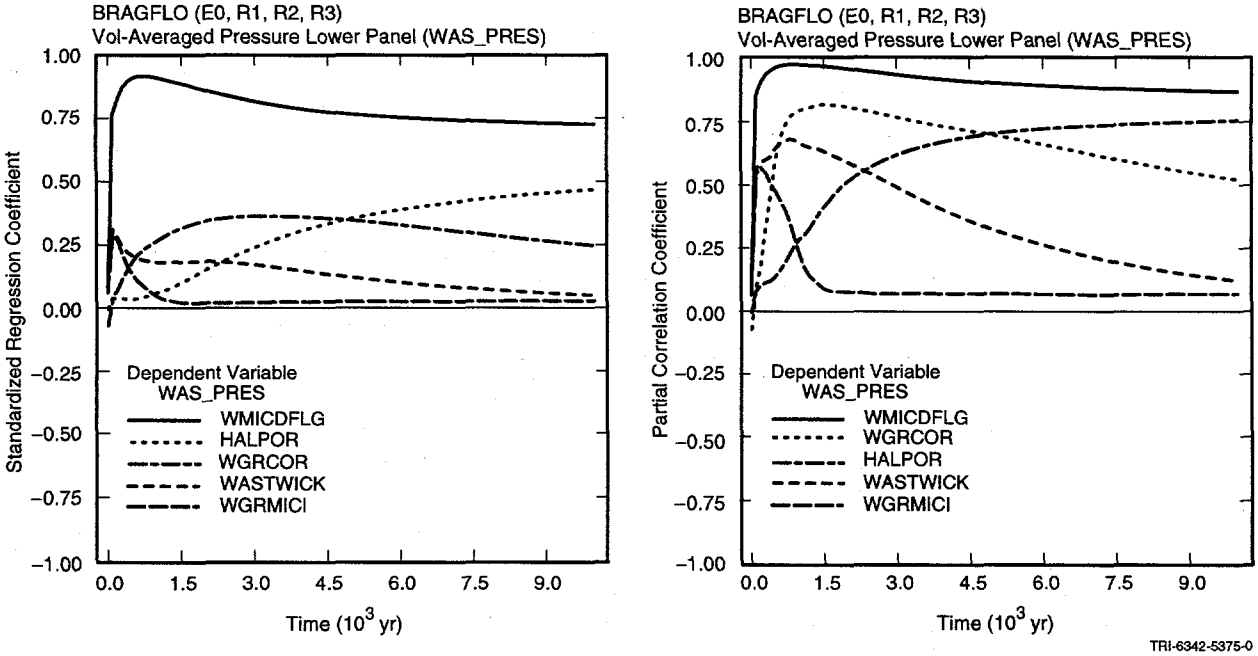


Fig. 8.3. Standardized regression coefficients (SRCs) and partial correlation coefficients (PCCs) for five variables having the largest PCCs, in absolute value, with pressure (Pa) in lower waste panel under undisturbed conditions (i.e., y = *E0:WAS_PRES*).

$prob_n(|\tilde{r}_{xy}|>|r_{xy}|)$ of obtaining a value $\tilde{r}_{xy}$ for which $|\tilde{r}_{xy}|$ exceeds $|r_{xy}|$ is given by

$$prob_n\left(|\tilde{r}_{xy}|>|r_{xy}|\right) = erfc\left(|r_{xy}|\sqrt{m}/\sqrt{2}\right), \quad (8.40)$$

where *erfc* is the complementary error function (i.e., $erfc(x) = (2/\sqrt{\pi})\int_x^\infty exp(-t^2)dt$) (p. 631, Press et al. 1992). Significance results obtained with the statistics in Eqs. (8.38) and (8.39) converge as m increases. Related significance results can also be defined for PCCs (Quade 1989).

As an example, CCs, SRCs and PCCs for y = *E0:WAS_PRES* at 10,000 yr are shown in Table 8.3. Of the 24 variables under consideration, 5 have CCs with p-values less than 0.1. The CCs and SRCs have similar values, with equality failing to exist because of small correlations between the 24 variables in the sample (see Table 5.1). The PCCs tend to be larger than the CCs and SRCs. Because PCCs provide a measure of the strength of linear relationships after corrections have been made for the effects of other variables, large PCCs have the potential to produce misleading impressions of variable importance; therefore, care should be exercised in the use and interpretation of PCCs. In particular, a large PCC does not necessarily imply that the corresponding input variable makes a large contribution to the uncertainty in the output variable under consideration. However, when the sampled variable values are independent (i.e., orthogonal), use of CCs, SRCs and PCCs will produce identical rankings of variable importance as previously noted. The effect of correlations within the sample can be seen in Table 8.3, with *SALPRES* ranked 5 with CCs and 6 with SRCs and PCCs.

8.5 Stepwise Regression Analysis

When many input variables are involved, the direct construction of a regression model containing all input variables as shown in Eq. (8.3) and illustrated in Table 8.2 may not be the best approach for several reasons. First, the large number of variables makes the regression model tedious to examine and unwieldy to display. Second, only a relatively small number of input variables typically has an impact on the output variable. As a result, there is no reason to include the remaining variables in the regression model. Third, correlated variables result in unstable regression coefficients (i.e., coefficients whose values are sensitive to the specific variables included in the regression model; see Sect. 8.7). When this occurs, the regression coefficients in a model containing all the input variables can give a misleading representation of variable importance. As a side point, if several input variables are highly correlated, consideration should be given to either removing all but one of the correlated variables or transforming the variables to correct for (i.e., remove) the correlations between them. Fourth, an overfitting of the data can result when variables are arbitrarily forced into the regression model. This phenomenon occurs when the regression model attempts to match the predictions associated with individual sample elements rather than match the trends shown by the sample elements collectively.

Table 8.3. Correlation Coefficients (CCs), Standardized Regression Coefficients (SRCs) and Partial Correlation Coefficients (PCCs) for y = *E0:WAS_PRES* at 10,000 yr

Variable	CC[b]			SRC[c]		PCC[d]	
Name[a]	p-Val	Rank	Value	Rank	Value	Rank	Value
WMICDFLG	0.0000	1.0	0.7124	1.0	0.7234	1.0	0.8642
HALPOR	0.0000	2.0	0.4483	2.0	0.4651	2.0	0.7469
WGRCOR	0.0000	3.0	0.2762	3.0	0.2460	3.0	0.5113
ANHPRM	0.0241	4.0	0.1302	4.0	0.1277	4.0	0.2953
SALPRES	0.0855	5.0	0.0993	6.0	0.0639	6.0	0.1526

[a] Variables for which CC with y has a p-value less than 0.1; variables ordered by p-values for CCs.
[b] p-value for CC, variable rank based on p-value for CC, and value of CC.
[c] Variable rank based on SRC and value for SRC for regression model containing 24 variables used in Table 8.2.
[d] Variable rank based on PCC and value for PCC calculated for 24 variables used in Table 8.2.

Stepwise regression analysis provides an alternative to constructing a regression model containing all the input variables. With this approach, a sequence of regression models is constructed. The first regression model contains the single input variable that has the largest impact on the uncertainty in the output variable (i.e., the input variable that has the largest correlation with the output variable y). The second regression model contains the two input variables that have the largest impact on the output variable: the input variable from the first step plus whichever of the remaining variables has the largest impact on the uncertainty not accounted for by the first variable (i.e., the input variable that has the largest correlation with the uncertainty in y that cannot be accounted for by the first variable). The third regression model contains the three input variables that have the largest impact on the output variable: the two input variables from the second step plus whichever of the remaining variables has the largest impact on the uncertainty not accounted for by the first two variables (i.e., the input variable that has the largest correlation with the uncertainty in y that cannot be accounted for by the first two variables). Additional models in the sequence are defined in the same manner until a point is reached at which further models are unable to meaningfully increase the amount of the uncertainty in the output variable that can be accounted for. Further, at each step of the process, the possibility exists for an already selected variable to be dropped out if this variable no longer has a significant impact on the amount of uncertainty in the output variable that can be accounted for by the regression model; this only occurs when correlations exist between the input variables.

Several aspects of stepwise regression analysis provide insights on the importance of the individual variables. First, the order in which the variables are selected in the stepwise procedure provides an indication of their importance, with the most important variable being selected first, the next most important variable being selected second, and so on. Second, the R^2 values (see Eq. (8.11)) at successive steps of the analysis also provide a measure of variable importance by indicating how much of the uncertainty in the dependent variable can be accounted for by all variables selected through each step. When the input variables are uncorrelated, the differences in the R^2 values for the regression models constructed at successive steps equals the fractions of the total uncertainty in the output variable that can be accounted for by the individual input variables being added at each step (see Eq. (8.14)). Third, the absolute values of the SRCs (see Eq. (8.15)) in the individual regression models provide an indication of variable importance. Further, the sign of an SRC indicates whether the input and output variable tend to increase and decrease together (a positive coefficient) or tend to move in opposite directions (a negative coefficient).

An important situation occurs when the input variables are uncorrelated. In this case, orderings of variable importance based on order of entry into the regression model, size of the R^2 values attributable to the individual variables, the absolute values of the SRCs, the absolute values of correlation coefficients, and the absolute values of the PCCs are the same. In situations where the input variables are believed to be uncorrelated, one of the important applications of the previously discussed restricted pairing technique of Iman and Conover (Sect. 5.5) is to ensure that the correlations between variables within a Latin hypercube or random sample are indeed close to zero. When variables are correlated, care must be used in the interpretation of the results of a regression analysis since the regression coefficients can change in ways that are basically unrelated to the importance of the individual variables as correlated variables are added to and deleted from the regression model (see Sect. 8.7 for an example of the effects of correlated variables on the outcomes of a regression analysis).

When the stepwise technique is used to construct a regression model, it is necessary to have some criterion to stop the construction process. When there are many independent variables, there is usually no reason to let the construction process continue until all the variables have been used. It is also necessary to have some criterion to determine when a variable is no longer needed and thus can be dropped from the regression model. As indicated earlier, this latter situation only occurs when the input variables are correlated. The usual criterion for making the preceding decisions is based on whether or not the regression coefficient associated with an input variable appears to be significantly different from zero. Specifically, an F-test or t-test is used to determine the probability that a regression coefficient with absolute value as large as or larger than the one constructed in the analysis would be obtained if, in reality, there was no relationship between the input and output variable, and, as a result, the apparent relationship that led to the constructed regression coefficient was due entirely to chance (see Eqs. (8.26), (8.27) and associated text). Sensitivity studies often use an α-value of 0.01 or 0.02 to add a variable to a regression model and a somewhat larger value to drop a variable from the model.

As models involving more variables are developed in a stepwise regression analysis, the possibility exists of overfitting the data. Overfitting occurs when the

regression model in essence "chases" the individual observations rather than following an overall pattern in the data. For example, it is possible to obtain a good fit to a set of points by using a polynomial of high degree. However, in doing so, it is possible to overfit the data and produce a spurious model that makes poor predictions.

To protect against overfit, the Predicted Error Sum of Squares (PRESS) criterion can be used to determine the adequacy of a regression model (Allen 1971). For a regression model containing q variables and constructed from m observations, PRESS is computed in the following manner. For $k = 1,2,...,m$, the kth observation is deleted from the original set of m observations and then a regression model containing the original q variables is constructed from the remaining $m - 1$ observations. With this new regression model, the value $\hat{y}_q(k)$ is estimated for the deleted observation y_k. Then, PRESS is defined from the preceding predictions and the m original observations by

$$PRESS_q = \sum_{k=1}^{m} \left(y_k - \hat{y}_q(k)\right)^2. \qquad (8.41)$$

The regression model having the smallest PRESS value is preferred when choosing between two competing models, as this is an indication of how well the basic pattern of the data has been matched versus an overfit or an underfit. In particular, PRESS values will decrease in size as additional variables are added to the regression model without an overfitting of the data (i.e., $PRESS_q > PRESS_{q+1}$), with an increase in the PRESS values (i.e., $PRESS_q < PRESS_{q+1}$) indicating an overfitting of the data. In addition to PRESS, there are also a number of other diagnostic tools that can be used to investigate the adequacy of regression models (Cook and Weisberg 1982, Belsley et al. 1980).

It is important to use scatterplots, PRESS values and other procedures to examine the reasonableness of regression models. This is especially true when regression models are used for sensitivity analysis. Such analyses often involve many input variables and large uncertainties in these variables. The appearance of spurious patterns is a possibility that must be checked for.

An example stepwise regression analysis follows for the variable y = *E0:WAS_PRES* previously analyzed with the regression model presented in Table 8.2. The first step selects the input variable x_j that has the largest impact on the output variable y. Specifically, this is defined to be the variable that has the largest correlation, in absolute value, with y (see Eqs. (8.29) and (8.31)). Thus, it is necessary to calculate the correlations between y and each of the 24 input variables under consideration. For illustration, Table 8.4 shows the 7 × 7 correlation matrix for y and the six input variables ultimately selected in the stepwise regression, although the full correlation matrix would actually be (24 + 1) × (24 + 1). Each element in the correlation matrix is the correlation between the variables in the corresponding row and column. As examination of the correlation matrix in Table 8.4 shows, *WMICDFLG* has the highest correlation with waste pressure, which is denoted by *WAS_PRES*. Thus, the first step in the analysis selects the variable *WMICDFLG*. Here and elsewhere in the stepwise procedure, the selection of variables to enter the regression model could equivalently be made on the basis of F-test or t-test values as defined in Eqs. (8.26) and (8.27). A regression model relating y to *WMICDFLG* is then developed as shown in Eq. (8.8) with $n = 1$ and $m = 300$. The resultant regression model is

$$\hat{y} = 8.94 \times 10^6 + 2.43 \times 10^6 \bullet WMICDFLG, \qquad (8.42)$$

which has an R^2 value of 0.508, an α-value of 0.0000, an SRC of 0.712 and a PRESS value of 1.20×10^{15}. This model is summarized as Step 1 in Table 8.5.

The second step selects the input variable x_j that has the largest impact on the uncertainty in the output variable y that cannot be accounted by *WMICDFLG*, the variable selected in the first step. This selection is made by defining a new variable

$$\begin{aligned}\tilde{y} &= y - \hat{y} \\ &= y - (8.94 \times 10^6 + 2.43 \times 10^6 \bullet WMICDFLG), \end{aligned} \qquad (8.43)$$

where $\hat{y}$ is defined in Eq. (8.42), and then calculating the correlations between $\tilde{y}$ and the remaining variables. The variable with the largest correlation, in absolute value, with $\tilde{y}$ is selected as the second variable for inclusion in the model. In this example, the selected variable is *HALPOR*. The regression model at this step will thus involve the two variables *WMICDFLG* and *HALPOR* and is constructed as shown in Eq. (8.8) with $n = 2$ and $m = 300$. The resultant regression model is

$$\begin{aligned}\hat{y} = {} & 6.89 \times 10^6 + 2.49 \times 10^6 \bullet WMICDFLG \\ & + 1.57 \times 10^8 \bullet HALPOR. \end{aligned} \qquad (8.44)$$

This model is summarized as Step 2 in Table 8.5.

Table 8.4. Correlation Matrix for Variables Selected in Stepwise Regression Analysis for Pressure in the Repository at 10,000 yr Under Undisturbed Conditions (i.e., y = *E0:WAS_PRES* at 10,000 yr in Fig. 7.5)

WMICDFLG	1.0000						
HALPOR	–0.0348	1.0000					
WGRCOR	0.0272	0.0216	1.0000				
ANHPRM	0.0008	–0.0039	0.0130	1.0000			
SHRGSSAT	–0.0026	0.0395	–0.0171	–0.0042	1.0000		
SALPRES	0.0560	–0.0072	0.0010	–0.0117	0.0061	1.0000	
E0:WAS_PRES	0.7124	0.4483	0.2762	0.1303	0.0820	0.0993	1.0000
	WMICDFLG	*HALPOR*	*WGRCOR*	*ANHPRM*	*SHRGSSAT*	*SALPRES*	*E0:WAS_PRES*

Table 8.5. Results of Stepwise Regression Analysis for Pressure in the Repository at 10,000 yr Under Undisturbed Conditions (i.e., y = *E0:WAS_PRES* at 10,000 yr in Fig. 7.5)

Step[a]	Variables[b]	SRC[c]	α-Values[d]	R^2 Values[e]	PRESS[f]
1	*WMICDFLG*	0.712	0.0000	0.508	1.20×10^{15}
2	*WMICDFLG*	0.729	0.0000	0.732	6.59×10^{14}
	HALPOR	0.474	0.0000		
3	*WMICDFLG*	0.722	0.0000	0.792	5.14×10^{14}
	HALPOR	0.468	0.0000		
	WGRCOR	0.246	0.0000		
4	*WMICDFLG*	0.722	0.0000	0.809	4.79×10^{14}
	HALPOR	0.469	0.0000		
	WGRCOR	0.245	0.0000		
	ANHPRM	0.128	0.0000		
5	*WMICDFLG*	0.722	0.0000	0.814	4.70×10^{14}
	HALPOR	0.466	0.0000		
	WGRCOR	0.246	0.0000		
	ANHPRM	0.129	0.0000		
	SHRGSSAT	0.070	0.0056		
6	*WMICDFLG*	0.718	0.0000	0.818	4.63×10^{14}
	HALPOR	0.466	0.0000		
	WGRCOR	0.246	0.0000		
	ANHPRM	0.129	0.0000		
	SHRGSSAT	0.070	0.0055		
	SALPRES	0.063	0.0012		

[a] Steps in the analysis.
[b] Variables selected at each step.
[c] Standardized regression coefficients (SRCs) for variables in the regression model at each step; see Eq. (8.15).
[d] p- or α-values for variables in the regression model at each step; see Eqs. (8.26), (8.27).
[e] R^2 value for the regression model at each step; see Eq. (8.11).
[f] Predicted error sum of squares (PRESS) value for the regression model at each step; see Eq. (8.41).

The third step selects the input variable x_j that has the largest impact on the uncertainty in the output variable y that cannot be accounted for by *WMICDFLG* and *HALPOR*, the two variables from the second step. This selection is made by defining a new variable

$$\begin{aligned}\tilde{y} &= y - \hat{y} \\ &= y - (6.89 \times 10^6 + 2.49 \times 10^6 \bullet \textit{WMICDFLG} \\ &\quad + 1.57 \times 10^8 \bullet \textit{HALPOR}),\end{aligned} \tag{8.45}$$

where $\hat{y}$ is defined in Eq. (8.44). The variable with the largest correlation, in absolute value, with $\tilde{y}$ is selected as the third variable for inclusion in the model. In this example, the selected variable is *WGRCOR*. The regression model for this step will thus involve the three variables *WMICDFLG*, *HALPOR* and *WGRCOR*. The resultant regression model is summarized as Step 3 in Table 8.5.

As shown in Table 8.5, the stepwise procedure then continues in the same manner through a total of six steps, until no more variables can be found with an α-value less than 0.02. At this point, the stepwise procedure stops.

At each step, the stepwise procedure also checks to see if any variable selected at a prior step now has an α-value that exceeds a specified level, which is 0.05 in this analysis. If such a situation occurs, the variable will be dropped from the analysis, with the possibility that it may be reselected at a later step as other variables are added and deleted from the model. This type of behavior only occurs when there are correlations between the input variables. As shown in the correlation matrix in Table 8.4, the restricted pairing technique has been successful in keeping the correlations between the input variables close to zero. Thus, no variables meet the criterion to be dropped from the regression model once they have been selected at a prior step.

Another result of this lack of correlation is that the regression coefficients do not change significantly as additional variables are added to the regression model. As examination of Table 8.5 shows, the regression coefficients for a specific variable are essentially the same in all regression models containing that variable. Further, as indicated in Eq. (8.14), the R^2 values obtained for successive models can be subtracted to obtain the contribution to the uncertainty in y due to the newly added variable. Thus, for example, *WMICDFLG* accounts for approximately 51% of the uncertainty in y (i.e., $R^2 = 0.508$), while *WMICDFLG* and *HALPOR* together account for approximately 73% of the uncertainty (i.e., $R^2 = 0.732$). As a result, *HALPOR* by itself accounts for approximately 73% – 51% = 22% of the uncertainty in y. Similar results hold for the other variables selected in the analysis.

Table 8.5 also reports the PRESS values for the regression models obtained at the individual steps in the analysis. A decreasing sequence of PRESS values indicates that the regression models are not overfitting the data on which they are based. An increase in the PRESS values suggests that a model is overfitting the data, and thus that the stepwise procedure should probably be stopped at the preceding step. As shown by the decreasing PRESS values in Table 8.5, the regression models in this analysis are probably not overfitting the data from which they were constructed.

Typically, a certain amount of discretion is involved in selecting the exact point at which to stop a stepwise regression analysis. Certainly, α-values and the behavior of PRESS values provide two criteria to consider in selecting a stopping point. Other criteria include the changes in the R^2 values that take place as additional variables are added to the regression models and whether or not spurious variables are starting to enter the regression models. When only very small changes in R^2 values are taking place (e.g., ≤ 0.01), there is often little reason to continue the stepwise process. When α-values approach or exceed 0.01 and a large number of input variables are being considered, it is fairly common to start getting spurious variables in the regression (see Fig. 1, Kleijnen and Helton 1999b). Such variables appear to have a small effect on the output variable which, in fact, is due to chance variation. In such situations, a natural stopping point may be just before spurious variables start being selected. Another possibility is to delete spurious variables from the regression model.

When the input variables are uncorrelated, a display of the results of a stepwise regression analysis as shown in Table 8.5 contains a large amount of redundant information. A more compact display can be obtained by listing the variables in the order that they entered in the regression model, the R^2 values obtained with the entry of successive variables into the regression model, and the SRCs for the variables contained in the final model. Table 8.6 shows what this summary looks like for the stepwise regression analysis in Table 8.5.

Numerous examples of the use of stepwise regression analysis in sampling-based sensitivity analyses are available in various articles by Helton et al. (1996, 1995b, 1989).

Table 8.6. Compact Summary of Stepwise Regression Analysis for Pressure in the Repository at 10,000 yr Under Undisturbed Conditions (i.e., y = *E0:WAS_PRES* at 10,000 yr in Fig. 7.5)

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

8.6 The Rank Transformation

Regression and correlation analyses often perform poorly when the relationships between the input and output variables are nonlinear. This is not surprising since such analyses are based on developing linear relationships between variables. The problems associated with poor linear fits to nonlinear data can often be mitigated by use of the rank transformation (Iman and Conover 1979, Conover and Iman 1981, Saltelli and Sobol' 1995). The rank transformation is a simple concept: data are replaced with their corresponding ranks and then the usual regression and correlation procedures are performed on these ranks. Specifically, the smallest value of each variable is assigned the rank 1, the next largest value is assigned the rank 2, and so on up to the largest value, which is assigned the rank m, where m denotes the number of observations. Further, averaged ranks are assigned to equal values of a variable. The analysis is then performed with these ranks being used as the values for the input and output variables. In essence, the use of rank-transformed data results in an analysis based on the strength of monotonic relationships rather than on the strength of linear relationships.

As an example, the strength of the monotonic relationship between x and y can be measured with Spearman's rank CC (RCC) for x and y, R_{xy}, which is simply Pearson's CC in Eq. (8.29) calculated on ranks. The test for zero rank correlation uses a table of quantiles for $|R_{xy}|$ (e.g., Table A10, Conover 1980). For a sample size of $m \geq 30$,

$$z = R_{xy}\sqrt{m-1} \tag{8.46}$$

approximately follows a standard normal distribution if the rank correlation between x and y is zero (p. 456, Conover 1980). Thus, similarly to Eq. (8.40) for r_{xy},

$$prob_n(|\tilde{R}_{xy}|>|R_{xy}|) = erfc(|R_{xy}|\sqrt{m-1}/\sqrt{2}), \tag{8.47}$$

where $prob_n(|\tilde{R}_{xy}|>|R_{xy}|$ is the probability that random variation would produce a value $\tilde{R}_{xy}$ larger in absolute value than the observed value R_{xy}. Further, standardized rank regression coefficients (SRRCs) and partial rank CCs (PRCCs) can be calculated analogously to the corresponding coefficients for raw data.

For perspective, analyses for y = *E0:BRAALIC* at 10,000 yr (i.e., the value at 10,000 yr in Fig. 8.4) with CCs, SRCs and PCCs calculated with both raw and rank-transformed data are presented in Table 8.7. The general patterns exhibited by the analyses with raw data and by the analyses with rank-transformed data are similar to those discussed in conjunction with Table 8.3. However, the two analyses differ in the importance assigned to individual variables. In particular, the analysis with rank-transformed data identifies *WMICDFLG* as the most important variable with an RCC of –0.6521; in contrast, the analysis with raw data identifies *WMICDFLG* as the second most important variable with a CC of –0.3210. The preceding is a nontrivial

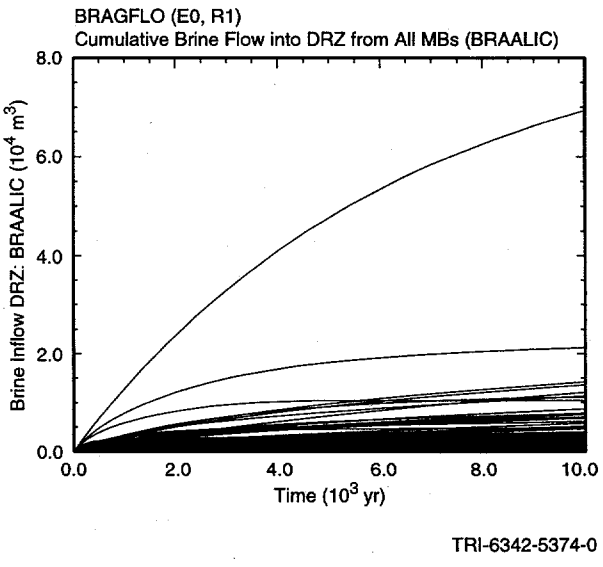


Fig. 8.4. Cumulative brine flow (m^3) into disturbed rock zone (DRZ) from all anhydrite marker beds (MBs) under undisturbed (i.e., E0) conditions (i.e., y = *E0:BRAALIC*).

Table 8.7. Correlation Coefficients (CCs, RCCs), Standardized Regression Coefficients (SRCs, SRRCs) and Partial Correlation Coefficients (PCCs, PRCCs) with Raw and Rank-Transformed Data for y = *E0:BRAALIC* at 10,000 yr in Fig. 8.4[a]

Raw Data

Variable		CC		SRC		PCC	
Name	p-Val	Rank	Value	Rank	Value	Rank	Value
ANHPRM	0.0000	1.0	0.5655	1.0	0.5568	1.0	0.6317
WMICDFLG	0.0000	2.0	–0.3210	2.0	–0.2931	2.0	–0.3878
WASTWICK	0.0045	3.0	–0.1639	4.0	–0.1451	4.0	–0.2075
WGRCOR	0.0048	4.0	–0.1628	3.0	–0.1669	3.0	–0.2370
ANHBCEXP	0.0095	5.0	–0.1497	5.0	–0.1155	5.0	–0.1663
WFBETCEL	0.0555	6.0	–0.1105	8.0	–0.0757	8.0	–0.1098
WRBRNSAT	0.0615	7.0	–0.1080	9.0	–0.0733	9.0	–0.1065
HALPOR	0.0934	8.0	–0.0969	6.0	–0.0993	6.0	–0.1435

Rank-Transformed Data

Variable		RCC		SRRC		PRCC	
Name	p-Val	Rank	Value	Rank	Value	Rank	Value
WMICDFLG	0.0000	1.0	–0.6521	1.0	–0.6533	1.0	–0.8787
ANHPRM	0.0000	2.0	0.5804	2.0	0.5937	2.0	0.8619
HALPRM	0.0014	3.0	0.1850	5.0	0.1443	5.0	0.3817
WGRCOR	0.0057	4.0	–0.1598	4.0	–0.1509	4.0	–0.3963
HALPOR	0.0087	5.0	–0.1518	3.0	–0.1539	3.0	–0.4031
WASTWICK	0.0405	6.0	–0.1185	7.0	–0.0948	7.0	–0.2617

[a] Table structure analogous to Table 8.3

difference because an RCC of –0.6521 implies that *WMICDFLG* can account for 42.5% of the uncertainty in y in rank-transformed space (i.e., $0.6521^2 \doteq 0.425$) while a CC of –0.3210 implies that *WMICDFLG* can account for only 10.3% of the uncertainty in y in the original untransformed space (i.e., $0.3210^2 \doteq 0.103$). Numerous other differences also exist.

Additional perspective on the use of raw and rank-transformed data in the analysis of y = *E0:BRAALIC* can be obtained from examination of the results of stepwise regression analyses (Table 8.8). In particular, the use of rank-transformed data leads to a regression model with 7 variables and an R^2 value of 0.869. In contrast, the use of raw data leads to a regression model with 6 variables and an R^2 value of only 0.496. Thus, the use of rank-transformed data is resulting in an analysis that can account for more of the uncertainty in y than can be accounted for in an analysis with raw data. As a result, the coefficients in Table 8.7 obtained with rank-transformed data (i.e., RCCs, SRRCs, PRCCs) are more informative with respect to the sources of the uncertainty in y than are the coefficients obtained with raw data.

When the relationship between the dependent and independent variables is linear, use of raw and rank-transformed data tends to produce similar results. When rank-transformed data are used and there are no ties in the data, the resulting values for regression coefficients and SRCs are equal; thus, the rank transform results in an automatic standardization of the data in this case.

The analysis with rank-transformed data is more effective than the analysis with raw data because the rank transformation tends to linearize the relationships between the independent variables (i.e., the x_j's) and the dependent variable (i.e., y). In particular, both *WMICDFLG* and *ANHPRM* show a stronger linear relationship with y after the rank transformation (Fig. 8.5). The rank transformation improves the analysis when nonlinear but monotonic relationships exist between the independent variables and the dependent

Table 8.8. Comparison of Stepwise Regression Analyses with Raw and Rank-Transformed Data for Cumulative Brine Flow over 10,000 yr under Undisturbed Conditions from the Anhydrite Marker Beds to the Disturbed Rock Zone that Surrounds the Repository (i.e., y = *E0:BRAALIC* at 10,000 yr in Fig. 8.4)

	Raw Data			Rank-Transformed Data		
Step[a]	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 with *ANHCOMP* and *HALCOMP* excluded from entry into regression model.
[c] Standardized regression coefficient (SRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.
[e] Standardized rank regression coefficients (SRRCs) in final regression model.

variable. When more complex relationships exist, the rank transformation may do little to improve the quality of an analysis. In such cases, more sophisticated procedures are required. For example, various tests can be used to check for deviations from randomness in scatterplots (Sects. 8.8, 8.9; also see Hamby 1994, Saltelli and Marivoet 1990, Kleijnen and Helton 1999a).

As for stepwise regression analyses, analyses with SRCs and PCCs of the type presented in Fig. 8.3 can often be improved with the use of rank-transformed data. When the rank transform is used, the resultant plots will contain SRRCs and PRCCs. As an example, the results of analyzing the cumulative brine inflows in Fig. 8.4 with both raw and rank-transformed data are presented in Fig. 8.6, with each plot frame showing the five variables with the largest, in absolute value, SRCs, PCCs, SRRCs and PRCCs as appropriate. As in the comparisons of stepwise regression analyses with raw and rank-transformed data (Table 8.8), the analyses with rank-transformed data in Fig. 8.6 produce outcomes that indicate stronger effects for individual variables than is the case for the analyses with raw data.

The rank transformation has become quite popular in sampling-based sensitivity analyses and many additional examples of its use exist (e.g., Sanchez and Blower 1997; Gwo et al. 1996; Helton et al. 1996, 1989; Hamby 1995; Blower and Dowlatabadi 1994; Whiting et al. 1993; MacDonald and Campbell 1986).

8.7 Effects of Correlations on Sensitivity Analyses

The presence of correlations between uncertain (i.e., sampled) variables can greatly complicate the interpretation of sensitivity analysis results. Regression-based sensitivity analyses for the variables in Fig. 7.4 will be used as an example (Table 8.9).

The regression analyses in Table 8.9 are all relatively successful in the sense that they have R^2 values between 0.86 and 0.91. However, inspection of the individual regression analyses indicates that there is an undesirable complication that results from the rank correlations of –0.99 that are assigned to the variable pairs (*ANHPRM*, *ANHCOMP*) and (*HALPRM*, *HALCOMP*) (Table 3.1). When no correlations exist between the sampled variables in the regression model, the regression coefficients will decrease monotonically in absolute value. In this case, an ordering of the variables by the absolute value of their regression coefficients provides a way to rank variable importance. However, when correlated variables are included in a regression model, the sizes and even the signs of the associated regression coefficients may not properly indicate the effects of these variables. This behavior appears in Table 8.9 for the pair (*HALPRM*, *HALCOMP*) in the regressions for Anhydrites a and b North (*E0:BRAABNIC*), Anhydrites a and b South (*E0:BRAABSIC*), MB 139 North (*E0:BRM39NIC*), MB 139 South

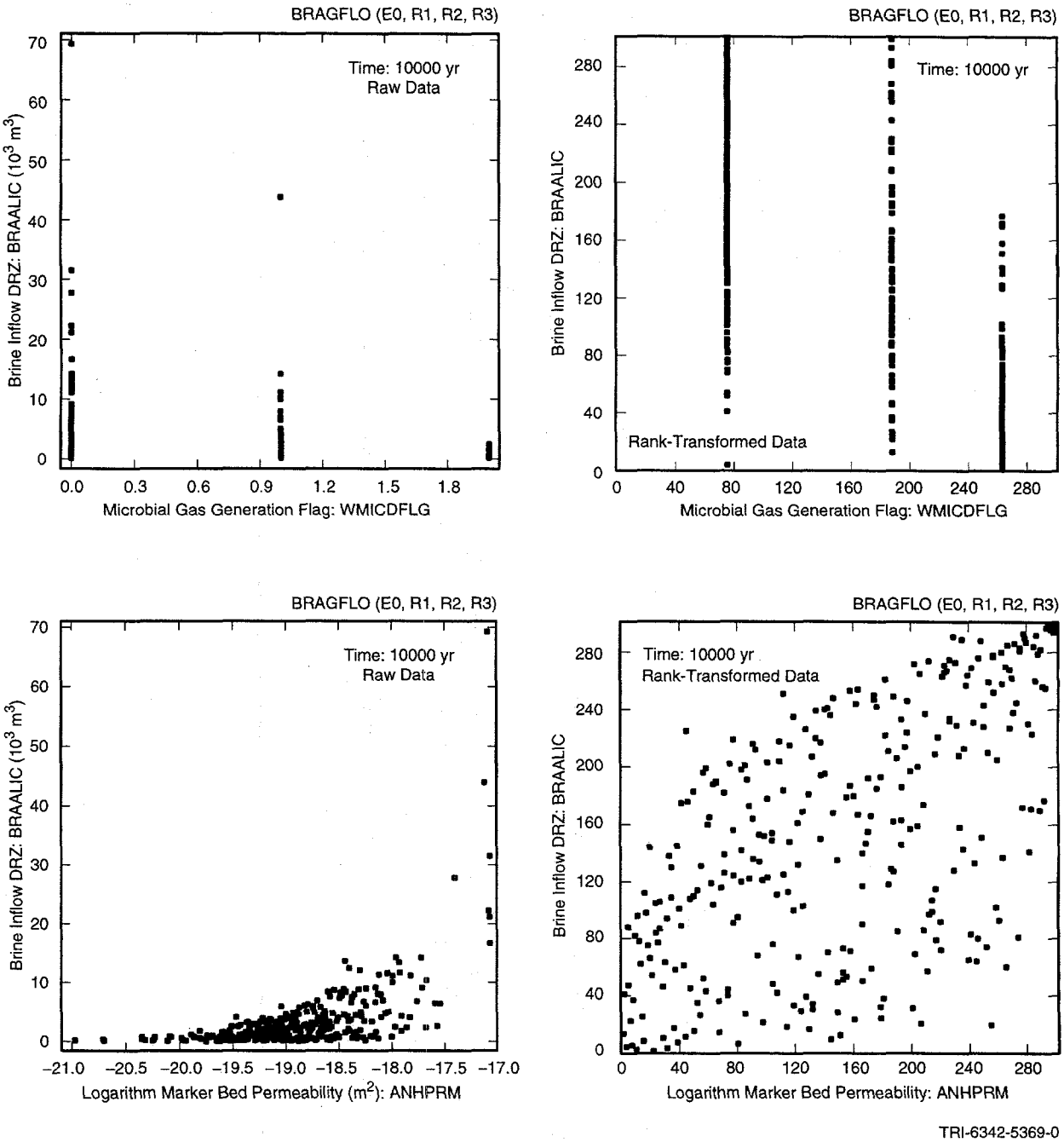


Fig. 8.5. Scatterplots for cumulative brine discharge (m^3) from the marker beds over 10,000 yr under undisturbed conditions (i.e., y = *E0:BRAALIC* at 10,000 yr in Fig. 8.4) versus microbial gas generation flag (*WMICDFLG*) and marker bed permeability (*ANHPRM*) with raw (i.e., untransformed) and rank-transformed data.

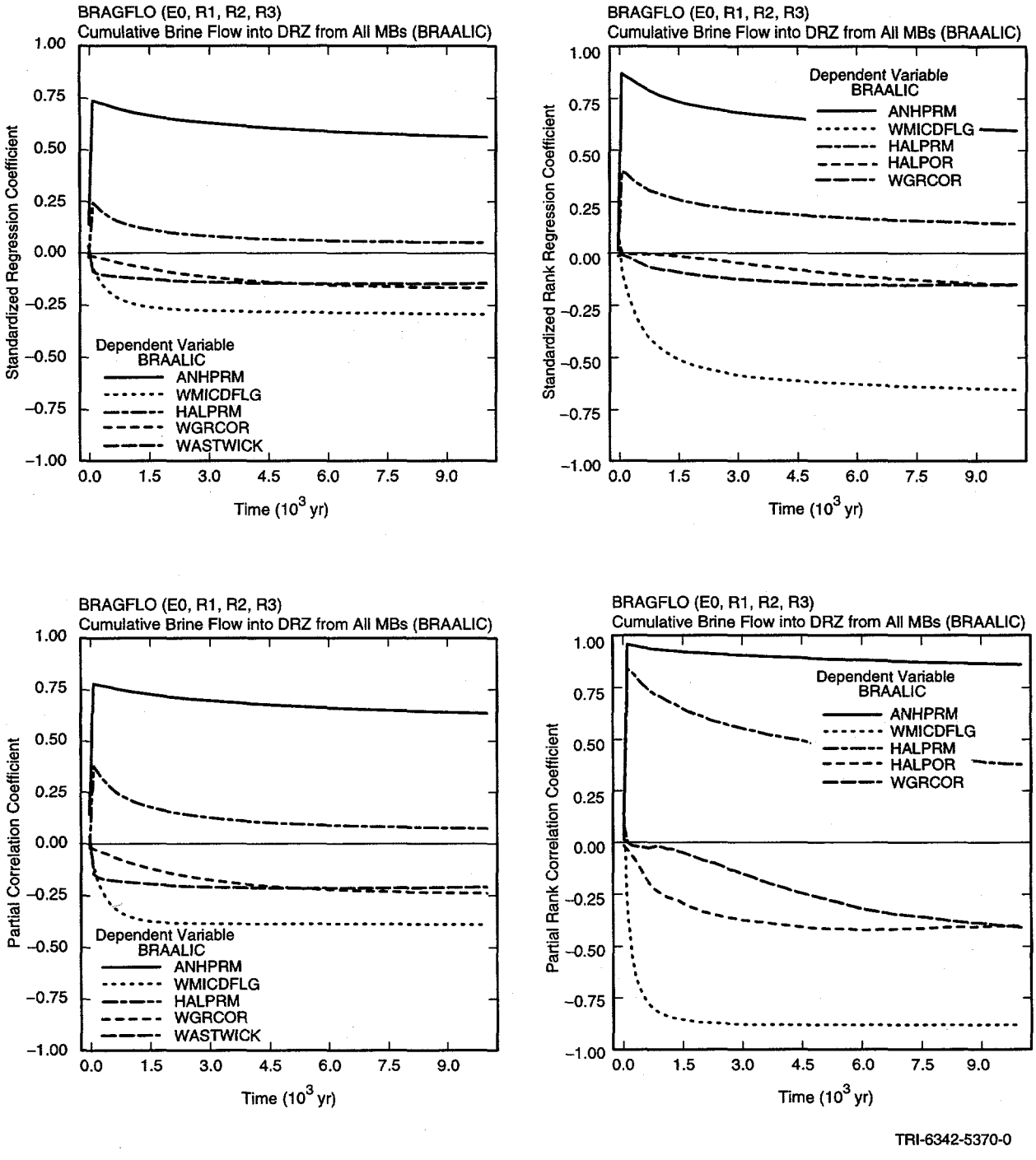


Fig. 8.6. Standardized regression coefficients (SRCs, SRRCs) and partial correlation coefficients (PCCs, PRCCs) calculated with raw and rank-transformed data for cumulative brine flow from anhydrite marker beds to disturbed rock zone (DRZ) under undisturbed conditions (i.e., y = *E0:BRAALIC* in Fig. 8.4) with *ANHCOMP* and *HALCOMP* excluded from calculation.

Table 8.9. Stepwise Regression Analyses with Rank-Transformed Data for Cumulative Brine Flow over 10,000 yr into DRZ (*E0:BRM38NIC, E0:BRM38SIC, E0:BRAABNIC, E0:BRAABSIC, E0:BRM39NIC, E0:BRM39SIC, E0:BRAALIC*) and into repository (*E0:BRNREPTC*) under Undisturbed Conditions (see Fig. 7.4)

	MB 138 North: *E0:BRM38NIC*			MB 138 South: *E0:BRM38SIC*			Anh a and b North: *E0:BRAABNIC*			Anh a and b South: *E0:BRAABSIC*		
Step[a]	Variable[b]	SRRC[c]	R^{2d}	Variable	SRRC	R^2	Variable	SRRC	R^2	Variable	SRRC	R^2
1	*ANHPRM*	0.75	0.54	*ANHPRM*	0.73	0.51	*WMICDFLG*	−0.66	0.43	*WMICDFLG*	−0.66	0.43
2	*WMICDFLG*	−0.52	0.80	*WMICDFLG*	−0.55	0.80	*ANHPRM*	0.60	0.79	*ANHPRM*	0.59	0.77
3	*HALCOMP*	0.21	0.84	*HALCOMP*	0.18	0.83	*HALPOR*	−0.15	0.81	*HALPOR*	−0.16	0.80
4	*HALPOR*	−0.11	0.86	*WGRCOR*	−0.13	0.85	*WGRCOR*	−0.16	0.84	*WGRCOR*	−0.16	0.83
5	*WGRCOR*	−0.12	0.87	*HALPOR*	−0.11	0.86	*SALPRES*	0.11	0.85	*SALPRES*	0.11	0.84
6	*SALPRES*	0.11	0.88	*SALPRES*	0.10	0.87	*WASTWICK*	−0.09	0.86	*HALPRM*	0.54	0.85
7	*WASTWICK*	−0.08	0.89	*WASTWICK*	−0.08	0.88	*HALPRM*	0.49	0.87	*WASTWICK*	−0.09	0.86
8	*WGRMICI*	−0.06	0.89	*WGRMICI*	−0.06	0.88	*HALCOMP*	0.40	0.87	*HALCOMP*	0.43	0.86
9	*SHRGSSAT*	−0.04	0.90	*SHRGSSAT*	−0.05	0.88	*SHRGSSAT*	−0.05	0.87	*SHRGSSAT*	−0.05	0.87

	MB 139 North: *E0:BRM39NIC*			MB 139 South: *E0:BRM39SIC*			MBs Total: *E0:BRAALIC*			Repository Total: *E0:BRNREPTC*		
Step	Variable	SRRC	R^2	Variable	SRRC	R^2	Variable	SRRC	R^2	Variable	SRRC	R^2
1	*WMICDFLG*	−0.65	0.42	*WMICDFLG*	−0.65	0.43	*WMICDFLG*	−0.65	0.43	*HALPOR*	0.88	0.77
2	*ANHPRM*	0.59	0.78	*ANHPRM*	0.57	0.75	*ANHPRM*	0.59	0.78	*WMICDFLG*	−0.26	0.85
3	*HALPOR*	−0.16	0.80	*HALPRM*	0.55	0.79	*HALPOR*	−0.16	0.80	*ANHPRM*	0.60	0.88
4	*HALPRM*	0.52	0.83	*HALPOR*	−0.16	0.81	*WGRCOR*	−0.15	0.82	*HALCOMP*	−0.09	0.89
5	*WGRCOR*	−0.15	0.85	*WGRCOR*	−0.15	0.84	*HALPRM*	0.51	0.85	*WRBRNSAT*	−0.09	0.89
6	*SALPRES*	0.12	0.86	*SALPRES*	0.12	0.85	*SALPRES*	0.12	0.86	*WGRCOR*	−0.08	0.90
7	*WASTWICK*	−0.10	0.87	*WASTWICK*	−0.10	.086	*WASTWICK*	−0.10	0.87	*ANHCOMP*	0.43	0.91
8	*HALCOMP*	0.37	0.88	*HALCOMP*	0.37	0.86	*HALCOMP*	0.37	0.87	*WASTWICK*	−0.06	0.91

[a] Steps in stepwise regression analysis.
[b] Variables listed in order of selection in regression analysis.
[c] Standardized rank regression coefficients (SRRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.

(*E0:BRM39SIC*) and MBs Total (*E0:BRAALIC*), and for the pair (*ANHPRM, ANHCOMP*) in the regression for Repository Total (*E0:BRNREPTC*). In particular, the existence of the strong correlations within the pairs (*HALPRM, HALCOMP*) and (*ANHPRM, ANHCOMP*) results in a nonmonotonic behavior of the associated regression coefficients.

As a more detailed example, explicit representations of the following three regression analyses for MBs Total (*E0:BRAALIC*) are shown in Table 8.10: (i) all 31 sampled variables allowed as candidates for inclusion in the regression model, (ii) *ANHCOMP* and *HALCOMP* excluded as candidates for inclusion in the regression model, and (iii) *ANHPRM* and *HALPRM* excluded as candidates for inclusion in the regression model. When all sampled variables are included as candidates, the regression coefficients decrease monotonically until Step 8, when *HALCOMP* enters the regression model. With entry of *HALCOMP*, the regression coefficient for *HALPRM* jumps from a value of 0.14 at Step 7 to a value of 0.51; further, *HALCOMP* has a regression coefficient of 0.37 even though it has essentially no effect on the R^2 value for the regression model (i.e., R^2 = 0.86889 at Step 7 and R^2 = 0.87203 at Step 8). When *ANHCOMP* and *HALCOMP* are excluded as candidates for entry into the regression model, a sequence of 7 regression models is produced that is identical to the first 7 regression models that are produced when all variables are allowed as candidates for inclusion. However, a different sequence of regression models is constructed when *ANHPRM* and *HALPRM* are excluded.

Table 8.10. Detailed Stepwise Regression Analyses with Rank-Transformed Data for Cumulative Brine Flow From all Marker Beds over 10,000 yr under Undisturbed Conditions (i.e., *y* = *E0:BRAALIC* in Fig. 7.4 and also at 10,000 yr in Fig. 8.4)

All Variables Included		*ANHCOMP* *HALCOMP* Excluded		*ANHPRM* *HALPRM* Excluded		All Variables		*ANHCOMP* *HALCOMP* Excluded		*ANHPRM* *HALPRM* Excluded	
Variable[a]	SRRC[b]	Variable	SRRC	Variable	SRRC	Variable	SRRC	Variable	SRRC	Variable	SRRC
Step 1[c]						Step 6					
WMICDFLG	−0.65	*WMICDFLG*	−0.65	*WMICDFLG*	−0.65	*WMICDFLG*	−0.66	*WMICDFLG*	−0.66	*WMICDFLG*	−0.66
R^2[d]	0.43	R^2	0.43	R^2	0.43	*ANHPRM*	0.59	*ANHPRM*	0.59	*ANHCOMP*	−0.59
Step 2						*HALPOR*	−0.16	*HALPOR*	−0.16	*HALPOR*	*−0.16*
						WGRCOR	−0.15	*WGRCOR*	−0.15	*WGRCOR*	*−0.15*
WMICDFLG	−0.66	*WMICDFLG*	−0.66	*WMICDFLG*	−0.67	*HALPRM*	0.14	*HALPRM*	0.14	*HALCOMP*	−0.14
ANHPRM	0.59	*ANHPRM*	0.59	*ANHCOMP*	−0.58	*SALPRES*	0.12	*SALPRES*	0.12	*SALPRES*	0.13
R^2	0.77	R^2	0.77	R^2	0.76	R^2	0.86	R^2	0.86	R^2	0.85
Step 3						Step 7					
WMICDFLG	−0.66	*WMICDFLG*	−0.66	*WMICDFLG*	−0.67	*WMICDFLG*	−0.66	*WMICDFLG*	−0.66	*WMICDFLG*	−0.66
ANHPRM	0.59	*ANHPRM*	0.59	*ANHCOMP*	−0.58	*ANHPRM*	0.59	*ANHPRM*	0.59	*ANHCOMP*	−0.58
HALPOR	−0.16	*HALPOR*	−0.16	*HALPOR*	−0.16	*HALPOR*	−0.16	*HALPOR*	−0.16	*HALPOR*	−0.16
R^2	0.80	R^2	0.80	R^2	0.79	*WGRCOR*	−0.15	*WGRCOR*	−0.15	*WGRCOR*	−0.15
Step 4						*HALPRM*	0.14	*HALPRM*	0.14	*HALCOMP*	−0.14
						SALPRES	0.12	*SALPRES*	0.12	*SALPRES*	0.13
WMICDFLG	−0.66	*WMICDFLG*	−0.66	*WMICDFLG*	−0.66	*WASTWICK*	−0.10	*WASTWICK*	−0.10	*WASTWICK*	−0.09
ANHPRM	0.60	*ANHPRM*	0.60	*ANHCOMP*	−0.58	R^2	0.87	R^2	0.87	R^2	0.85
HALPOR	−0.16	*HALPOR*	−0.16	*HALPOR*	−0.16	Step 8					
WGRCOR	−0.15	*WGRCOR*	−0.15	*WGRCOR*	−0.15						
R^2	0.82	R^2	0.82	R^2	0.81	*WMICDFLG*	−0.65	No additional		No additional	
Step 5						*ANHPRM*	0.59	variable		variable	
						HALPOR	−0.16	selected		selected	
WMICDFLG	−0.65	*WMICDFLG*	−0.65	*WMICDFLG*	−0.66	*WGRCOR*	−0.15				
ANHPRM	0.59	*ANHPRM*	0.59	*ANHCOMP*	−0.58	*HALPRM*	0.51				
HALPOR	−0.16	*HALPOR*	−0.16	*HALPOR*	−0.16	*SALPRES*	0.12				
WGRCOR	−0.15	*WGRCOR*	−0.15	*WGRCOR*	−0.15	*WASTWICK*	−0.10				
HALPRM	0.15	*HALPRM*	0.15	*HALCOMP*	−0.14	*HALCOMP*	0.37				
R^2	0.85	R^2	0.85	R^2	0.83	R^2	0.87				

[a] Variables in regression model.
[b] Standardized rank regression coefficients (SRRCs) for variables in regression model.
[c] Steps in stepwise regression analysis.
[d] R^2 value for regression model.

In this case, *ANHPRM* and *HALPRM* are replaced in the regression models with *ANHCOMP* and *HALCOMP*, and the signs of the regression coefficients are reversed. Thus, *ANHCOMP* and *HALCOMP* appear with negative regression coefficients where *ANHPRM* and *HALPRM* appear with positive regression coefficients. In contrast, *HALPRM* and *HALCOMP* both have positive regression coefficients when they appear together in the regression model constructed at Step 8 when all variables are included as candidates for entry into the analysis. Thus, care must be used in interpreting regression analyses that involve highly correlated variables.

8.8 Identification of Nonmonotonic Patterns

Sometimes regression-based sensitivity analyses perform very poorly. The rank transformation has been introduced as a possible analysis procedure for such situations (Sect. 8.5). However, when viewed broadly, the rank transformation provides only a variant on linear regression analysis, with a model that seeks to identify linear relationships being replaced by a model that seeks to identify monotonic relationships. A more general approach is to attempt to determine if the scatterplots of a dependent (i.e., predicted) variable versus individual independent (i.e., sampled) variables appear to display nonmonotonic patterns.

As an example, time-dependent pressure in the repository subsequent to an E2 intrusion and an associated sensitivity analysis based on PRCCs are shown in Fig. 8.7, with the PRCCs having small values after the occurrence of the drilling intrusion at 1000 yr. Further, as indicated by the regression analyses in Table 8.11 for pressure at 10,000 yr, the use of neither raw nor rank-

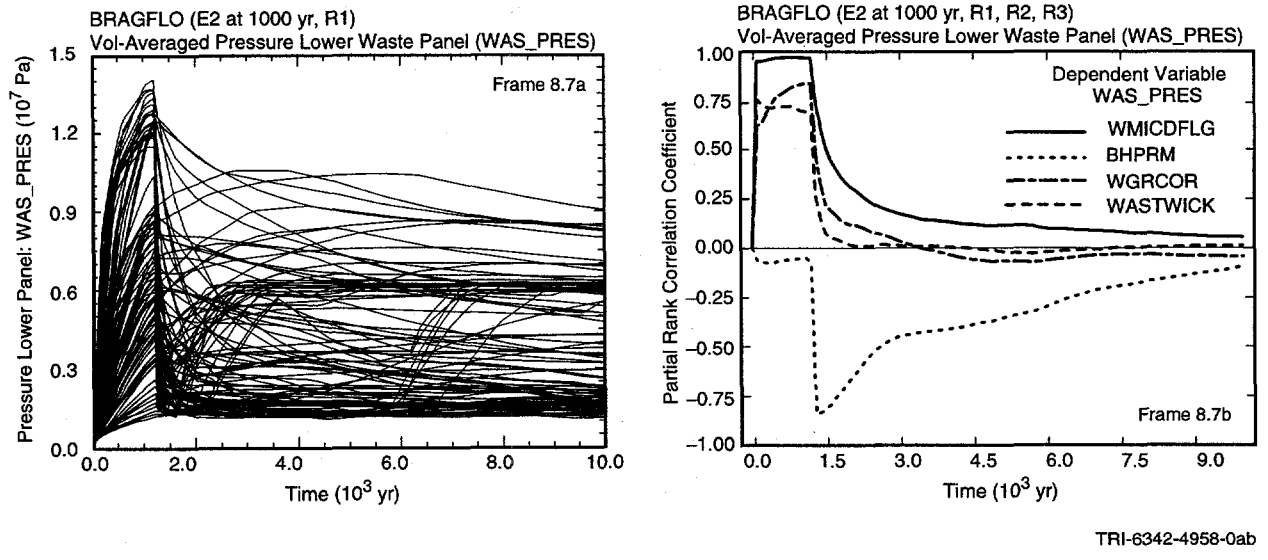


Fig. 8.7. Uncertainty and sensitivity analysis results for repository pressure for an E2 intrusion into lower waste panel at 1000 yr (i.e., y = *E2:WAS_PRES*).

Table 8.11. Stepwise Regression Analyses with Raw and Rank-Transformed Data with Pooled Results from Replicates R1, R2 and R3 (i.e., for a total of 300 observations) for y = *E2:WAS_PRES* at 10,000 yr

	Raw Data, *E2:WAS_PRES*			Rank-Transformed Data, *E2:WAS_PRES*		
Step[a]	Variable[b]	SRC[c]	R^{2d}	Variable[b]	SRRC[e]	R^{2d}
1	*HALPRM*	0.37	0.14	*HALPRM*	0.36	0.13
2	*ANHPRM*	0.24	0.20	*ANHPRM*	0.24	0.19
3	*HALPOR*	0.14	0.22	*HALPOR*	0.14	0.20

[a] Steps in stepwise regression analysis.
[b] Variables listed in order of selection in regression analysis with *ANHCOMP* and *HALCOMP* excluded from entry into regression model because of −0.99 rank correlation within the pairs (*ANHPRM, ANHCOMP*) and (*HALPRM, HALCOMP*).
[c] Standardized regression coefficients (SRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.
[e] Standardized rank regression coefficients (SRRCs) in final regression model.

transformed data in a stepwise regression analysis produced a very successful regression model (i.e., R^2 values of 0.22 and 0.20 result for raw and rank-transformed data, respectively). Yet, unless there is some type of error in the calculations, the uncertainty in the sampled variables must be giving rise to the variations in the pressure curves in Fig. 8.7a.

As discussed in Sect. 8.1, the examination of scatterplots often provides an effective way to identify influential variables. In particular, the examination of scatterplots shows that *BHPRM* is the dominant variable with respect to the uncertainty in repository pressure subsequent to an E2 intrusion (Fig. 8.2). This is rather disconcerting as *BHPRM* was not identified in either the PRCC analysis in Fig. 8.7b or the regression analyses in Table 8.11. Thus, the clearly dominant variable has been completely missed in the formal analyses in Fig. 8.7b and Table 8.11, and was only identified by an exhaustive examination of the scatterplots for the individual variables. Clearly, some type of formal procedure for identifying patterns in scatterplots is desirable; otherwise, the analyst is confronted with the requirement to manually examine large numbers of scatterplots and also to subjectively assess the relative strengths of the individual patterns appearing in these plots.

In this section, three procedures for identifying nonmonotonic patterns are introduced. Each of these procedures is based on determining if some measure of central tendency for the dependent variable is a function of individual independent variables. In particular, the F-test for equal means, the χ^2-test for equal medians, and the Kruskal-Wallis test are introduced as means of determining if measures of central tendency for a dependent variable change as a function of the values of individual independent variables (Sect. 5, Kleijnen and Helton 1999a). For convenience, the preceding tests will be designated as tests for common means (CMNs), common medians (CMDs) and common locations (CLs).

The procedures discussed in this section involve an assessment of the relationship between a dependent and an independent variable. For notational convenience, these variables will be represented by y and x, respectively. This assessment is based on dividing the values of x (i.e., x_k, $k = 1, 2, \ldots, m$) into nX classes and then testing to determine if y has a common measure of central tendency across these classes. The required classes are obtained by dividing the range of x into a sequence of mutually exclusive and exhaustive subintervals containing equal numbers of sampled values (Fig. 8.8). When an x is discrete (e.g., see *WMICDFLG* in Fig. 8.5), individual classes are defined for each of the distinct values. For notational convenience, let q, $q = 1, 2, \ldots, nX$, designate the individual classes into which the values of x have been divided; let $\mathcal{X}_q$ designate the set such that $k \in \mathcal{X}_q$ only if x_k belongs to class q; and let nX_q equal the number of elements contained in $\mathcal{X}_q$ (i.e., the number of x_k's associated with class q).

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 have been divided (e.g., the intervals defined on the abscissas of the scatterplots in Fig. 8.8). Specifically, if the y values conditional on each class of x values are normally distributed with equal expected values, then

$$F = \frac{\left[\sum_{q=1}^{nX} nX_q \bar{y}_q^2 - m\bar{y}^2\right] /(nX-1)}{\left[\sum_{k=1}^{m} y_k^2 - \sum_{q=1}^{nX} nX_q \bar{y}_q^2\right] /(m-nX)} \tag{8.48}$$

follows an F-distribution with $(nX - 1, m - nX)$ degrees of freedom, where $\bar{y}_q = \sum_{k \in \mathcal{X}_q} y_k / nX_k$ and $\bar{y}$ is defined in conjunction with Eq. (8.15). Given that the indicated assumptions hold, the probability of obtaining an F-statistic of value $\tilde{F}$ that exceeds the value of F in Eq. (8.48) can be estimated by $prob_F(\tilde{F} > F \mid nX-1, m-nX)$ as defined in Eq. (8.23). A low probability (i.e., p-value) of obtaining a larger value for F suggests that the observed pattern involving x and y did not arise by chance and hence that x has an effect on the behavior of y.

The χ^2-test for contingency tables can be used to test for the equality of the median values of y for the classes into which the values of x have been divided (pp. 143-178, Conover 1980). First, the median, $y_{0.5}$, is estimated for all m observations. Specifically,

$$y_{0.5} = \begin{cases} y_{(0.5m)} & \text{if } 0.5\,m \text{ is an integer} \\ [y_{([0.5m])} + y_{([0.5m]+1)}]/2 & \text{otherwise,} \end{cases} \tag{8.49}$$

where $y_{(k)}$, $k = 1, 2, \ldots, m$, denotes the ordering of the y-values such that $y_{(k)} \le y_{(k+1)}$ and $[\sim]$ designates the greatest integer function (p. 14, David 1981). The individual classes of x values are then further subdivided on the basis of whether y values fall above or below $y_{0.5}$ (Fig. 8.8). For class q, let nX_{1q} equal the number of y

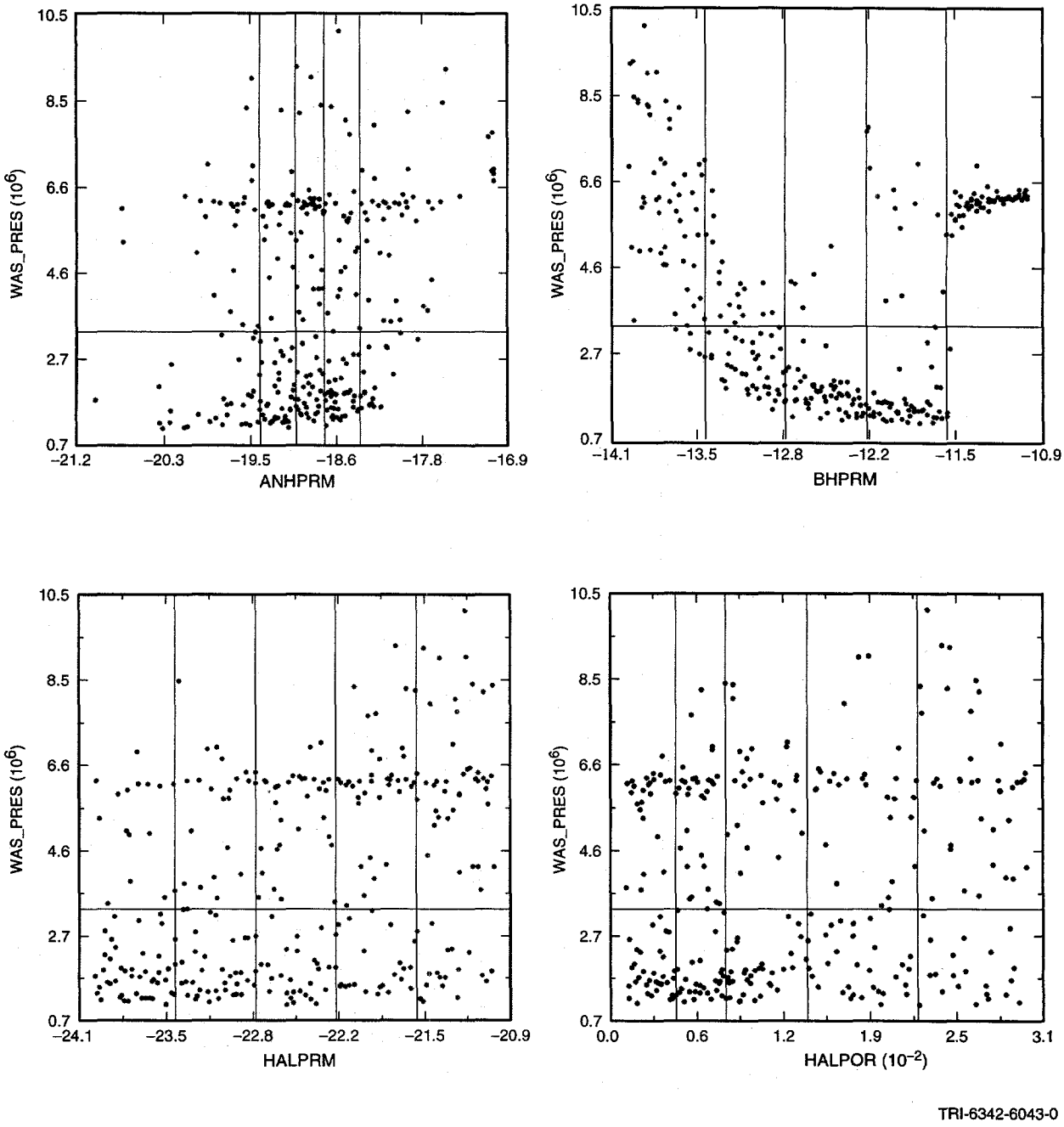


Fig. 8.8. Partitioning of the ranges of *ANHPRM*, *BHPRM*, *HALPRM* and *HALPOR* into $nX = 5$ classes in an analysis for y = *E2:WAS_PRES* at 10,000 yr; horizontal lines correspond to the median, $y_{0.5}$, of y.

values that exceed $y_{0.5}$, and let nX_{2q} equal the number of y values that are less than or equal to $y_{0.5}$.

The result of this partitioning is a $2 \times nX$ contingency table with nX_{rq} observations in each cell. The following statistic can now be defined:

$$T = \sum_{q=1}^{nX} \sum_{r=1}^{2} (nX_{rq} - nE_{rq})^2 / nE_{rq}, \qquad (8.50)$$

where

$$nE_{rq} = \left(\sum_{r=1}^{2} nX_{rq} \right) \left(\sum_{q=1}^{nX} nX_{rq} \right) / nX$$

and corresponds to the expected number of observations in cell (r, q). If the individual classes of x values, $q = 1, 2, \ldots, nX$, have equal medians, then T approximately follows a χ^2 distribution with $(nX - 1)(2 - 1) = nX - 1$ degrees of freedom (p. 156, Conover 1980). The probability $prob_{\chi^2}(\tilde{T} > T | nX - 1)$ of obtaining a value $\tilde{T}$ that exceeds T in the presence of equal medians is given by

$$prob_{\chi^2}(\tilde{T} > T | nX - 1) = Q[(nX - 1)/2, T/2], \qquad (8.51)$$

where $Q(a, b)$ designates the complement of the incomplete gamma function (p. 215, Press et al. 1992). A small value (i.e., p-value) for $prob_{\chi^2}(\tilde{T} > T | nX - 1)$ indicates that the y's conditional on individual classes have different medians and hence that x has an influence on y. To maintain the validity of the χ^2-test in the analysis of contingency tables, Conover suggests using a partition in which $nE_{rq} \geq 1$ (p. 156, Conover 1980).

The Kruskal-Wallis test statistic, T, is based on rank-transformed data and uses the same classes of x values as the F-statistic in Eq. (8.48) (pp. 229-230, Conover 1980). Specifically,

$$T = \left[\sum_{q=1}^{nX} (R_q^2 / nX_q) - m(m+1)^2 / 4 \right] \Big/ s^2, \qquad (8.52)$$

where

$$R_q = \sum_{k \in {}_{\circ}X_q} r(y_k),$$

$$s^2 = \left[\sum_{k=1}^{m} r(y_k)^2 - m(m+1)^2 / 4 \right] \Big/ (m-1),$$

and $r(y_k)$ denotes the rank of y_k. If the y values conditional on each class of x values have the same distribution, then the statistic T in Eq. (8.52) approximately follows a χ^2 distribution with $nX - 1$ degrees of freedom (pp. 230 - 231, Conover 1980). Given this approximation, the probability $prob_{\chi^2}(\tilde{T} > T | nX - 1)$ of obtaining a value $\tilde{T}$ that exceeds T in the presence of identical y distributions for the individual classes is given by Eq. (8.51). A small value for $prob_{\chi^2}(\tilde{T} > T | nX - 1)$ (i.e., a p-value) indicates that the 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 has an effect on y.

For y = *E2:WAS_PRES*, the three tests for non-monotonic relationships introduced in this section (i.e., CMNs, CMDs, CLs) all identify *BHPRM* as the most influential variable (Table 8.12). In contrast, the effect of *BHPRM* was missed in analyses based on correlation coefficients with raw and rank-transformed data (Table 8.12). Further, the three tests assign identical rankings to all variables with p-values below 0.1. After *BHPRM*, the next two most important variables as indicated by p-values are *HALPRM* and *ANHPRM*. These variables were also indicated as having effects with correlation coefficients with raw and rank-transformed data; however, as indicated by the low R^2 values in the associated regression models (i.e., 0.20 and 0.19 in Table 8.11), these variables by themselves are not very effective in accounting for the uncertainty in y in a regression-based analysis.

8.9 Identification of Random Patterns

The three tests described in the preceding section attempt to identify departures from monotonic trends. An even less restrictive approach to identifying influential variables is to determine if the scatterplot for the points (x_k, y_k), $k = 1, 2, \ldots, m$, appears to be random conditional on the marginal distributions for x and y. Specifically, the χ^2-test can be used to indicate if the pattern appearing in a scatterplot appears to be

Table 8.12. Comparison of Variable Rankings with Different Analysis Procedures for y = *E2:WAS_PRES* at 10,000 yr and a Maximum of Five Classes of Values for Each Variable (i.e., $nX = 5$)

Variable	CC[b]		RCC[c]		CMN: 1 × 5[d]		CMD: 2 × 5[e]		CL: 1 × 5[f]	
Name[a]	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
HALPRM	1.0	0.0000	1.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0000	3.0	0.0002	3.0	0.0007	3.0	0.0000
HALPOR	3.0	0.0090	3.0	0.0184	5.0	0.0415	5.0	0.0700	5.0	0.0940
ANHBCEXP	7.0	0.1786	8.0	0.2373	4.0	0.0405	4.0	0.0595	4.0	0.0602
BHPRM	10.0	0.3651	6.0	0.1704	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANRBRSAT	19.0	0.7133	14.0	0.4378	7.0	0.1513	6.0	0.0823	7.0	0.1304

[a] Variables for which at least one of the tests (i.e., CC, RCC, CMN:1×5, CMD:2×5, CL: 1×5) has a *p*- or α-value less than 0.1; variables ordered by *p*-values for CCs
[b] Ranks and *p*-values for CCs
[c] Ranks and *p*-values for RCCs
[d] Ranks and *p*-values for CMNs test with 1×5 grid
[e] Ranks and *p*-values for CMDs test with 2×5 grid
[f] Ranks and *p*-values for CLs (Kruskal-Wallis) test with 1×5 grid

nonrandom (Sect. 7, Kleijnen and Helton 1999a, Wagner 1995). For convenience, the χ^2-test for nonrandom patterns will be denoted as a test for statistical independence (SI).

With the χ^2- test, the values for the sampled variable (i.e., the x values on the abscissa) are divided into classes (Fig. 8.9). As in Sect. 8.8, let q, q = 1, 2, ..., nX, designate the individual classes into which the values of x have been divided; let $\mathcal{X}_q$ designate the set such that $k \in \mathcal{X}_q$ only if x_k belongs to class q; and let nX_q equal the number of elements contained in $\mathcal{X}_q$ (i.e., the number of x_k's associated with class q). Similarly, the values for the dependent variable (i.e., the y values on the ordinate) are also divided into classes (Fig. 8.9). For notational convenience, let p, p = 1, 2, ..., nY, designate the individual classes into which the values of y are divided; let $\mathcal{Y}_p$ designate the set such that $k \in \mathcal{Y}_p$ only if y_k belongs to class p; and let nY_p equal the number of elements contained in $\mathcal{Y}_p$ (i.e., the number of y_k's associated with class p). Typically, the classes $\mathcal{X}_q$ and $\mathcal{Y}_p$ are defined by ordering the x_k's and y_k's, respectively, and then requiring the individual classes to have similar numbers of elements (i.e., the nX_q are approximately equal for q = 1, 2, ..., nX, and the nY_p are approximately equal for p = 1, 2, ..., nY).

The partitioning of x and y into nX and nY classes in turn partitions (x, y) into nX nY classes (Fig. 8.9), where (x_k, y_k) belongs to class (q, p) only if x_k belongs to class q of the x values (i.e., $k \in \mathcal{X}_q$) and y_k belongs to class p of the y values (i.e., $k \in \mathcal{Y}_p$). For notational convenience, let O_{qp} denote the set such that $k \in O_{qp}$ only if $k \in \mathcal{X}_q$ (i.e., x_k is in class q of x values) and also $k \in \mathcal{Y}_p$ (i.e., y_k is in class p of y values), and let nO_{qp} equal the number of elements contained in O_{qp}. Further, if x and y are independent, then

$$nE_{qp} = (nY_p / m)(nX_q / m)m = nY_p \, nX_q / m \qquad (8.53)$$

is an estimate of the expected number of observations (x_k, y_k) that should fall in class (q, p).

The following statistic can be defined:

$$T = \sum_{q=1}^{nX} \sum_{p=1}^{nY} (nO_{qp} - nE_{qp})^2 / nE_{qp} \, . \qquad (8.54)$$

Asymptotically, T follows a χ^2-distribution with (nX–1) (nY–1) degrees of freedom when x and y are independent. Thus, the probability $prob_{\chi^2}[\tilde{T} > T \mid (nX-1)(nY-1)]$ of obtaining a value of $\tilde{T}$ that exceeds T when x and y are independent is given by Eq. (8.51).

The preceding probability provides a way to identify scatterplots that appear to display a significant relationship (i.e., pattern) involving the x and y variables on the abscissa and ordinate. In particular, $prob_{\chi^2}[\tilde{T} > T \mid (nX-1)(nY-1)]$ is the probability that a larger value of the statistic would occur due to chance variation (i.e., a p-value). A small p-value indicates that,

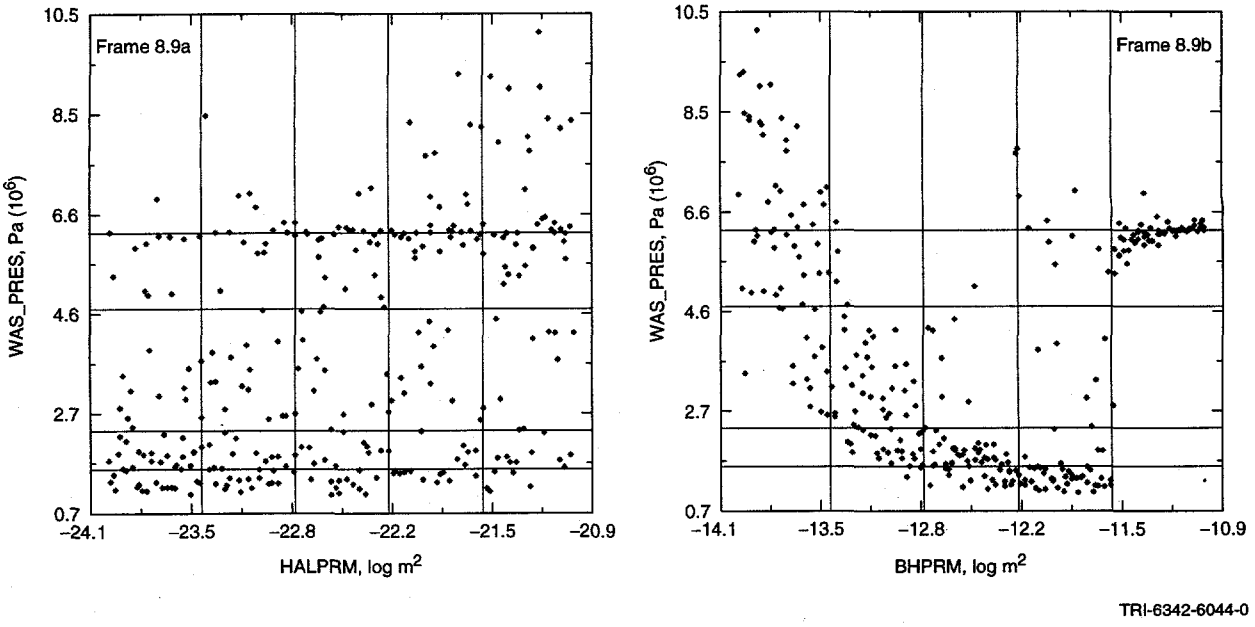


Fig. 8.9. Examples of the partitioning of the ranges of x = *HALPRM* and x = *BHPRM* into $nX = 5$ classes and the range of y = *E2:WAS_PRES* at 10,000 yr into $nY = 5$ classes.

under the assumptions of the test, an outcome equal to or greater than the observed value of the statistic is unlikely to occur due to chance. Thus, the implication is that the pattern in the scatterplot arose from some underlying relationship involving x and y rather than from chance alone.

As an example, a ranking of variable importance based on p-values for y = *E2:WAS_PRES* and a 5×5 grid (Fig. 8.9) is given in Table 8.13 under the heading SI: 5×5. The most important variable is *BHPRM*, which is consistent with the well-defined pattern in the corresponding scatterplot in Fig. 8.9b. In contrast, this pattern is completely missed by the regression analyses in Table 8.11. The next most important variable is *HALPRM*, with an effect that is discernible but rather weak in the corresponding scatterplot (Fig. 8.9a). In particular, the dependent variable tends to increase as *HALPRM* increases but with much noise around this trend. After *BHPRM* and *HALPRM*, small possible effects are indicated for *WGRCOR* and *ANHPRM*, with the corresponding scatterplots showing what are at best rather weak patterns (Fig. 8.8). After *BHPRM*, *HALPRM*, *WGRCOR* and *ANHPRM*, the p-values increase rapidly (Table 8.13), and there is little reason to believe that the ordering of the remaining variables on the basis of their p-values is due to anything other than chance. Similar variable rankings were also obtained in the analyses with CMNs, CMDs and CLs (Table 8.12).

The χ^2-statistic for identifying nonrandom patterns is based on superimposing grids on the scatterplots under consideration (Fig. 8.9). As a result, the outcome of such an analysis can depend on the grid selected for use. In particular, different grids can lead to different orderings of variable importance, although the identification of strong patterns is probably relatively insensitive to reasonable grid selections (i.e., grids that do not have an excessive number of cells relative to the number of points in the scatterplots under consideration). As an example, a ranking of variable importance based on p-values for y = *E2:WAS_PRES* and a 10×10 grid is given in Table 8.13 under the heading SI: 10×10. The rankings with 5×5 and 10×10 grids produce similar but not identical results, with both grids resulting in the identification of *BHPRM* as the most important variable and the identification of *BHPRM*, *HALPRM*, *WGRCOR* and *ANHPRM* as the four most important variables. Both analyses suggest that none of the remaining variables have a discernible effect on *E2:WAS_PRES*. Similar robustness is also present in the analyses of y with CMNs, CMDs and CLs (Table 24, Kleijnen and Helton 1999a; see Tables 8, 14 and 19 of Kleijnen and Helton 1999a for comparisons with additional variables).

The p-values used to identify important variables in Table 8.13 are calculated with statistical assumptions that are not fully satisfied. In particular, the sample from the x's consists of three pooled LHSs rather than a

random sample (see Eq. (5.16)). A Monte Carlo simulation can be used to assess if the use of formal statistical procedures to determine p-values is producing misleading results. Specifically, a large number of samples (10,000 in this example) of the form

$$(x_k, y_k), k = 1, 2, \ldots, 300, \tag{8.55}$$

can be generated by pairing the 300 values for x (i.e., the 300 values for the particular x under consideration contained in the samples indicated in Eq. (5.16)) with the 300 predicted values for y (i.e., the 300 values for y that resulted from the use of the sample elements indicated in Eq. (5.16)). The specific pairing algorithm used was to randomly and without replacement assign an x value to each y value, which is similar to bootstrapping (Efron and Tibshirani 1993) except that the sampling is being performed without replacement. This random assignment was repeated 10,000 times to produce 10,000 samples of the form in Eq. (8.55). Each of the 10,000 samples can be used to calculate the value of the χ^2-statistic. The resulting empirical distribution of the χ^2-statistic can then be used to estimate the p-value for the χ^2-statistic actually observed in the analysis. Comparison of the p-value obtained from Eq. (8.51) with the p-value obtained from the empirical distribution provides an indication of the robustness of the variable rankings with respect to possible deviations from the assumptions underlying the formal statistical procedure in Eq. (8.51).

As indicated by comparing the results in columns SI: 5 × 5 and SIMC: 5 × 5 in Table 8.13, the analytical determination of p-values in Eq. (8.51) and the just described Monte Carlo determination of p-values are producing similar results. Thus, at least in this example, the variable rankings are not being adversely impacted by the use of Eq. (8.51). Similar comparisons were also obtained in the analyses for y with CMNs, CMDs and CLs (Table 24, Kleijnen and Helton 1999a; see Tables 8, 14 and 19 of Kleijnen and Helton 1999a, for comparisons with additional variables).

Table 8.13. Comparison of Variable Rankings with the χ^2-statistic for y = *E2:WAS_PRES* at 10,000 yr Obtained with a Maximum of Five Classes of x Values (i.e., $nX = 5$) and Analytic Determination of p-Values with Variable Rankings Obtained with (i) a Maximum of Ten Classes of x values (i.e., $nX = 10$) and Analytic Determination of p-Values and (ii) a Maximum of Five Classes of x Values (i.e., $nX = 5$) and Monte Carlo Determination of p-Values (Table 23, Kleijnen and Helton 1999a; see Table 10.23, Kleijnen and Helton 1999c, for omitted results)

Variable	SI: 5 × 5[b]		SI: 10 × 10[c]		SIMC: 5 × 5[d]	
Name[a]	Rank	p-Val	Rank	p-Val	Rank	p-Val
BHPRM	1.0	0.0000	1.0	0.0000	1.5	0.0000
HALPRM	2.0	0.0002	4.0	0.0082	1.5	0.0000
WGRCOR	3.0	0.0002	2.0	0.0028	3.0	0.0002
ANHPRM	4.0	0.0049	3.0	0.0032	4.0	0.0033
SHRGSSAT	5.0	0.0698	22.0	0.8482	5.0	0.0699
SHBCEXP	6.0	0.1010	15.0	0.3495	6.0	0.0989
WGRMICI	7.0	0.1985	11.0	0.1646	7.0	0.2013
ANHBCVGP	8.0	0.2427	14.0	0.3398	8.0	0.2380
			...			
SHPRMHAL	24.0	0.9064	24.0	0.8863	24.0	0.9102
SHPRMCON	25.0	0.9898	20.0	0.5316	25.0	0.9933

[a] Twenty-six (25) variables from Table 3.1 included in analysis, with (i) *ANHCOMP* and *HALCOMP* not included because of the −0.99 rank correlations within the pairs (*ANHPRM*, *ANHCOMP*) and (*HALPRM*, *HALCOMP*) and (ii) *BPCOMP*, *BPINTPRS*, *BPPRM* and *BPVOL* not included because brine pocket properties are not relevant to the E2 intrusion under consideration.

[b] Variable rankings and p-values obtained with a maximum of five classes of x values (i.e., $nX = 5$), five classes of y values (i.e., $nY = 5$) and analytic determination of p-values (see Eqs. (8.54) and (8.51)). Discrete variables (e.g., *WMICDFLG*, which has only three distinct values) are divided into less than nX classes when they have less than nX distinct values.

[c] Variable rankings and p-values obtained with a maximum of ten classes of x values (i.e., $nX = 10$), ten classes of y values (i.e., $nY = 10$) and analytic determination of p-values.

[d] Variable rankings and p-values obtained with a maximum of five classes of x values (i.e., $nX = 5$), five classes of y values (i.e., $nY = 5$) and Monte Carlo determination of p-values.

9.0 Test Problems

Selected test problems from Campolongo et al. 2000 will now be used to illustrate sampling-based methods for uncertainty and sensitivity analysis. No attempt is made to present results for all test problems. Rather, the problems to be discussed were selected because they either provided representative results or interesting analysis challenges. To illustrate the effects of sampling procedures, each problem was evaluated with 10 independent LHSs of size 100 each and also 10 independent random samples of size 100 each. Sensitivity analysis results will be presented for 1 LHS of size 100 (i.e., *nLHS* = 100); in some instances, sensitivity analysis results will also be presented for the 1000 sample elements that result from pooling the 10 LHSs (i.e., *nLHS* = 1000). The sensitivity analysis procedures and/or measures considered will include correlation coefficients (CCs), rank correlation coefficients (RCCs), common means (CMNs), common locations (CLs), common medians (CMDs), statistical independence (SI), standardized regression coefficients (SRCs), partial correlation coefficients (PCCs), standardized rank regression coefficients (SRRCs), partial rank correlation coefficients (PRCCs), stepwise regression analysis with raw and rank-transformed data, and examination of scatterplots. It is hoped that the presentation of these results will help the reader develop insights with respect to the behavior and effectiveness of the techniques under consideration.

9.1 Linear Test Problems

The first linear test problem (Model 1 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = \sum_{i=1}^{3} x_i, \quad \mathbf{x} = [x_1, x_2, x_3], \tag{9.1}$$

with $x_i : U(\overline{x}_i - \sigma_i, \overline{x}_i + \sigma_i)$, $\overline{x}_i = 3^{i-1}$, $\sigma_i = 0.5\,\overline{x}_i$ for i = 1, 2, 3, and x:$U(a, b)$ used to indicate that x has a uniform distribution on $[a, b]$.

The distributions assigned to the x_i lead to a distribution for $f(\mathbf{x})$, with Latin hypercube sampling tending to produce more stable estimates of this distribution than random sampling (Fig. 9.1).

For the *nLHS* = 100, CCs, RCCs, CMNs, CLs, CMDs and SI all identify x_3 as the most important variable; CCs and RCCs also indicate an effect for x_2 (Table 9.1). Due to the large size of the p-values (i.e., > 0.05), CMNs, CLs, CMDs and SI do not indicate an effect for x_2, and none of the tests indicate an effect for x_1.

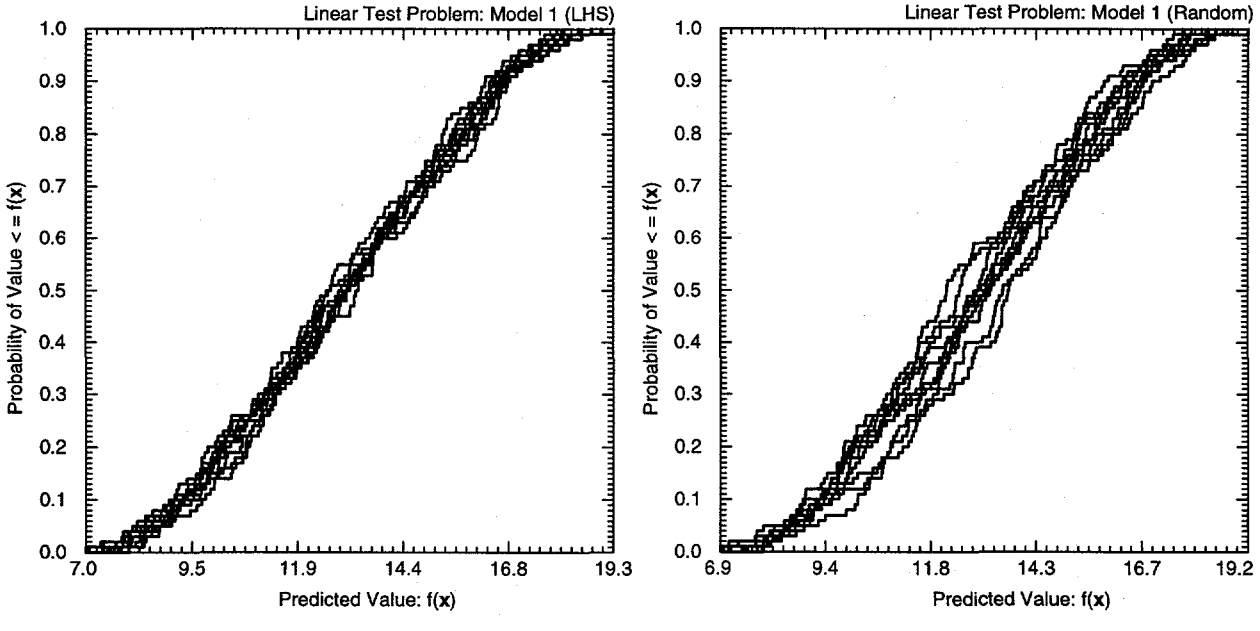


Fig. 9.1. Stability of estimated CDF for linear test problem with Model 1 (see Eq. (9.1)).

For *nLHS* = 1000, CCs, RCCs, CMNs, CLs, CMDs and SI identify x_3 and x_2 as the two most important variables (Table 9.1). Further, CCs, RCCs, CMNs and CLs also indicate an effect for x_1. Thus, as might be expected, the larger sample is leading to more resolution in the sensitivity analysis. However, CCs and RCCs were able to identify the two most important variables with a sample of size 100.

Examination of scatterplots clearly shows the dominant effect of x_3 (Fig. 9.2). The effect of x_2 is barely discernible in the scatterplot for *nLHS* = 100 but is easily seen for *nLHS* = 1000. The scatterplots for x_1 (not shown) indicate no visually discernible effect for *nLHS* = 100 and a barely discernible effect for *nLHS* = 1000.

In addition to various tests of significance (Table 9.1) and the examination of scatterplots (Fig. 9.2), various coefficient values (e.g., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) can also be used to assess variable importance (Table 9.2). In Table 9.2 and other similar tables in this section, CCs and RCCs are calculated between individual pairs of variables, and SRCs and SRRCs are calculated with all sampled variables included in the regression model (i.e., x_1, x_2, x_3 in this example; see Eq. (9.1)). In the complete absence of correlations between the sampled variable values, corresponding CCs and SRCs would be the same and so would corresponding RCCs and SRRCs. As indicated by the similarity of the values for CCs and SRCs and also for RCCs and SRRCs, there is little correlation between the sampled variables. Further, because an exact linear model is under consideration, PCCs and PRCCs are equal to one. Thus, for a linear model, PCCs and PRCCs provide no information on the importance of individual variables. Because of the linearity of the model, the sample of size *nLHS* = 1000 gives results almost identical to those in Table 9.2 for *nLHS* = 100.

An alternative summary of the SRCs and SRRCs in Table 9.2 is to present the sensitivity results in the form of a stepwise regression analysis (Table 9.3). Then, variable importance is indicated by the order in which the variables entered the regression model, the sizes and signs of the SRCs or SRRCs, and the changes in R^2 values as additional variables are added to the regression model. Because a linear model is under consideration, the stepwise process ultimately produces a regression model with an R^2 value of 1.00. However, the last variable added to the regression model (i.e., x_1) has little effect and only raises the R^2 value from 0.99 to 1.00. The regression coefficients do not provide

Table 9.1. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Linear Test Problem with Model 1 (see Eq. (9.1))

Sample Size: *nLHS* = 100

Variable	CC[b]		RCC[c]		CMN[d]		CL[e]		CMD[f]		SI[g]	
Name[a]	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_3	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_2	2.0	0.0015	2.0	0.0027	2.0	0.0502	2.0	0.0779	2.0	0.5249	2.0	0.2954
x_1	3.0	0.5091	3.0	0.5694	3.0	0.7528	3.0	0.7089	3.0	0.7358	3.0	0.8392

Sample Size: *nLHS* = 1000

Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_3	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_2	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000
x_1	3.0	0.0007	3.0	0.0017	3.0	0.0155	3.0	0.0313	3.0	0.4748	3.0	0.1164

[a] Variables ordered by *p*-values for CCs.
[b] Ranks and *p*-values for CCs.
[c] Ranks and *p*-values for RCCs.
[d] Ranks and *p*-values for CMNs test with 1×5 grid.
[e] Ranks and *p*-values for CLs (Kruskal-Wallis) test with 1×5 grid.
[f] Ranks and *p*-values for CMDs text with 2×5 grid.
[g] Ranks and *p*-values for SI test with 5×5 grid.

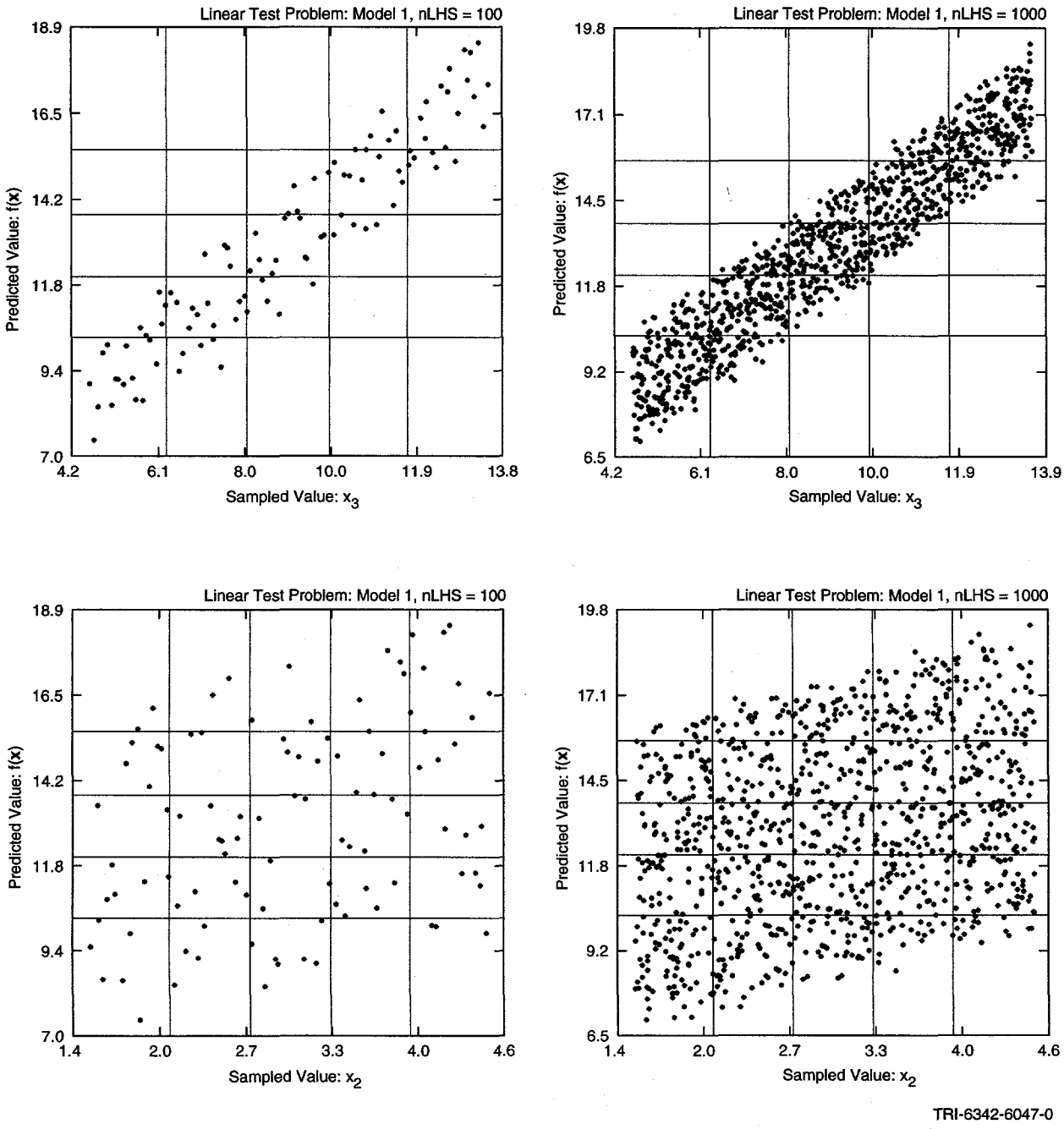


Fig. 9.2. Scatterplots for linear test problem with Model 1 (see Eq. (9.1)).

Table 9.2. Sensitivity Results Based on Coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) and Sample Size *nLHS* = 100 for Linear Test Problem with Model 1 (see Eq. (9.1))

Variable	CC[b]		SRC[b]		PCC[b]		RCC[b]		SRRC[b]		PRCC[b]	
Name[a]	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value
x_3	1	0.9439	1	0.9459	2	1.000	1	0.9466	1	0.9482	2	1.000
x_2	2	0.3175	2	0.3156	2	1.000	2	0.3018	2	0.2987	2	1.000
x_1	3	0.0660	3	0.1054	2	1.000	3	0.0572	3	0.0976	2	1.000

[a] Variables ordered by *p*-values for CCs.
[b] Ranks and values for CCs, SRCs, PCCs, RCCs, SRRCs and PRCCs as indicated.

Table 9.3. Sensitivity Results Based on Stepwise Regression Analysis with Raw (i.e., Untransformed) Data and Sample Size *nLHS* = 100 for Linear Test Problem with Model 1 (see Eq. (9.1))

Variable[a]	R^{2b}	RC[c]	SRC[d]	*p*-Value[e]
x_3	0.89098	1.0000E+00	9.4588E–01	0.0000E+00
x_2	0.98891	1.0000E+00	3.1558E–01	0.0000E+00
x_1	1.00000	1.0000E+00	1.0541E–01	0.0000E+00

[a] Variables in order of entry into regression model.
[b] Cumulative R^2 value with entry of each variable into regression model; see Eq. (8.11).
[c] Regression coefficients (RCs); see Eq. (8.8).
[d] Standardized regression coefficients (SRCs); see Eq. (8.15).
[e] For variable in row (i.e., x_j), *p*- or α-value for addition of x_j to regression model containing remaining variables; see Footnote n, Table 8.1.

information on variable importance (i.e., they are all 1.00); rather, it is the SRCs that provide an indication of variable importance. The results in Table 9.3 are for raw data; use of rank-transformed data produces similar results.

When a linear relationship exists between a predicted variable and multiple input variables, stepwise regression analysis provides more information on variable importance than simply examining correlation coefficients. First, the changes in R^2 values as additional variables are added to the regression model provides an indication of how much uncertainty can be accounted for by each variable. For example, the R^2 values produced with the addition of each variable to the regression model in Table 9.3 are 0.89, 0.99 and 1.00, respectively. Thus, the last variable selected (i.e., x_1) only changes the R^2 value from 0.99 to 1.00. Second, the *F*-test for the sequential addition of variables to the regression model is more sensitive than the test for the significance of a single CC. For example, the *p*-value obtained with *nLHS* = 100 for the CC associated with x_1 is 0.5091 (Table 9.1); in contrast, the *p*-value for the entry of x_1 into the regression model that already contains x_3 and x_2 is less than 10^{-4}.

The second linear test problem (Model 3 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = \sum_{i=1}^{22} c_i (x_i - 1/2), \qquad \mathbf{x} = [x_1, x_2, \ldots, x_{22}], \quad (9.2)$$

with $x_i : U(0, 1)$ and $c_i = (i - 11)^2$ for $i = 1, 2, \ldots, 22$.

Latin hypercube and random sampling produce estimates of similar stability for the CDF for $f(\mathbf{x})$ (Fig. 9.3). This is different from the first linear function, where Latin hypercube sampling produced more stable estimates (Fig. 9.1). This stability probability results from the fact that the model can be written as

$$f(\mathbf{x}) = c_{22}(x_{22} - 1/2) + \sum_{i=1}^{10} c_i \left[(x_i - 1/2) + (x_{22-i} - 1/2)\right], \quad (9.3)$$

which tends to smooth the effects of the random sampling.

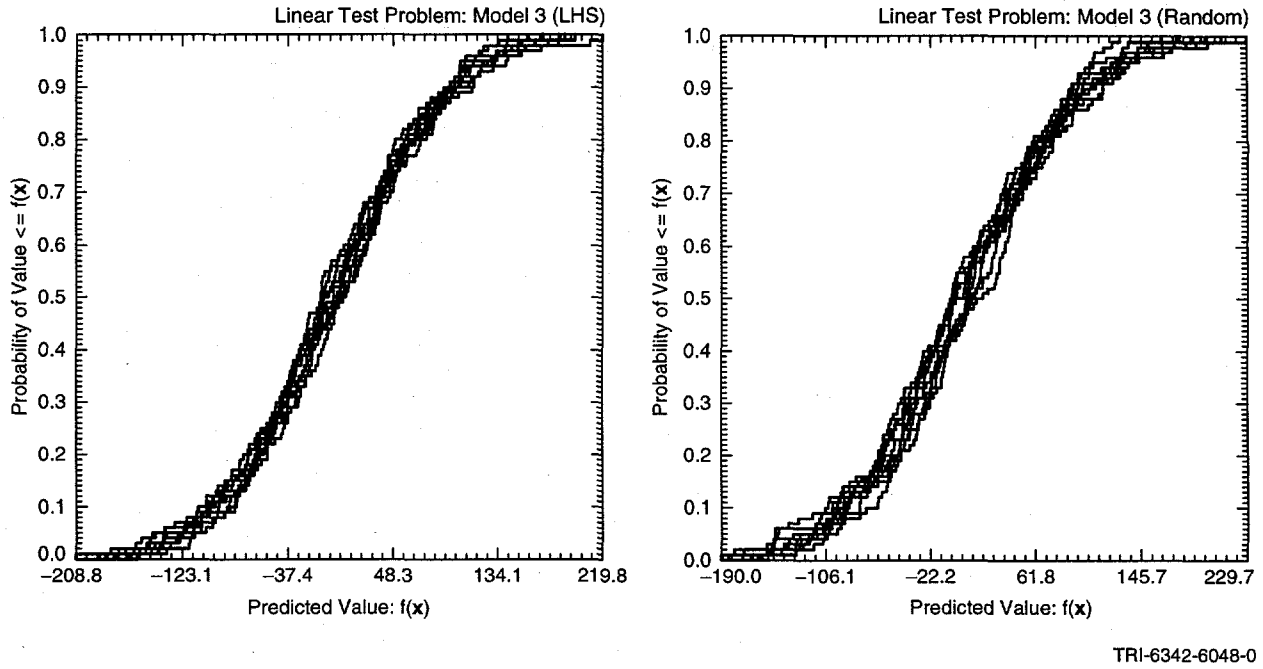


Fig. 9.3. Stability of estimated CDF for linear test problem with Model 3 (see Eq. (9.2)).

For the LHS of size *nLHS* = 100, CCs and RCCs identify the same variables as affecting f (i.e., x_{22}, x_{21}, x_1, x_{20}, x_3, x_2, x_{19}, x_{18}, x_4 with p-values less than 0.05) (Table 9.4). Similar identifications are also made for CMNs and CLs; in contrast, CMDs and SI fail to identify some of the variables identified by CCs and RCCs. For the LHS of size *nLHS* = 1000, all tests identify more variables as affecting f (Table 9.4). Further, there is more agreement between the tests on the most important variables (i.e., smallest p-values). However, a number of variables are not identified as having an effect on f by any of the tests (e.g., x_7, x_{15}, x_{14}, x_8, x_9, x_{12}, x_{13}, x_{11}, x_{10} have p-values greater than 0.05 for most tests).

Given that a linear model is under consideration, stepwise regression provides a more informative summary of variable effects than the coefficients in Table 9.4 (Table 9.5). In particular, the stepwise regression analysis with *nLHS* = 100 identifies the effects of all 21 variables that influence the evaluation of f (i.e., all variables except x_{11}, which has a coefficient of zero). The results for *nLHS* = 1000 (not shown) are essentially identical with those for *nLHS* = 100; thus, no improvement in the results of the stepwise regression analysis is obtained by increasing the sample size. Thus, the tests of significance used with the stepwise regression analysis are more effective in identifying the effects of individual variables than the tests used in conjunction with Table 9.4. In particular, the stepwise regression in Table 9.5 correctly identifies the effects of all variables influencing f with a sample of size *nLHS* = 100; the test based on CCs in Table 9.4 does not identify the effects of all variables with a sample of size *nLHS* = 1000 (i.e., same variables have p-values greater than 0.1).

The cumulative R^2 values with the entry of each variable into the regression model are shown in Table 9.5. The increase in the R^2 value with the entry of a variable shows the fraction of the total uncertainty that can be accounted for by that variable in a linear regression model (e.g., x_{21} accounts for a fraction 0.36279 – 0.20948 = 0.15331 of the total uncertainty). As indicated by the incremental R^2 values, no single variable dominates the uncertainty in f.

For perspective, scatterplots for the first two variables selected in the stepwise process (i.e., x_{22}, x_{21}) are shown in Fig. 9.4. Although the patterns are discernible, they are not strong, which is consistent with the incremental R^2 values of 0.20948 and 0.15311 associated with x_{22} and x_{21}.

Both regression coefficients and SRCs are given in Table 9.5. The SRCs are a better measure of variable importance because they incorporate the effects of a variable's distribution and also remove the effects of units. Except for the effects of correlations within a

Table 9.4. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Linear Test Problem with Model 3 (see Eq. (9.2))[a]

Sample Size: $nLHS = 100$

Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val
x_{22}	1.0	0.0000	1.0	0.0000	1.0	0.0001	1.0	0.0002	1.5	0.0009	2.0	0.0208
x_{21}	2.0	0.0001	2.0	0.0002	2.0	0.0004	4.0	0.0018	9.0	0.1074	1.0	0.0086
x_1	3.0	0.0002	3.0	0.0003	5.0	0.0024	5.0	0.0043	7.0	0.0289	8.5	0.1137
x_{20}	4.0	0.0003	4.0	0.0005	7.0	0.0070	7.0	0.0131	11.0	0.2311	5.0	0.0615
x_3	5.0	0.0015	5.0	0.0016	6.0	0.0064	6.0	0.0086	4.0	0.0103	3.0	0.0239
x_2	6.0	0.0028	7.0	0.0037	3.0	0.0006	2.0	0.0012	3.0	0.0051	6.0	0.0791
x_{19}	7.0	0.0037	6.0	0.0025	10.0	0.0699	10.0	0.0445	5.0	0.0123	10.5	0.1785
x_{18}	8.0	0.0238	8.0	0.0197	8.0	0.0318	8.0	0.0294	6.0	0.0146	12.0	0.2202
x_4	9.0	0.0444	9.0	0.0289	9.0	0.0399	9.0	0.0295	10.0	0.1991	8.5	0.1137
x_{17}	10.0	0.1095	10.0	0.1135	12.0	0.1476	12.0	0.1515	15.5	0.4060	10.5	0.1785
x_{16}	11.0	0.1379	11.0	0.1154	20.0	0.7358	18.0	0.6699	13.5	0.3546	4.0	0.0316
x_5	12.0	0.2668	12.0	0.3349	4.0	0.0012	3.0	0.0016	1.5	0.0009	7.0	0.1010
x_6	13.0	0.4991	13.0	0.4195	18.0	0.6822	19.0	0.6835	18.0	0.5249	14.0	0.4186
x_9	14.0	0.5118	18.0	0.6595	17.0	0.3711	16.0	0.4258	18.0	0.5249	15.5	0.4884
x_7	15.0	0.5261	16.0	0.5194	19.0	0.7351	20.0	0.7596	15.5	0.4060	18.0	0.5987
x_8	16.0	0.5368	15.0	0.5006	14.0	0.3476	17.0	0.4307	18.0	0.5249	13.0	0.3239
x_{12}	17.0	0.5487	14.0	0.4632	16.0	0.3656	14.0	0.3570	20.0	0.7358	18.0	0.5987
x_{13}	18.0	0.7118	17.0	0.6491	13.0	0.2392	13.0	0.2676	13.5	0.3546	15.5	0.4884
x_{14}	19.0	0.8221	21.0	0.9223	21.0	0.9511	21.0	0.9651	22.0	0.9825	18.0	0.5987
x_{15}	20.0	0.8317	20.0	0.7924	22.0	0.9922	22.0	0.9929	21.0	0.9384	20.0	0.6359
x_{11}	21.0	0.8909	19.0	0.7495	11.0	0.0716	11.0	0.1020	8.0	0.0404	21.0	0.7440
x_{10}	22.0	0.9217	22.0	0.9840	15.0	0.3507	15.0	0.3963	12.0	0.3084	22.0	0.7776

Sample Size: $nLHS = 1000$

Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val
x_{22}	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_{21}	2.0	0.0000	2.0	0.0000	2.0	0.0000	3.0	0.0000	2.0	0.0000	3.0	0.0000
x_1	3.0	0.0000	3.0	0.0000	3.0	0.0000	2.0	0.0000	3.0	0.0000	2.0	0.0000
x_2	4.0	0.0000	4.0	0.0000	4.0	0.0000	4.0	0.0000	6.0	0.0000	4.0	0.0000
x_{20}	5.0	0.0000	5.0	0.0000	5.0	0.0000	5.0	0.0000	4.0	0.0000	5.0	0.0000
x_3	6.0	0.0000	6.0	0.0000	6.0	0.0000	6.0	0.0000	5.0	0.0000	6.0	0.0000
x_{19}	7.0	0.0000	7.0	0.0000	7.0	0.0000	7.0	0.0000	7.0	0.0000	7.0	0.0000
x_{18}	8.0	0.0000	8.0	0.0000	8.0	0.0000	8.0	0.0000	8.0	0.0000	9.0	0.0001
x_4	9.0	0.0000	9.0	0.0000	9.0	0.0000	9.0	0.0000	9.0	0.0000	10.0	0.0003
x_5	10.0	0.0000	11.0	0.0000	10.0	0.0000	10.0	0.0000	10.0	0.0002	8.0	0.0000
x_{17}	11.0	0.0000	10.0	0.0000	11.0	0.0002	11.0	0.0002	12.0	0.0040	12.0	0.0121
x_{16}	12.0	0.0011	12.0	0.0002	12.0	0.0124	12.0	0.0035	11.0	0.0004	13.0	0.1164
x_6	13.0	0.0018	13.0	0.0014	13.0	0.0252	13.0	0.0212	13.0	0.0206	11.0	0.0019
x_7	14.0	0.0637	14.0	0.0776	14.0	0.0267	14.0	0.0697	16.0	0.1538	14.0	0.1164
x_{15}	15.0	0.0959	15.0	0.0892	22.0	0.6771	20.0	0.5827	21.0	0.7431	21.0	0.6691
x_{14}	16.0	0.2579	18.0	0.3909	18.0	0.4664	19.0	0.5414	18.0	0.4809	15.0	0.1843
x_8	17.0	0.3165	16.0	0.2949	19.0	0.5583	18.0	0.4750	14.0	0.0425	17.0	0.2509
x_9	18.0	0.3907	19.0	0.4178	20.0	0.5701	21.0	0.7113	22.0	0.9437	18.0	0.2899
x_{12}	19.0	0.4606	17.0	0.3616	16.0	0.3026	16.0	0.1976	15.0	0.1402	16.0	0.1944
x_{13}	20.0	0.4625	21.0	0.5867	15.0	0.1438	15.0	0.1348	17.0	0.2792	19.0	0.2954
x_{11}	21.0	0.6626	20.0	0.5806	17.0	0.3446	17.0	0.3143	19.0	0.4932	20.0	0.4530
x_{10}	22.0	0.7892	22.0	0.7117	21.0	0.6605	22.0	0.7753	20.0	0.5512	22.0	0.9950

[a] Table structure same as in Table 9.1.

Table 9.5. Sensitivity Results Based on Stepwise Regression Analysis with Raw (i.e., Untransformed) Data and Sample Size *nLHS* = 100 for Linear Test Problem with Model 3 (see Eq. (9.2))[a]

Variable	R^2	RC	SRC	*p*-Value
x_{22}	0.20948	1.2100E+02	4.6052E–01	2.7828E–08[b]
x_{21}	0.36279	1.0000E+02	3.8038E–01	2.7828E–08
x_1	0.50981	1.0000E+02	3.8141E–01	2.7828E–08
x_{20}	0.63339	8.1000E+01	3.0763E–01	2.7828E–08
x_2	0.73563	8.1000E+01	3.0830E–01	2.7828E–08
x_3	0.80541	6.4000E+01	2.4338E–01	2.7828E–08
x_{19}	0.86382	6.4000E+01	2.4317E–01	2.7828E–08
x_{18}	0.90285	4.9000E+01	1.8642E–01	2.7828E–08
x_4	0.93449	4.9000E+01	1.8614E–01	2.7828E–08
x_5	0.95728	3.6000E+01	1.3677E–01	2.7828E–08
x_{17}	0.97297	3.6000E+01	1.3665E–01	2.7828E–08
x_6	0.98146	2.5000E+01	9.5070E–02	2.7828E–08
x_{16}	0.98978	2.5000E+01	9.5121E–02	2.7828E–08
x_{15}	0.99340	1.6000E+01	6.0789E–02	2.7828E–08
x_7	0.99710	1.6000E+01	6.0905E–02	2.7828E–08
x_8	0.99833	9.0000E+00	3.4256E–02	2.7828E–08
x_{14}	0.99950	9.0000E+00	3.4263E–02	2.7828E–08
x_9	0.99974	4.0000E+00	1.5206E–02	2.7828E–08
x_{13}	0.99997	4.0000E+00	1.5225E–02	2.7828E–08
x_{10}	0.99999	9.9999E–01	3.8041E–03	2.7828E–08
x_{12}	1.00000	1.0000E+00	3.8018E–03	2.7828E–08
x_{11}	1.00000	–3.0113E–05	–1.1426E–07	2.6792E–01

[a] Table structure same as in Table 9.3.
[b] Identical values result from lack of resolution in algorithm used in the calculation of very small *p*-values.

sample, CCs and SRCs are the same; thus, the CCs between the x_i and $f(\mathbf{x})$ are also available from Table 9.5. For example, Fig. 9.4 contains scatterplots with associated CCs of approximately 0.46052 for x_{22} and 0.38038 for x_{21}.

9.2 Monotonic Test Problems

The first monotonic test problem (Model 4 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = x_1 + x_2^4, \quad \mathbf{x} = [x_1, x_2], \tag{9.4}$$

with $x_i : U(0, 1)$ for $i = 1, 2$ (Model 4a), $x_i : U(0, 3)$ for $i = 1, 2$ (Model 4b) or $x_i : U(0, 5)$ for $i = 1, 2$ (Model 4c). Thus, f is the same in Models 4a, 4b, and 4c, but the distributions assigned to the x_i change. In the following, Models 4a and 4c will be considered as this incorporates the two extremes in the behavior of f.

Latin hypercube sampling produces more stable estimates of the CDFs for Models 4a and 4c than is the case for random sampling (Fig. 9.5). This stability is particularly noticeable for Model 4c, where the value of $f(\mathbf{x})$ is dominated by a strong nonlinear relationship involving x_2; in this problem, the stratification associated with Latin hypercube sampling produces CDF estimates that are much more stable than those obtained with random sampling.

Sensitivity analysis for Model 4c is not very interesting due to the dominance of $f(\mathbf{x})$ by x_2 (Fig. 9.6), with the result that all of the sensitivity analysis procedures under consideration identify x_2 as the dominant variable. Sensitivity analysis is more interesting for Model 4a as both x_1 and x_2 affect $f(\mathbf{x})$. Therefore, only sensitivity analysis for Model 4a will be discussed.

All procedures identify x_1 and x_2 as affecting $f(\mathbf{x})$ for Model 4a and the sample of size *nLHS* = 100 (Table 9.6). The well-defined effects of x_1 and x_2 can be seen in the corresponding scatterplots (Fig. 9.7). The patterns are better defined in the scatterplots for *nLHS* = 1000 but still easily recognizable in the scatterplots for *nLHS* = 100.

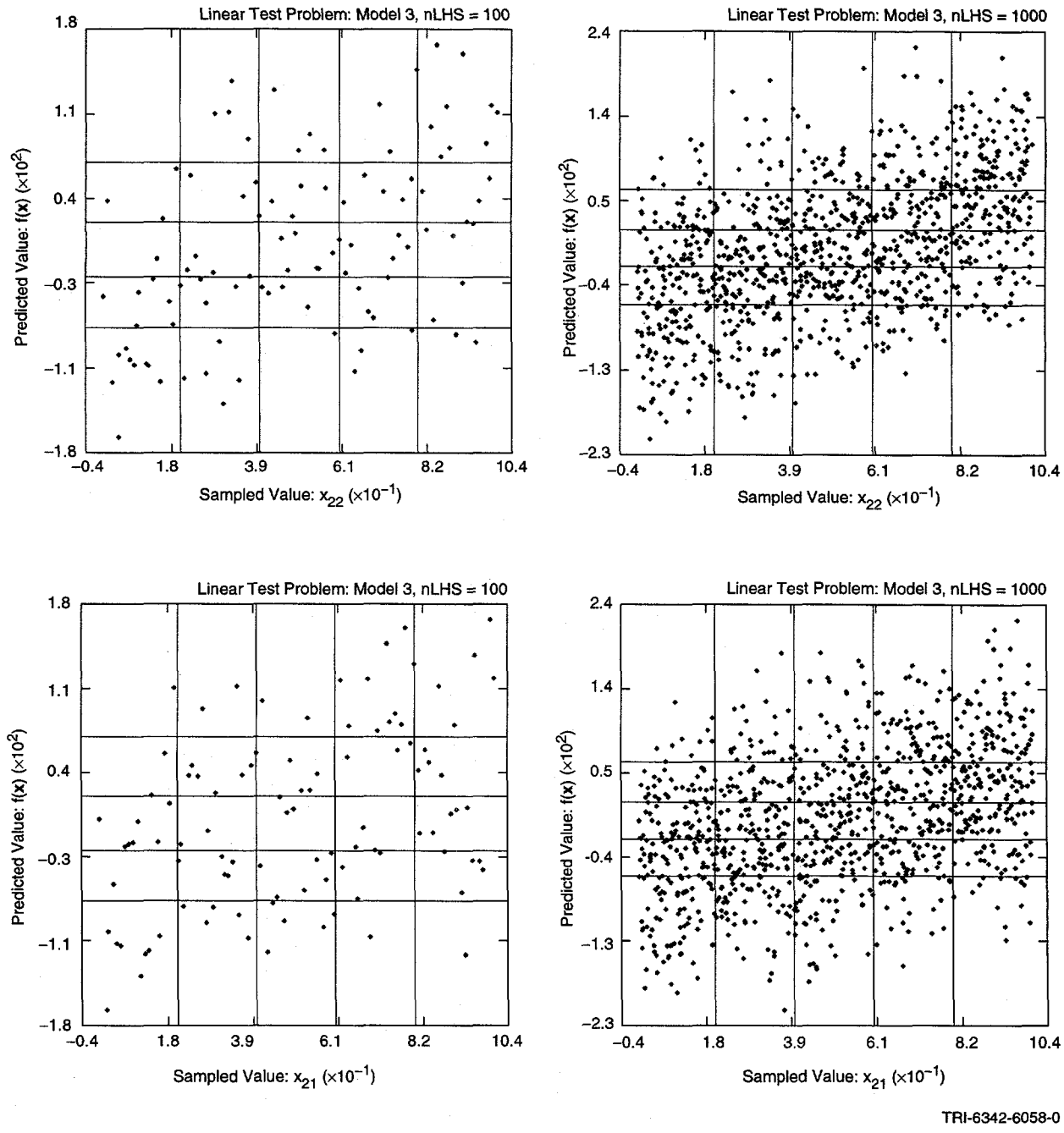


Fig. 9.4. Scatterplots for linear test problem with Model 3 (see Eq. (9.2)).

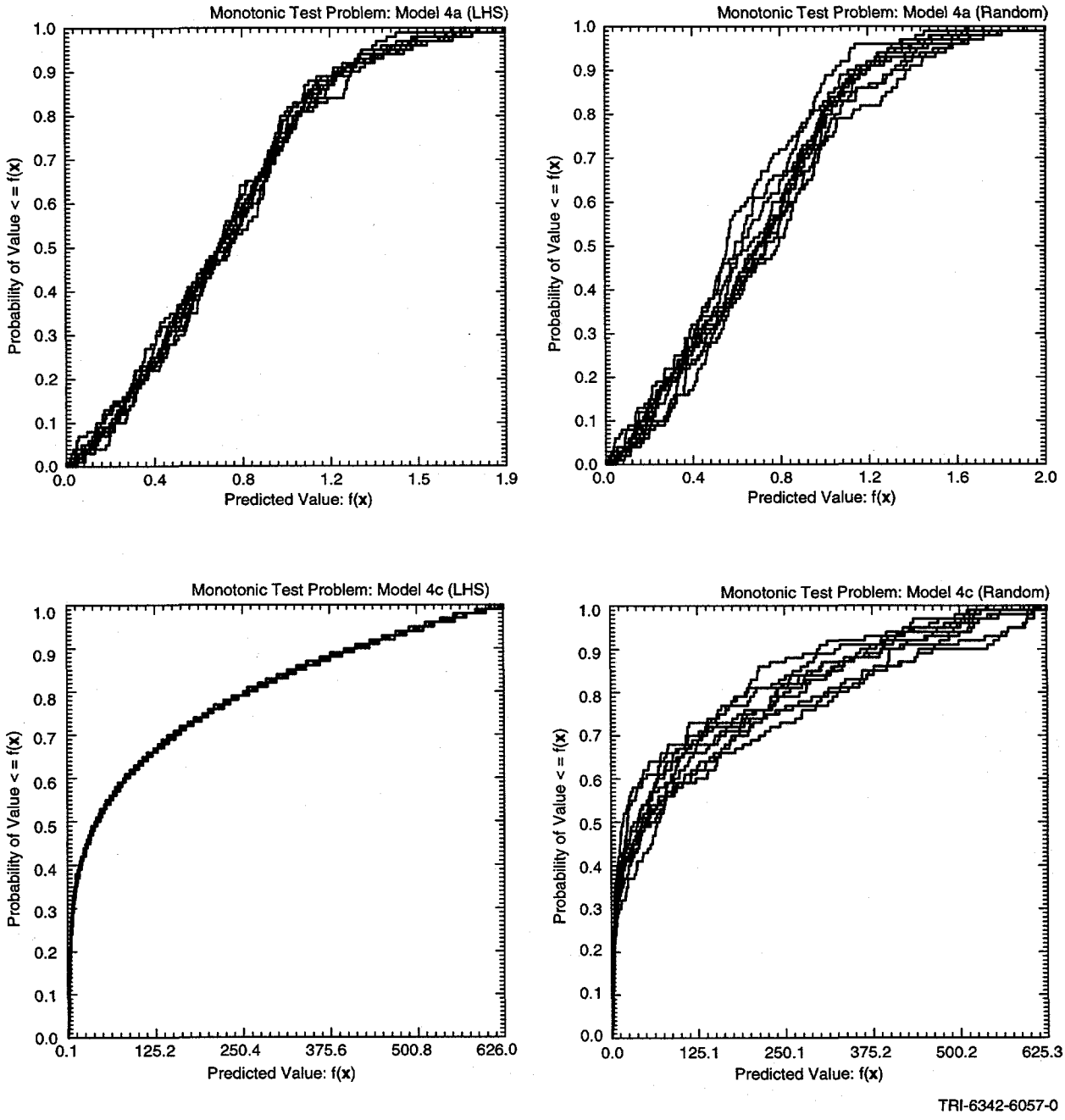


Fig. 9.5. Stability of estimated CDFs for monotonic test problem with Models 4a and 4c (see Eq. (9.4)).

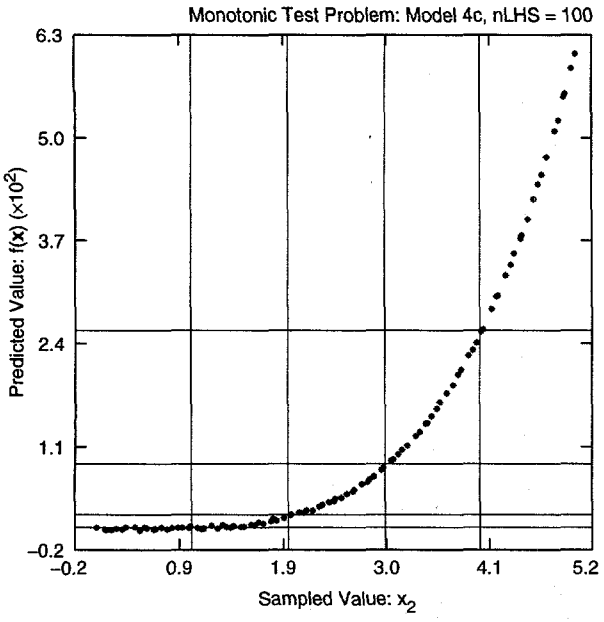


Fig. 9.6. Scatterplot with *nLHS* = 100 for monotonic test problem with Model 4c (see Eq. (9.4)).

For perspective, various coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) involving x_1, x_2 and $f(\mathbf{x})$ are presented in Table 9.7. As should be the case, CCs and SRCs are similar in size and PCCs are larger than CCs and SRCs; similar patterns also hold for RCCs, SRRCs and PRCCs. In this example, the coefficients calculated with raw (i.e., untransformed) data have values that are similar to those of the corresponding coefficients calculated with rank-transformed data. Thus, the problem is not as nonlinear over the distributions of x_1 and x_2 as might be suggested by the definition of f in Eq. (9.4), which is consistent with the linear trends appearing in the scatterplots in Fig. 9.7. The use of samples of size *nLHS* = 100 and *nLHS* = 1000 produce similar coefficient values. Thus, the behavior of the function is being adequately captured with *nLHS* = 100, and little is gained by using a large sample size (although the scatterplots are more visually appealing for *nLHS* = 1000).

The sensitivity results for Model 4a can also be summarized as the outcome of a stepwise regression analysis (Table 9.8). As already observed in conjunction with Table 9.7, x_1 is identified as having a stronger effect on the uncertainty in $f(\mathbf{x})$ than x_2, and analyses with raw (i.e., untransformed) data and rank-transformed data produce similar results. Use of the sample of size *nLHS* = 1000 produces little improvement in the regression analyses, with R^2 values for the final regression model changing from 0.88580 and 0.87966 with raw and rank-transformed data with *nLHS* = 100 to 0.88356 and 0.88482 for *nLHS* = 1000 (regressions not shown). Thus, as previously noted, increasing the sample size in this example does not improve the results of the sensitivity analysis.

The use of regression analysis with rank-transformed data rather than raw data produced no improvement in the resultant regression model for Model 4a (Table 9.8). However, the potential exists for considerable improvement when the dependent variable is a nonlinear but monotonic function of the independent variable(s). For example, a nonlinear but monotonic relationship exists between x_2 and $f(\mathbf{x})$ for Model 4c (Fig. 9.6). In the analysis of this model, a regression with rank-transformed data relating $f(\mathbf{x})$ to x_2 with *nLHS* = 100 produces a regression model with an R^2 value of 0.97574; the corresponding regression with raw data produces a regression model with an R^2 value of 0.75003.

The second monotonic test problem (Model 5 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = \exp\left(\sum_{i=1}^{6} b_i x_i\right) - \prod_{i=1}^{6} (e^{b_i} - 1)/b_i,$$
$$\mathbf{x} = [x_1, x_2, \ldots, x_6], \tag{9.5}$$

Table 9.6. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Monotonic Test Problem with Model 4a (see Eq. (9.4)) and *nLHS* = 100[a]

Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_1	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_2	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0004	2.0	0.0000

[a] Table structure same as in Table 9.1.

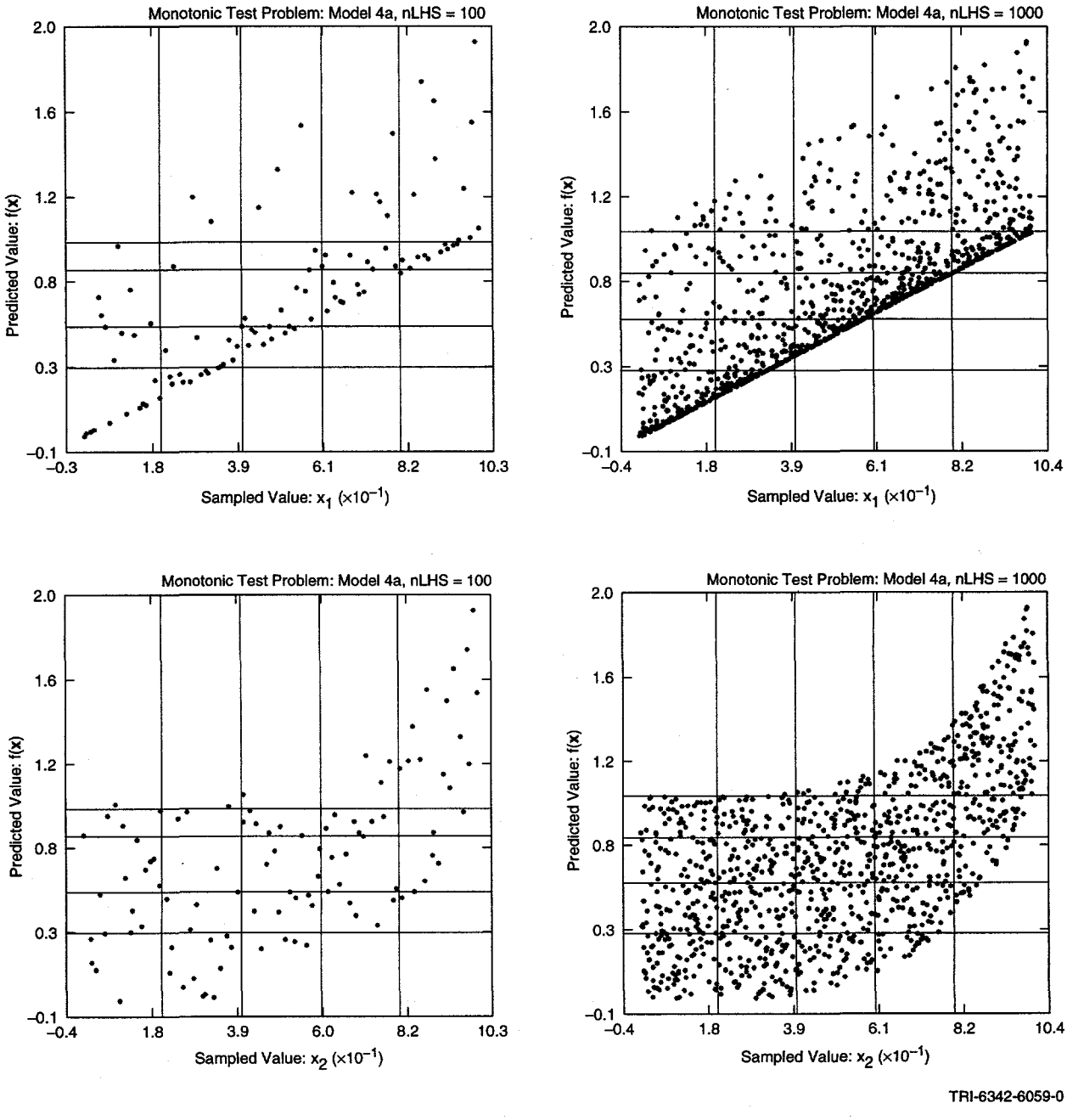


Fig. 9.7. Scatterplots for monotonic test problem with Model 4a (see Eq. (9.4)).

Table 9.7. Sensitivity Results Based on Coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) for Monotonic Test Problem with Model 4a (see Eq. (9.4))

Sample Size: *nLHS* = 100							
Variable	CC[b]			SRC[c]		PCC[c]	
Name[a]	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.0	0.7367	1.0	0.7401	1.0	0.9097
x_2	0.0000	2.0	0.5814	2.0	0.5857	2.0	0.8662

Variable	RCC[e]			SRRC[f]		PRCC[f]	
Name[d]	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.0	0.7688	1.0	0.7723	1.0	0.9122
x_2	0.0000	2.0	0.5322	2.0	0.5373	2.0	0.8401

Sample Size: *nLHS* = 1000							
Variable	CC			SRC		PCC	
Name	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.5	0.7310	1.0	0.7263	1.0	0.9051
x_2	0.0000	1.5	0.5967	2.0	0.5910	2.0	0.8660

Variable	RCC			SRRC		PRCC	
Name	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.5	0.7531	1.0	0.7489	1.0	0.9108
x_2	0.0000	1.5	0.5692	2.0	0.5637	2.0	0.8567

[a] Variables ordered by *p*-values for CCs.
[b] *p*-values, ranks and values for CCs.
[c] Ranks and values for SRCs and PCCs as indicated.
[d] Variables ordered by *p*-values for RCCs.
[e] *p*-values, ranks and values for RCCs.
[f] Ranks and values for SRRCs and PRCCs as indicated.

Table 9.8. Sensitivity Results Based on Stepwise Regression Analysis for Monotonic Test Problem with Model 4a (see Eq. (9.4)) and Sample Size *nLHS* = 100[a]

Raw Data				
Variable	R^2	RC	SRC	*p*-Value
x_1	0.54273	1.0070E+00	7.4014E–01	2.7828E–08
x_2	0.88580	7.9861E–01	5.8573E–01	2.7828E–08

Rank-Transformed Data				
Variable	R^2	RRC[b]	SRRC[c]	*p*-Value
x_1	0.59099	7.7229E–01	7.7229E–01	2.7828E–08
x_2	0.87966	5.3728E–01	5.3728E–01	2.7828E–08

[a] Table structure same as in Table 9.3.
[b] Rank regression coefficient (RRC).
[c] Standardized rank regression coefficient (SRRC).

with $b_1 = 1.5$, $b_2 = b_3 = \cdots = b_6 = 0.9$ and $x_i : U(0, 1)$ for $i = 1, 2, \ldots, 6$.

Latin hypercube sampling produces more stable estimates of the CDF for $f(\mathbf{x})$ than does random sampling (Fig. 9.8). However, the distribution has a long tail to the right, and both sampling procedures show considerable variation across replicates in the largest observed value for $f(\mathbf{x})$. Thus, if accurate estimates of the upper quantiles of the CDF were required, then it would be necessary to use a large sample size or possibly switch to an importance sampling procedure. For functions that are as inexpensive to evaluate as f, it would be wasteful to invest the effort to design an importance sampling procedure. However, as the cost of evaluating the function (i.e., model) increases, at some point use of importance sampling may become cost effective.

All tests (i.e., CCs, RCCs, CMNs, CLs, CMDs, SI) identify x_1 as the most important variable for $nLHS$ = 100 (Table 9.9); further, CCs and RCCs identify effects for all six x_i. Given the definition of f, x_1 is the most important variable with respect to the uncertainty in $f(\mathbf{x})$, and $x_2, x_3, \ldots, x_6$ have equal-sized effects on this uncertainty. For $nLHS = 1000$, all tests identify effects for all six x_i.

The coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) involving the x_i and $f(\mathbf{x})$ are presented in Table 9.10. The largest coefficients involve x_1; x_2, x_3, ..., x_6 have similar-sized coefficients; CCs and SRCs are essentially equal, as is the case for RCCs and SRRCs; PCCs and PRCCs are larger than the corresponding CCs and RCCs, respectively; and all coefficients are positive, which is consistent with the use of the x_i in the definition of $f(\mathbf{x})$. Samples of size $nLHS$ = 100 and $nLHS = 1000$ produce similar coefficient estimates.

The scatterplots for x_1 and x_2 show discernible, but not particularly strong, patterns (Fig. 9.9). As should be the case given the definition of $f(\mathbf{x})$ and the distributions assigned to the x_i, the scatterplots for x_1 show somewhat stronger patterns than the scatterplots for x_2. The scatterplots for x_3, x_4, x_5, x_6 are similar to those for x_2.

The sensitivity results for Model 5 can also be presented as stepwise regression analyses with raw and rank-transformed data (Table 9.11). The regression analyses with both raw and rank-transformed data identify the effects associated with all six x_i's. Further, the regression analyses with rank-transformed data produce models with higher R^2 values than the regression analyses with raw data (i.e., 0.94119 versus 0.74993 for $nLHS$ = 100 and 0.96285 versus 0.80030 for $nLHS$ = 1000). There is little difference in the regression results obtained with $nLHS$ = 100 and $nLHS$ = 1000 (not shown).

9.3 Nonmonotonic Test Problems

The first nonmonotonic test problem (Model 7 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = \prod_{i=1}^{8} g_i(x_i), \quad \mathbf{x} = [x_1, x_2, \ldots, x_8]$$
$$= \prod_{i=1}^{8} \frac{|4x_i - 2| + a_i}{1 + a_i} \qquad (9.6)$$

with $a_1 = 0$, $a_2 = 1$, $a_3 = 4.5$, $a_4 = 9$, $a_5 = a_6 = a_7 = a_8 = 99$, and x_i: $U(0, 1)$ for $i = 1, 2, \ldots, 8$.

Latin hypercube sampling produces estimates of the CDF for $f(\mathbf{x})$ that are more stable than those produced by random sampling (Fig. 9.10).

Tests based on CCs and RCCs fail to identify any of the x_i as affecting $f(\mathbf{x})$ for $nLHS$ = 100 and also for $nLHS$ = 1000 (Table 9.12). In contrast, tests based on CMNs, CLs, CMDs and SI identify significant effects for x_1 and x_2 for both $nLHS$ = 100 and $nLHS$ = 1000, with the exception that the SI test does not identify x_2 for $nLHS$ = 100. In addition, smaller effects are indicated for x_3 (CMN, CL, CMD) and x_4 (CMN, CL, CMD, SI) for $nLHS$ = 1000. Tests based on CCs and RCCs fail to identify the effects of x_1 and x_2 on $f(\mathbf{x})$ because these effects are both nonlinear and nonmonotonic (Fig. 9.11). In contrast, such effects are readily identified by CMNs, CLs, CMDs and SI. All the coefficients involving $f(\mathbf{x})$ and the x_i's (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) are essentially zero; similarly, the regression analyses with raw and rank-transformed data produce no meaningful results.

The second nonmonotonic test problem (Model 8 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = h(x_2) \sum_{i=0}^{[x_2/2]} c_i(x_2) g_i(x_1, x_2), \quad \mathbf{x} = [x_1, x_2], \quad (9.7)$$

where h, c_i and g_i are defined by

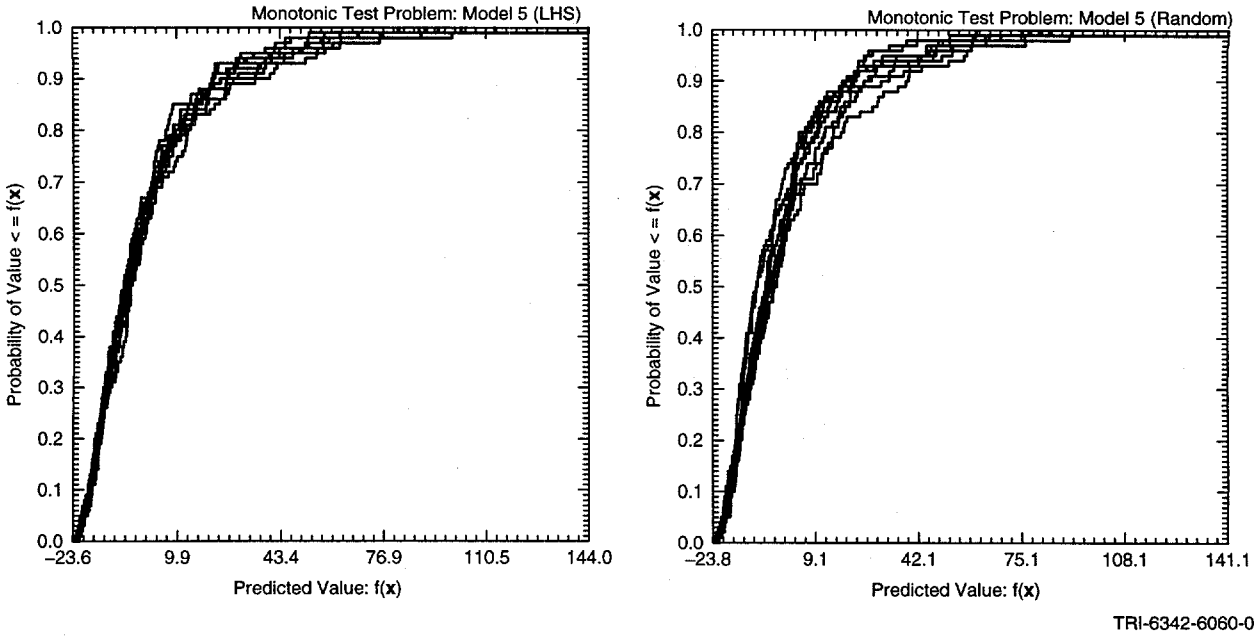


Fig. 9.8. Stability of estimated CDF for monotonic test problem with Model 5 (see Eq. (9.5)).

Table 9.9. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Monotonic Test Problem with Model 5 (see Eq. (9.5))[a]

Sample Size: $nLHS = 100$												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val
x_1	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_4	2.0	0.0005	5.0	0.0009	3.0	0.0161	4.0	0.0095	6.0	0.1468	5.0	0.2436
x_5	3.0	0.0007	6.0	0.0029	2.0	0.0006	2.0	0.0011	3.0	0.0342	2.0	0.0156
x_2	4.0	0.0041	4.0	0.0007	4.0	0.0211	5.0	0.0098	2.0	0.0087	3.0	0.0180
x_6	5.0	0.0051	3.0	0.0004	6.0	0.0840	6.0	0.0184	4.0	0.0477	6.0	0.4530
x_3	6.0	0.0052	2.0	0.0003	5.0	0.0464	3.0	0.0070	5.0	0.0780	4.0	0.0540
Sample Size: $nLHS = 1000$												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val
x_1	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_5	2.0	0.0000	6.0	0.0000	3.0	0.0000	4.0	0.0000	5.0	0.0000	5.0	0.0000
x_2	3.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000
x_4	4.0	0.0000	4.0	0.0000	5.0	0.0000	3.0	0.0000	3.0	0.0000	4.0	0.0000
x_3	5.0	0.0000	3.0	0.0000	4.0	0.0000	6.0	0.0000	6.0	0.0000	6.0	0.0000
x_6	6.0	0.0000	5.0	0.0000	6.0	0.0000	5.0	0.0000	4.0	0.0000	3.0	0.0000

[a] Table structure same as in Table 9.1.

Table 9.10. Sensitivity Results Based on Coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) for Monotonic Test Problem with Model 5 (see Eq. (9.5))[a]

Sample Size: *nLHS* = 100							
Variable	CC			SRC		PCC	
Name	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.0	0.5078	1.0	0.5223	1.0	0.7221
x_4	0.0005	2.0	0.3459	3.0	0.3446	3.0	0.5673
x_5	0.0007	3.0	0.3371	2.0	0.3509	2.0	0.5739
x_2	0.0041	4.0	0.2868	5.0	0.2952	5.0	0.5080
x_6	0.0051	5.0	0.2803	6.0	0.2837	6.0	0.4929
x_3	0.0052	6.0	0.2793	4.0	0.2973	4.0	0.5108

Variable	RCC			SRRC		PRCC	
Name	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.0	0.5852	1.0	0.6013	1.0	0.9273
x_3	0.0003	2.0	0.3596	2.0	0.3763	2.0	0.8404
x_6	0.0004	3.0	0.3591	3.0	0.3669	3.0	0.8339
x_2	0.0007	4.0	0.3405	4.0	0.3456	4.0	0.8183
x_4	0.0009	5.0	0.3334	5.0	0.3317	5.0	0.8071
x_5	0.0029	6.0	0.2992	6.0	0.3142	6.0	0.7912

Sample Size: *nLHS* = 1000							
Variable	CC			SRC		PCC	
Name	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.0	0.5259	1.0	0.5217	1.0	0.7594
x_5	0.0000	2.0	0.3412	2.0	0.3367	2.0	0.6017
x_2	0.0000	3.0	0.3297	4.0	0.3241	4.0	0.5871
x_4	0.0000	4.0	0.3275	3.0	0.3251	3.0	0.5882
x_3	0.0000	5.0	0.3274	5.0	0.3220	5.0	0.5846
x_6	0.0000	6.0	0.3032	6.0	0.3044	6.0	0.5629

Variable	RCC			SRRC		PRCC	
Name	*p*-Value	Rank	Value	Rank	Value	Rank	Value
x_1	0.0000	1.0	0.5960	1.0	0.5917	1.0	0.9508
x_2	0.0000	2.0	0.3624	2.0	0.3558	2.0	0.8792
x_3	0.0000	3.0	0.3553	4.0	0.3486	3.0	0.8751
x_4	0.0000	4.0	0.3484	5.0	0.3462	5.0	0.8736
x_6	0.0000	5.0	0.3467	3.0	0.3486	4.0	0.8751
x_5	0.0000	6.0	0.3431	6.0	0.3380	6.0	0.8687

[a] Table structure same as in Table 9.7.

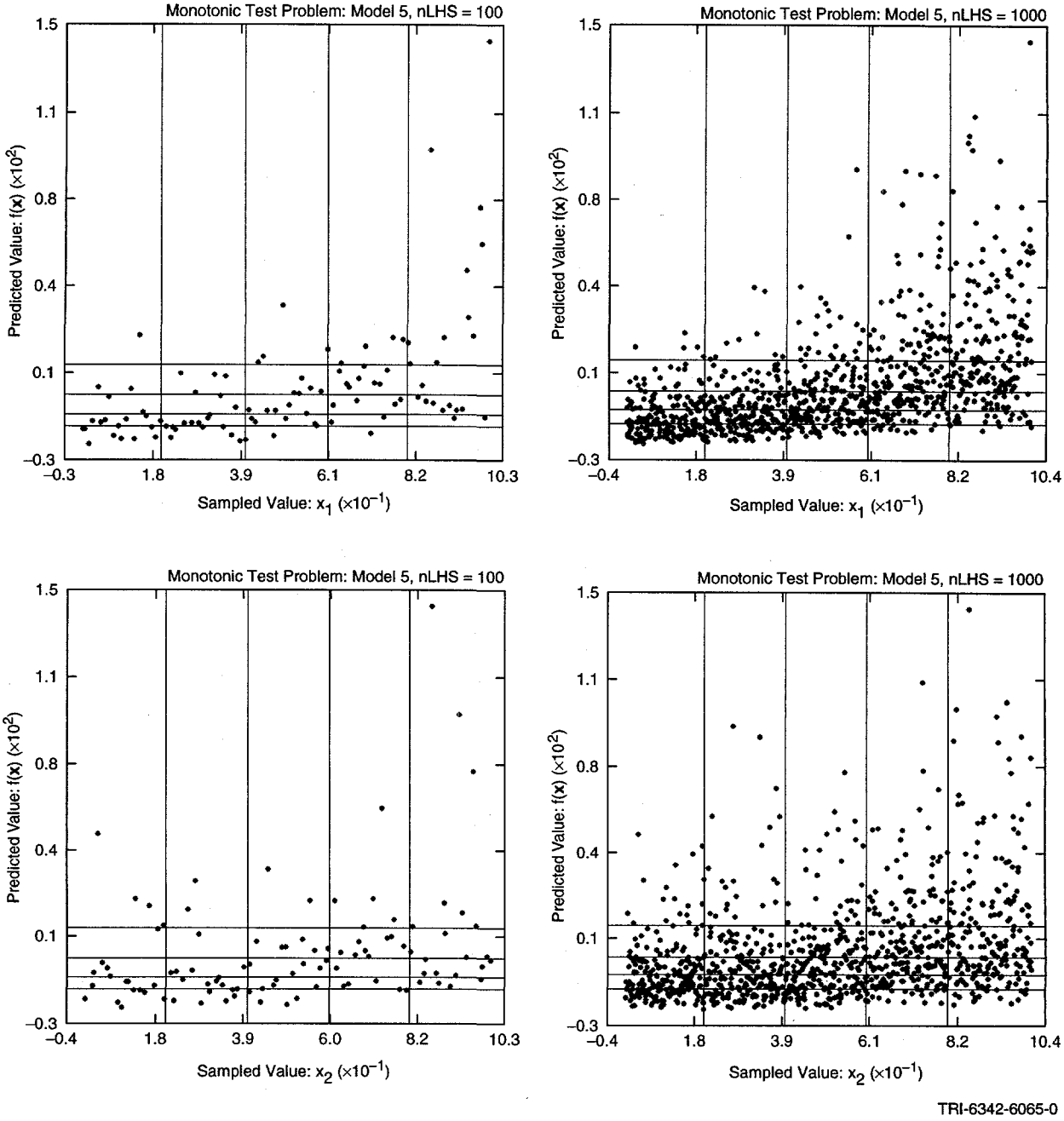


Fig. 9.9. Scatterplots for monotonic test problem with Model 5 (see Eq. (9.5)).

Table 9.11. Sensitivity Results Based on Stepwise Regression Analysis for Monotonic Test Problem with Model 5 (see Eq. (9.5)) and Sample Size *nLHS* = 100[a]

Raw Data				
Variable	R^2	RC	SRC	*p*-Value
x_1	0.25787	4.4071E+01	5.2230E–01	2.7828E–08
x_5	0.37674	2.9727E+01	3.5091E–01	2.9036E–08
x_4	0.49249	2.9194E+01	3.4459E–01	2.9872E–08
x_2	0.58539	2.4960E+01	2.9519E–01	1.7598E–07
x_3	0.66967	2.5164E+01	2.9734E–01	1.5130E–07
x_6	0.74993	2.4008E+01	2.8369E–01	4.1674E–07
Rank-Transformed Data				
Variable	R^2	RRC	SRRC	*p*-Value
x_1	0.34245	6.0130E–01	6.0130E–01	2.7828E–08
x_6	0.48424	3.6689E–01	3.6689E–01	2.7828E–08
x_3	0.62262	3.7628E–01	3.7628E–01	2.7828E–08
x_2	0.73162	3.4561E–01	3.4561E–01	2.7828E–08
x_4	0.84275	3.3165E–01	3.3165E–01	2.7828E–08
x_5	0.94119	3.1419E–01	3.1419E–01	2.7828E–08

[a] Table structure same as in Table 9.8.

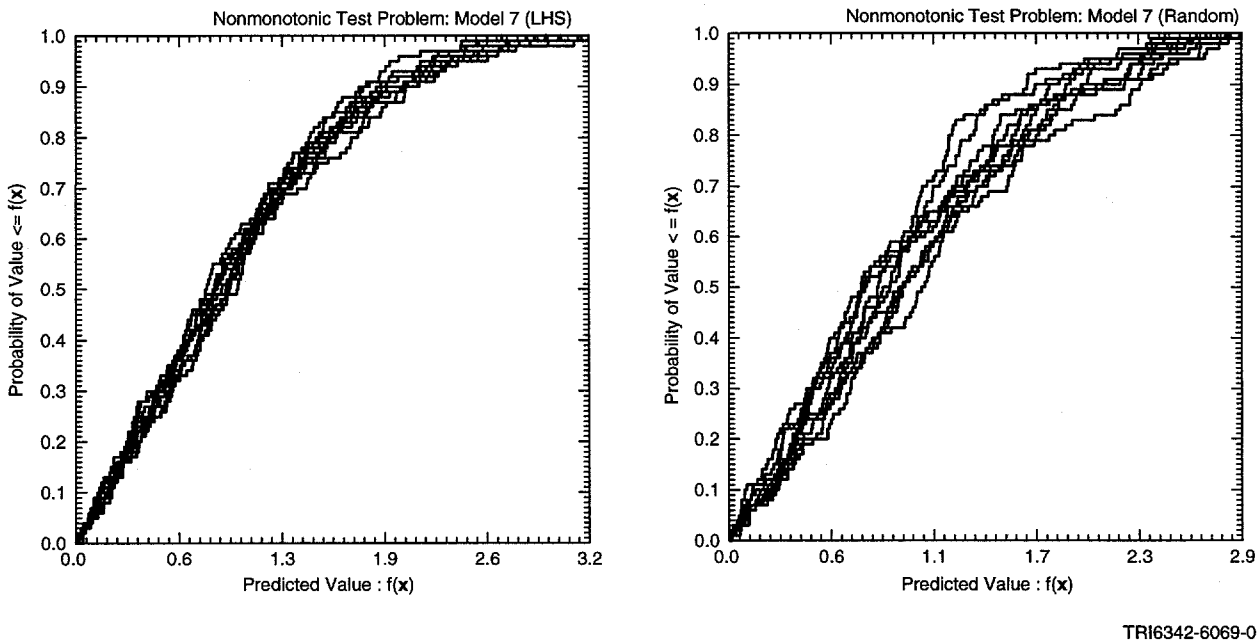


Fig. 9.10. Stability of estimated CDF for nonmonotonic test problem with Model 7 (see Eq. (9.6)).

Table 9.12. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Nonmonotonic Test Problem with Model 7 (see Eq. (9.6))[a]

Sample Size *nLHS* = 100												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_1	1.0	0.1657	1.0	0.2382	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_3	2.0	0.4400	3.0	0.4666	3.0	0.2294	3.0	0.3469	7.0	0.7358	3.0	0.1010
x_8	3.0	0.4518	2.0	0.4090	7.0	0.7298	8.0	0.7661	7.0	0.7358	8.0	0.9489
x_6	4.0	0.4566	6.0	0.5905	5.0	0.6637	6.0	0.7193	4.5	0.5918	7.0	0.8666
x_7	5.0	0.4758	4.0	0.5528	8.0	0.7360	7.0	0.7623	7.0	0.7358	6.0	0.6728
x_5	6.0	0.6796	5.0	0.5860	6.0	0.7179	5.0	0.4218	4.5	0.5918	2.0	0.0698
x_2	7.0	0.7545	8.0	0.9833	2.0	0.0010	2.0	0.0055	2.0	0.0206	4.0	0.1601
x_4	8.0	0.9581	7.0	0.9002	4.0	0.4531	4.0	0.3902	3.0	0.0916	5.0	0.5615
Sample Size *nLHS* = 1000												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_7	1.0	0.2089	1.0	0.1838	7.0	0.7123	7.0	0.7153	6.0	0.2873	6.0	0.4153
x_6	2.0	0.2644	3.0	0.2813	8.0	0.8882	8.0	0.7586	7.0	0.6411	8.0	0.9394
x_8	3.0	0.2943	2.0	0.2345	6.0	0.6228	6.0	0.4925	8.0	0.7652	7.0	0.6544
x_4	4.0	0.3376	4.0	0.4287	4.0	0.0045	4.0	0.0140	4.0	0.0224	3.0	0.0156
x_2	5.0	0.6614	6.0	0.9430	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000
x_1	6.0	0.7620	8.0	0.9708	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_5	7.0	0.8045	7.0	0.9433	5.0	0.4128	5.0	0.3011	5.0	0.1712	5.0	0.2412
x_3	8.0	0.9197	5.0	0.7315	3.0	0.0001	3.0	0.0034	3.0	0.0220	4.0	0.1178

[a] Table structure same as in Table 9.1.

$$h(x_2) = 2^{-x_2}, c_i(x_2) = (-1)^i \binom{x_2}{i} \binom{2x_2 - 2i}{x_2},$$

$$g_i(x_1, x_2) = x_1^{x_2 - 2i},$$

and x_1: $U(-1, 1)$, x_2: $DU(5)$, [~] designates the greatest integer function, and x: $DU(n)$ indicates that x has a uniform distribution over the integers $j = 1, 2, \ldots, n$.

Latin hypercube sampling and random sampling produce estimates of the CDF for $f(\mathbf{x})$ that exhibit similar stability (Fig. 9.12). This behavior is in contrast to the other examples, where Latin hypercube sampling tends to produce more stable CDF estimates than random sampling.

For *nLHS* = 100, tests based on CMNs and SI identify an effect for x_1 (i.e., *p*-values < 0.05) (Table 9.13). The test based on CLs with a *p*-value of 0.0723 also suggests an effect for x_1. None of the remaining tests (i.e., CCs, RCCs, CMDs) indicates an effect for x_1. The test based on SI with a *p*-value of 0.0698 suggests a possible effect for x_2; none of the other tests have *p*-values that suggest an effect for x_2. For *nLHS* = 1000, all tests indicate an effect for x_1, and the test based on SI also indicates an effect for x_2.

This example has complex patterns involving x_1 and x_2 (Fig. 9.13). These patterns partially emerge for *nLHS* = 100 and are readily apparent for *nLHS* = 1000. Of the tests under consideration, the test based on SI is most effective in identifying these patterns. Due to the complexity of the relations involving x_1, x_2 and $f(\mathbf{x})$, none of the previously considered coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) have values that provide any useful insights on these relationships. Similarly, stepwise regression analyses with raw and rank-transformed data fail to provide any useful insights.

The third nonmonotonic test problem (Model 9 in Campolongo et al. 2000) is defined by

$$f(\mathbf{x}) = \sin x_1 + A \sin^2 x_2 + B x_3^4 \sin x_1,$$
$$\mathbf{x} = [x_1, x_2, x_3], \tag{9.8}$$

with $A = 7$, $B = 0.1$, and x_i: $U(-\pi, \pi)$ for $i = 1, 2, 3$.

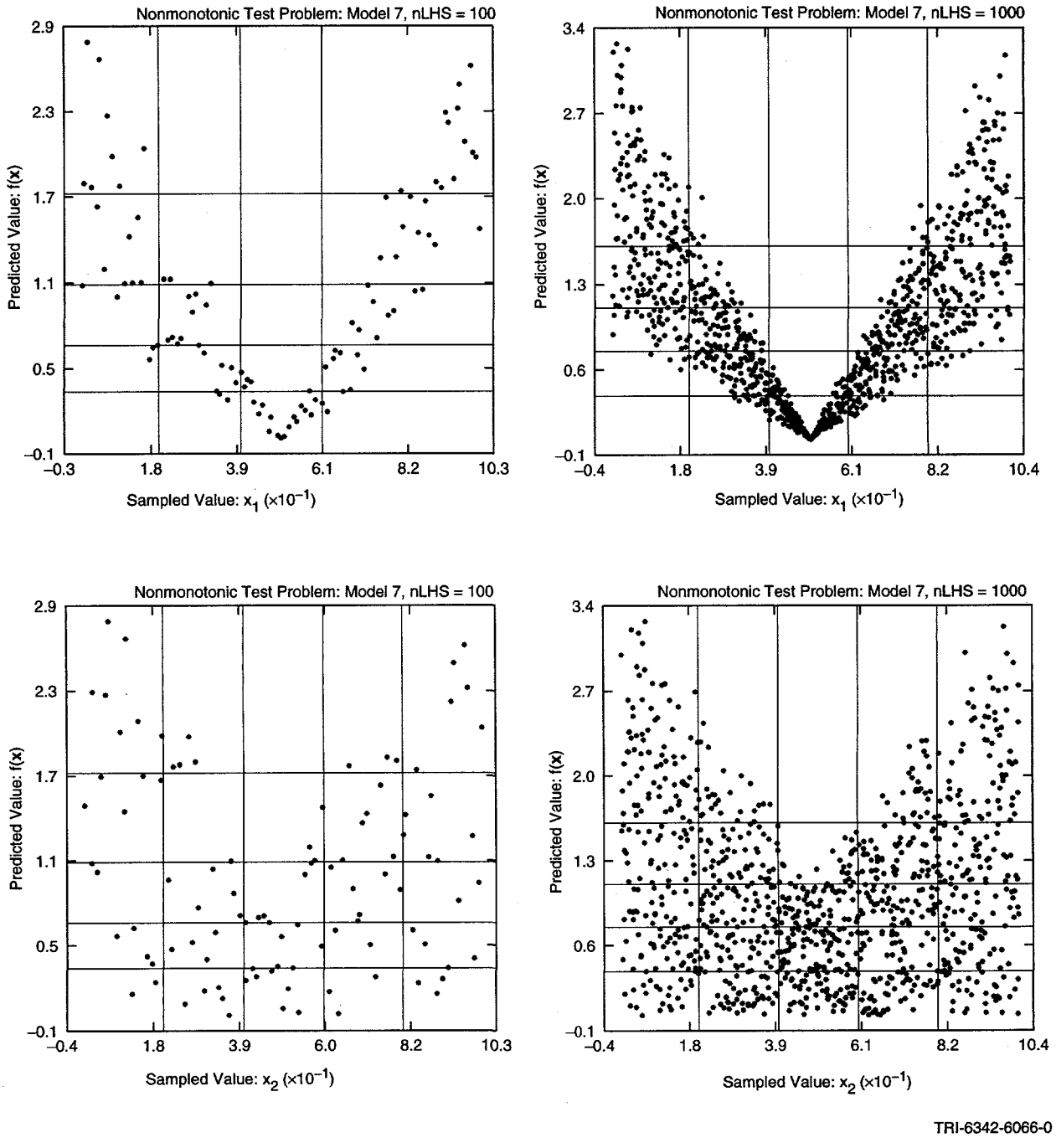


Fig. 9.11. Scatterplots for nonmonotonic test problem with Model 7 (see Eq. (9.6)).

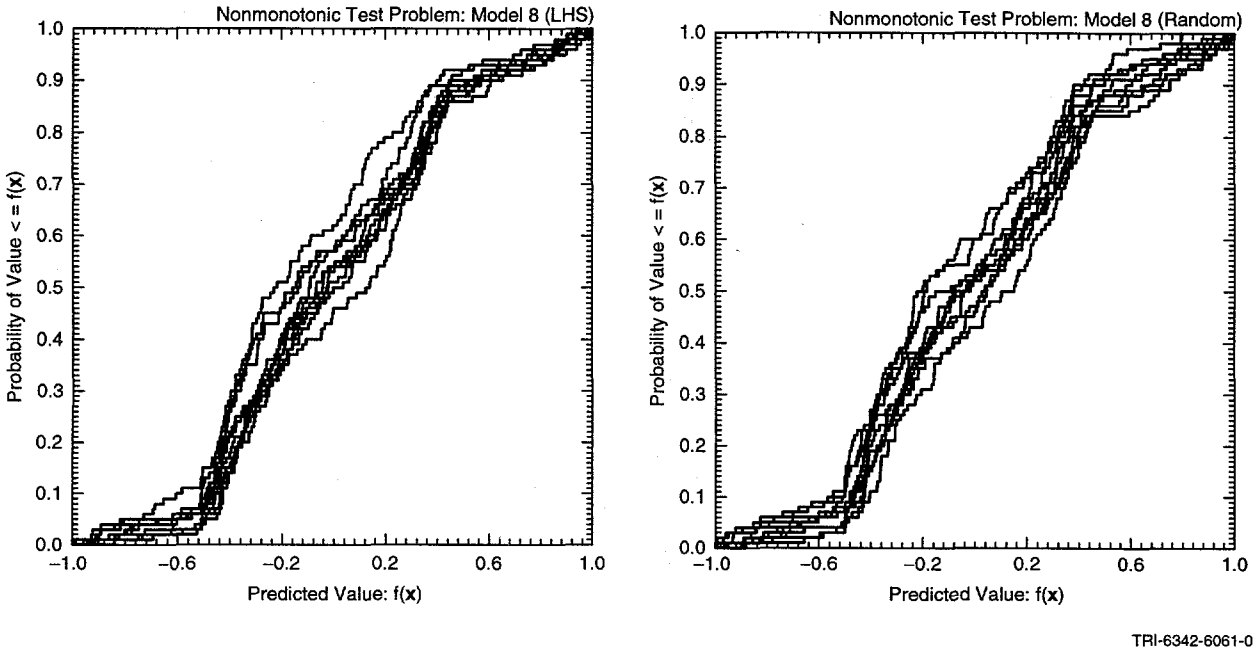


Fig. 9.12. Stability of estimated CDF for nonmonotonic test problem with Model 8 (see Eq. (9.7)).

Table 9.13. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Nonmonotonic Test Problem with Model 8 (see Eq. (9.7))[a]

Sample Size *nLHS* = 100												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_1	1.0	0.1968	2.0	0.3458	1.0	0.0346	1.0	0.0723	1.0	0.1468	1.0	0.0003
x_2	2.0	0.2412	1.0	0.2722	2.0	0.7078	2.0	0.7449	2.0	0.9384	2.0	0.0698

Sample Size *nLHS* = 1000												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
x_1	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
x_2	2.0	0.6222	2.0	0.0659	2.0	0.9090	2.0	0.2553	2.0	0.1847	2.0	0.0000

[a] Table structure same as in Table 9.1.

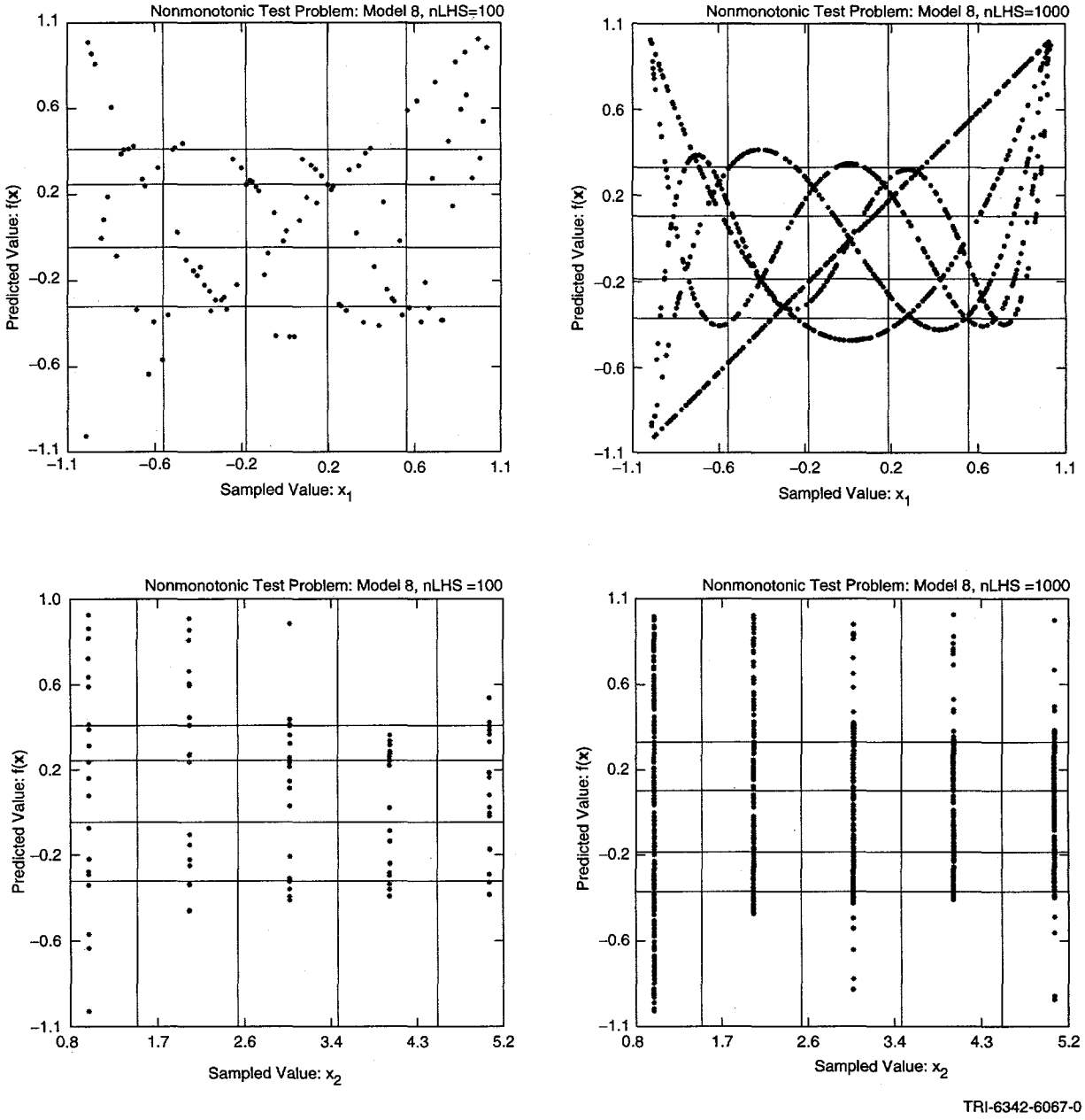


Fig. 9.13. Scatterplots for nonmonotonic test problem with Model 8 (see Eq. (9.7)).

For this example, the CDF estimates obtained with Latin hypercube sampling are more stable than those obtained with random sampling (Fig. 9.14).

In sensitivity analyses with *nLHS* = 100, all tests identify x_1 as affecting $f(\mathbf{x})$ (Table 9.14). In addition, the CMNs, CLs, CMDs and SI tests also identify an effect for x_2. None of the tests identifies an effect for x_3. For *nLHS* = 1000, all tests indicate an effect for x_1, and tests based on CMNs, CLs, CMDs and SI indicate an effect for x_2. In contrast, CCs and RCCs fail to indicate an effect for x_2. Further, the test based on SI also identifies an effect for x_3.

Examination of scatterplots clearly shows that x_1, x_2 and x_3 have readily discernible influences on $f(\mathbf{x})$ (Fig. 9.15). The tests based on CCs and RCCs are completely missing the nonlinear and nonmonotonic patterns induced in $f(\mathbf{x})$ by x_2 and x_3. Tests based on CCs and RCCs are able to identify a slight increasing pattern in the relationship between x_1 and $f(\mathbf{x})$; but this is only part of the patterns appearing in Fig. 9.15. Tests based on CMNs, CLs and CMDs identify the pattern associated with x_2 but fail to identify the pattern associated with x_3 that tends to produce similar means and medians across the entire range of x_3. In contrast, this pattern was detected by the test for SI with *nLHS* = 1000.

Due to the lack of strong linear or monotonic relationships between x_1, x_2, x_3 and $f(\mathbf{x})$, individual coefficients (i.e., CCs, SRCs, PCCs, RCCs, SRRCs, PRCCs) are close to zero and provide little useful information to help in determining the effects of x_1, x_2 and x_3 on $f(\mathbf{x})$. For the same reasons, stepwise regression analysis with raw or rank-transferred data is not very informative.

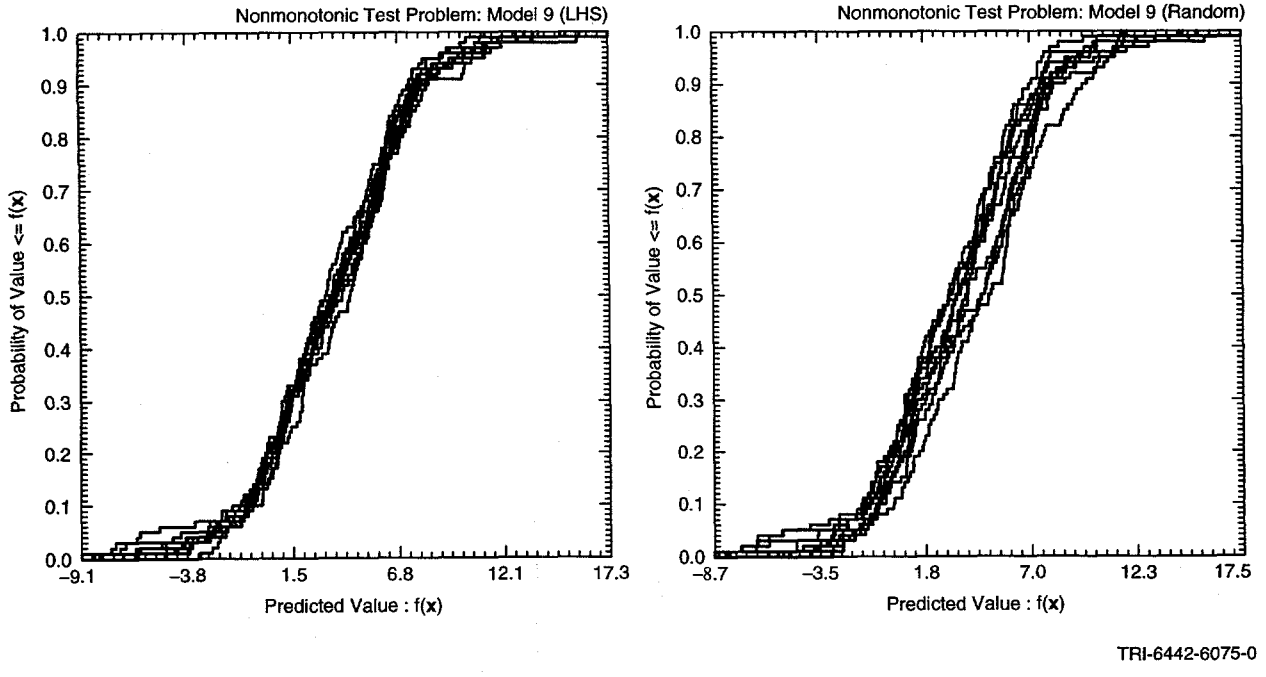


Fig. 9.14. Stability of estimated CDF for nonmonotonic test problem with Model 9 (see Eq. (9.8)).

Table 9.14. Sensitivity Results Based on CCs, RCCs, CMNs, CLs, CMDs and SI for Nonmonotonic Test Problem with Model 9 (see Eq. (9.8))[a]

Sample Size *nLHS* = 100												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
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												
Variable	CC		RCC		CMN		CL		CMD		SI	
Name	Rank	*p-Val*	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
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] Table structure same as in Table 9.1.

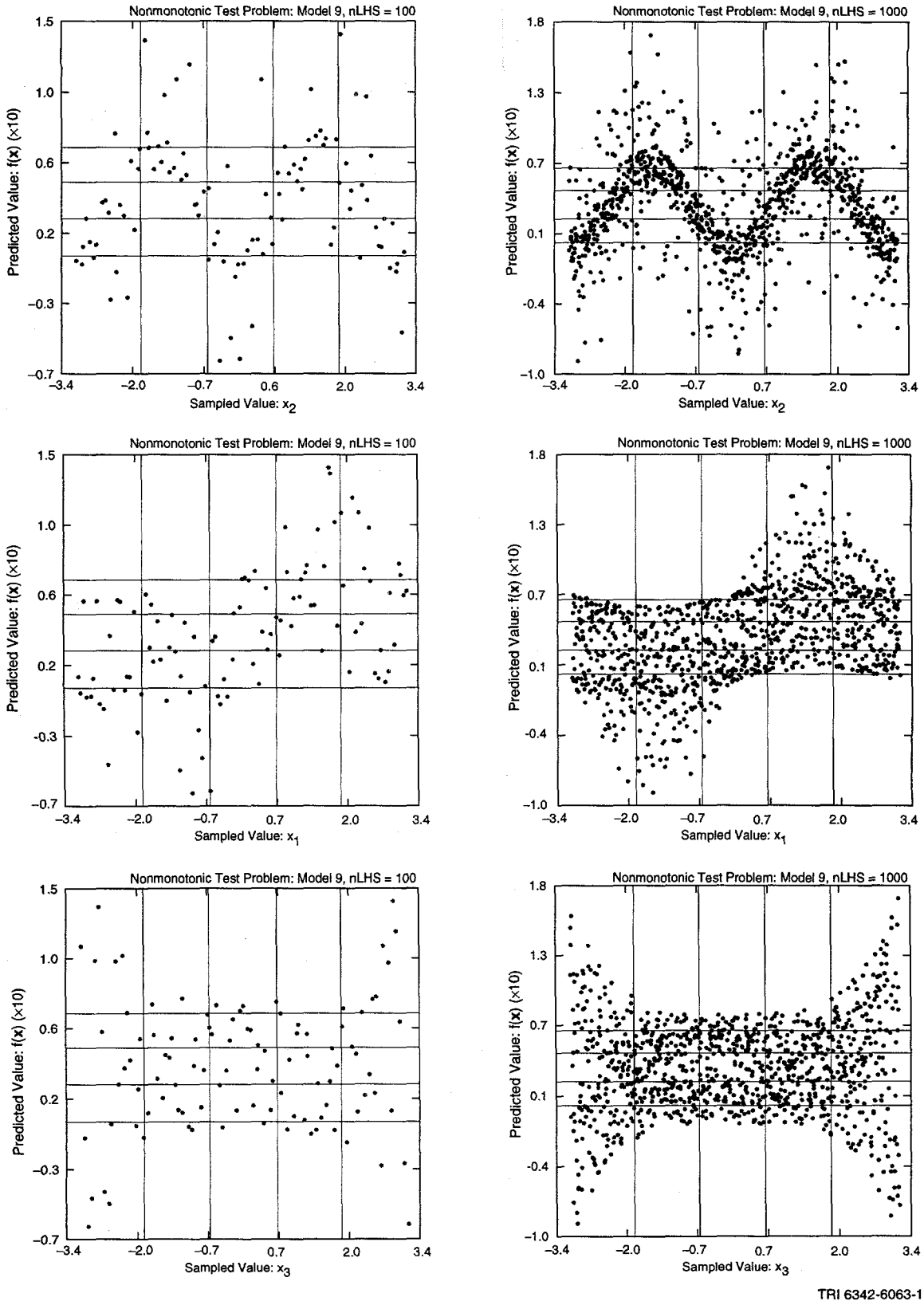


Fig. 9.15. Scatterplots for nonmonotonic test problem with Model 9 (see Eq. (9.8)).

10.0 Performance Assessment for the Waste Isolation Pilot Plant

10.1 Stochastic and Subjective Uncertainty

As indicated in Sect. 2, many large analyses involve two distinct sources of uncertainty: stochastic or aleatory uncertainty, which arises because the system under study can behave in many different ways, and subjective or epistemic uncertainty, which arises from an inability to specify an exact value for a quantity that is assumed to have a constant value within a particular analysis. An example of such an analysis is the U.S. Nuclear Regulatory Commission's reassessment of the risk from commercial nuclear reactors in the United States (commonly referred to as the NUREG-1150 analysis after its report number), where stochastic uncertainty arose from the many possible accidents that could occur at the power plants under study and subjective uncertainty arose from the many uncertain quantities required in the estimation of the probabilities and consequences of these accidents (U.S. NRC 1990-1991, Breeding et al. 1992, Helton and Breeding 1993). Numerous other examples also exist (e.g., McKone 1994, Allen et al. 1996, Price et al. 1996, Payne et al. 1992, PLG 1983, PLG 1982).

This presentation will use the PA carried out in support of the DOE's 1996 Compliance Certification Application (CCA) for the WIPP as an example of an analysis involving both stochastic and subjective uncertainty. Parts of this analysis involving the model for two-phase flow in the BRAGFLO program (Sect. 3) have already been introduced and used to illustrate uncertainty and sensitivity analysis in the presence of subjective uncertainty. Although the analyses with BRAGFLO were an important part of the 1996 WIPP PA, they constitute only one component of a large analysis. The following provides a high-level overview of sampling-based uncertainty and sensitivity analysis in the 1996 WIPP PA. The need to treat both stochastic and subjective uncertainty in the 1996 WIPP PA arose from regulations promulgated by the U.S. Environmental Protection Agency (EPA) and briefly summarized in the next paragraph.

The following is the central requirement in the EPA's regulation for the WIPP, 40 CFR 191, Subpart B, and the primary determinant of the conceptual and computational structure of the 1996 WIPP PA (p. 38086, U.S. EPA 1985):

> § 191.13 Containment requirements:
>
> (a) Disposal systems for spent nuclear fuel or high-level or transuranic radioactive wastes shall be designed to provide a reasonable expectation, based upon performance assessments, that cumulative releases of radionuclides to the accessible environment for 10,000 years after disposal from all significant processes and events that may affect the disposal system shall: (1) Have a likelihood of less than one chance in 10 of exceeding the quantities calculated according to Table 1 (Appendix A)[1]; and (2) Have a likelihood of less than one chance in 1,000 of exceeding ten times the quantities calculated according to Table 1 (Appendix A).
>
> (b) Performance assessments need not provide complete assurance that the requirements of 191.13(a) will be met. Because of the long time period involved and the nature of the events and processes of interest, there will inevitably be substantial uncertainties in projecting disposal system performance. Proof of the future performance of a disposal system is not to be had in the ordinary sense of the word in situations that deal with much shorter time frames. Instead, what is required is a reasonable expectation, on the basis of the record before the implementing agency, that compliance with 191.13(a) will be achieved.

The EPA also promulgated 40 CFR 194 (U.S. EPA 1996), where the following elaboration on the intent of 40 CFR 191.13 is given (pp. 5242-5243, U.S. EPA 1996):

> § 194.34 Results of performance assessments.
>
> (a) The results of performance assessments shall be assembled into "complementary, cumulative distribution functions" (CCDFs) that represent the probability of

[1] Radionuclide releases normalized to amount of radioactive material placed in the disposal facility; see U.S. EPA 1985 or Helton 1993a for a description of the normalization process.

exceeding various levels of cumulative release caused by all signficant processes and events. (b) Probability distributions for uncertain disposal system parameter values used in performance assessments shall be developed and documented in any compliance application. (c) Computational techniques, which draw random samples from across the entire range of the probability distributions developed pursuant to paragraph (b) of this section, shall be used in generating CCDFs and shall be documented in any compliance application. (d) The number of CCDFs generated shall be large enough such that, at cumulative releases of 1 and 10, the maximum CCDF generated exceeds the 99th percentile of the population of CCDFs with at least a 0.95 probability. (e) Any compliance application shall display the full range of CCDFs generated. (f) Any compliance application shall provide information which demonstrates that there is at least a 95 percent level of statistical confidence that the mean of the population of CCDFs meets the containment requirements of § 191.13 of this chapter.

In addition to the requirements in 40 CFR 191.13 and 40 CFR 194.34 just quoted, 40 CFR 191 and 40 CFR 194 contain many additional requirements for the certification of the WIPP for the disposal of TRU waste. However, it is the indicated requirements that determine the overall structure of the 1996 WIPP PA.

Together, 191.13(a) and 194.34(a) lead to a CCDF and boundary line as illustrated in Fig. 10.1, with the CCDF for releases to the accessible environment required to fall below the boundary line. The CCDF derives from disruptions that could occur in the future and is thus characterizing the effects of stochastic uncertainty. In contrast, 194.34(b) and (c) require the characterization and propagation of the effects of subjective uncertainty. Ultimately, this uncertainty will lead to a distribution of CCDFs of the form illustrated in Fig. 10.1, with this distribution deriving from subjective uncertainty.

Analyses that maintain a distinction between stochastic and subjective uncertainty require the introduction of a probability space $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ for stochastic uncertainty and a probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ for

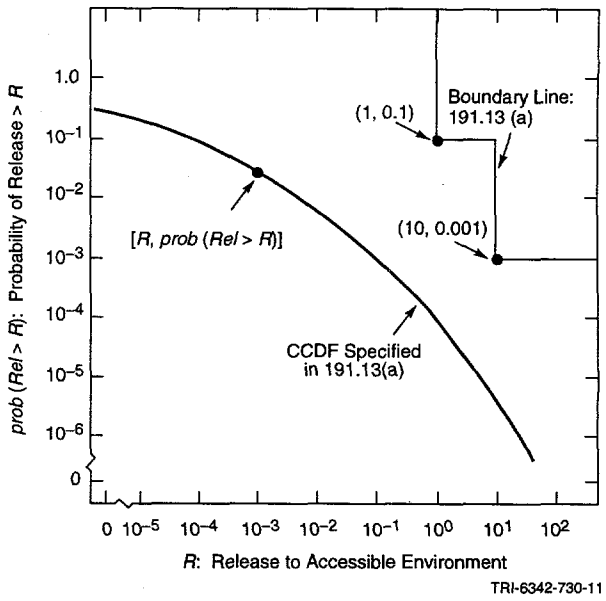


Fig. 10.1. Boundary line and associated CCDF specified in 40 CFR 191, Subpart B (Fig. 2, Helton et al. 1998b).

subjective uncertainty (Sect. 2). The probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ for subjective uncertainty used in the 1996 WIPP PA has already been introduced in Sect. 3, with Table 3.1 listing 31 of the 57 uncertain variables associated with the elements $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$. Specifically, $\mathbf{x}_{su}$ is a vector of the form

$$\mathbf{x}_{su} = [x_1, x_2, \ldots, x_{57}] \tag{10.1}$$

in the 1996 WIPP PA, where x_1, x_2,..., x_{31} are listed in Table 3.1 and the remaining elements of $\mathbf{x}_{su}$ (i.e., x_{32}, x_{33}, ..., x_{57}) are listed in Table 5.2.1 of Helton et al. 1998a. The probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ was defined by specifying distributions for the elements of $\mathbf{x}_{su}$ as indicated in Eq. (1.3) and illustrated in Fig. 4.3.

In the 1996 WIPP PA, the probability space $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ for stochastic uncertainty derives from the many different disruptions that could occur at the WIPP over the 10,000 yr regulatory time frame imposed on it. In particular, regulatory guidance and extensive review of potential features, events and processes (FEPs) that could affect the WIPP led to the elements $\mathbf{x}_{st}$ of the sample space $\mathcal{S}_{st}$ being defined as vectors of the form

$$\mathbf{x}_{st} = [\underbrace{t_1, l_1, e_1, b_1, p_1, \mathbf{a}_1}_{1^{st}\text{ intrusion}}, \underbrace{t_2, l_2, e_2, b_2, p_2, \mathbf{a}_2}_{2^{nd}\text{ intrusion}}, \ldots,$$

$$\underbrace{t_n, l_n, e_n, b_n, p_n, \mathbf{a}_n}_{n^{th}\text{ intrusion}}, t_{min}], \tag{10.2}$$

where n is the number of exploratory drilling intrusions for natural resources (i.e., oil or gas) that occur in the immediate vicinity of the repository, t_i is the time (yr) of the i^{th} intrusion, l_i designates the location of the i^{th} intrusion, e_i designates the penetration of an excavated or nonexcavated area by the i^{th} intrusion, b_i designates whether or not the i^{th} intrusion penetrates pressurized brine in the Castile Fm, p_i designates the plugging procedure used with the i^{th} intrusion (i.e., continuous plug, two discrete plugs, three discrete plugs), $\mathbf{a}_i$ designates the type of waste penetrated by the i^{th} intrusion (i.e., no waste, contact-handled (CH) waste, remotely handled (RH) waste), and t_{min} is the time at which potash mining occurs within the land withdrawal boundary (Chapt. 3, Helton et al. 1998a). The definition of ($\mathcal{S}_{st}$, $\mathcal{J}_{st}$, p_{st}) was then completed by assigning a distribution to each element of $\mathbf{x}_{st}$ (Chapt. 3, Helton et al. 1998a).

The FEPs review process also led to the identification of processes and associated models for use in the estimation of consequences (e.g., normalized radionuclide releases to the accessible environment in the context of the EPA regulations) for elements $\mathbf{x}_{st}$ of $\mathcal{S}_{st}$ (Fig. 10.2, Table 10.1). Symbolically, this estimation process can be represented by

$$\begin{aligned} f(\mathbf{x}_{st}) &= f_C(\mathbf{x}_{st}) + f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] \\ &+ f_{DBR}\{\mathbf{x}_{st}, f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})], f_B(\mathbf{x}_{st})\} \\ &+ f_{MB}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] + f_{DL}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] \\ &+ f_S[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] \\ &+ f_{S-T}\{\mathbf{x}_{st,0}, f_{S-F}(\mathbf{x}_{st,0}), f_{N-P}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]\}, \end{aligned} \tag{10.3}$$

where $f(\mathbf{x}_{st})$ ~ normalized radionuclide release to the accessible environment associated with $\mathbf{x}_{st}$ and, in general, many additional consequences, $\mathbf{x}_{st}$ ~ particular future under consideration, $\mathbf{x}_{st,0}$ ~ future involving no drilling intrusions but a mining event at the same time t_{min} as in $\mathbf{x}_{st}$, $f_C(\mathbf{x}_{st})$ ~ cuttings and cavings release to accessible environment for $\mathbf{x}_{st}$ calculated with CUTTINGS_S, $f_B(\mathbf{x}_{st})$ ~ results calculated for $\mathbf{x}_{st}$ with BRAGFLO (in practice, $f_B(\mathbf{x}_{st})$ is a vector containing a large amount of information including time-dependent pressures and saturations for gas and brine),

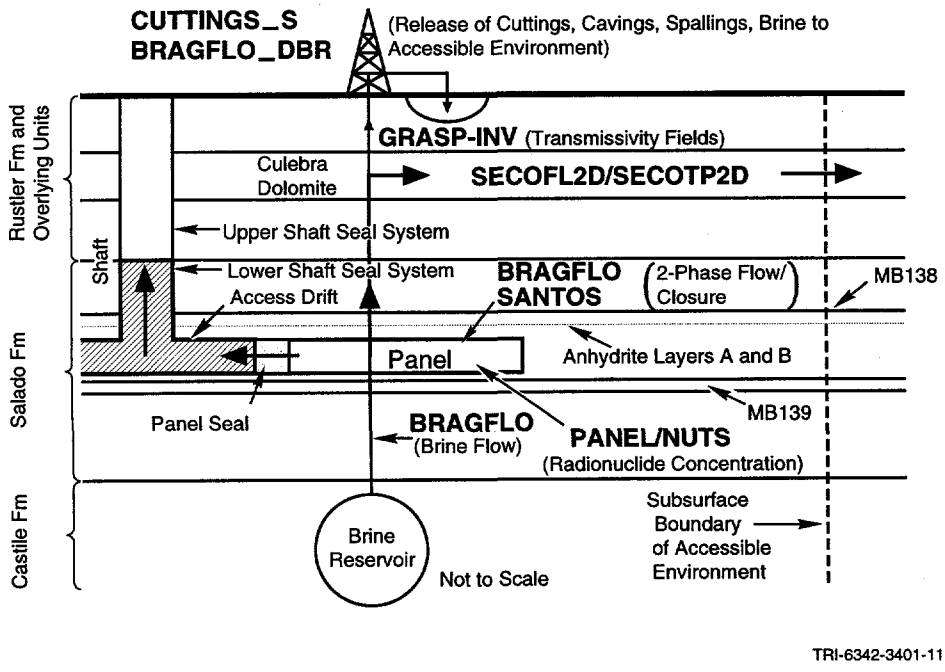


Fig. 10.2. Computer programs (models) used in 1996 WIPP PA (Fig. 5, Helton et al. 1998b).

Table 10.1. Summary of Computer Models Used in the 1996 WIPP PA (Table 1, Helton et al. 1998b).

BRAGFLO: Calculates multiphase flow of gas and brine through a porous, heterogeneous reservoir. Uses finite difference procedures to solve system of nonlinear partial differential equations that describes the mass conservation of gas and brine along with appropriate constraint equations, initial conditions and boundary conditions. Additional information: Bean et al. 1996; Sect. 4.2, Helton et al. 1998a.

BRAGFLO_DBR: Special configuration of BRAGFLO model used in calculation of dissolved radionuclide releases to the surface (i.e., direct brine releases) at the time of a drilling intrusion. Uses initial value conditions obtained from calculations performed with BRAGFLO and CUTTINGS_S. Additional information: Stoelzel et al. 1996; Sect. 4.7, Helton et al. 1998a.

CUTTINGS_S: Calculates the quantity of radioactive material brought to the surface in cuttings and cavings and also in spallings generated by an exploratory borehole that penetrates a waste panel, where cuttings designates material removed by the drillbit, cavings designates material eroded into the borehole due to shear stresses resulting from the circular flow of the drilling fluid (i.e., mud), and spallings designates material carried to the borehole at the time of an intrusion due to the flow of gas from the repository to the borehole. Spallings calculation uses initial value conditions obtained from calculations performed with BRAGFLO. Additional information: Berglund 1992, 1996; Sects. 4.5, 4.6, Helton et al. 1998a.

GRASP-INV: Generates transmissivity fields (estimates of transmissivity values) conditioned on measured transmissivity values and calibrated to steady-state and transient pressure data at well locations using an adjoint sensitivity and pilot-point technique. Additional information: LaVenue and Rama Rao 1992, LaVenue 1996.

NUTS: Solves system of partial differential equations for radionuclide transport in vicinity of repository. Uses brine volumes and flows calculated by BRAGFLO as input. Additional information: Stockman et al. 1996; Sect. 4.3, Helton et al. 1998a.

PANEL: Calculates rate of discharge and cumulative discharge of radionuclides from a waste panel through an intruding borehole. Discharge is a function of fluid flow rate, elemental solubility and radionuclide inventory. Uses brine volumes and flows calculated by BRAGFLO as input. Based on solution of system of linear ordinary differential equations. Additional information: Stockman et al. 1996; Sect. 4.4, Helton et al. 1998a.

SANTOS: Solves quasistatic, large deformation, inelastic response of two-dimensional solids with finite element techniques. Used to determine porosity of waste as a function of time and cumulative gas generation, which is an input to calculations performed with BRAGFLO. Additional information: Bean et al. 1996; Stone 1997a, 1997b; Sect. 4.2.3, Helton et al. 1998a.

SECOFL2D: Calculates single-phase Darcy flow for groundwater flow in two dimensions. The formulation is based on a single partial differential equation for hydraulic head using fully implicit time differencing. Uses transmissivity fields generated by GRASP-INV. Additional information: Roache 1993, Ramsey and Wallace 1996; Sect. 4.8, Helton et al. 1998a.

SECOTP2D: Simulates transport of radionuclides in fractured porous media. Solves two partial differential equations: one provides two-dimensional representation for convective and diffusive radionuclide transport in fractures and the other provides one-dimensional representation for diffusion of radionuclides into rock matrix surrounding the fractures. Uses flow fields calculated by SECOFL2D. Additional information: Roache 1993, Ramsey and Wallace 1996; Sect. 4.9, Helton et al. 1998a.

$f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ spallings release to accessible environment for $\mathbf{x}_{st}$ calculated with the spallings model contained in CUTTINGS_S, $f_{DBR}\{\mathbf{x}_{st}, f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})], f_B(\mathbf{x}_{st})\}$ ~ direct brine release to accessible environment for $\mathbf{x}_{st}$ calculated with a modified version of BRAGFLO designated BRAGFLO_DBR, $f_{MB}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release through anhydrite marker beds to accessible environment for $\mathbf{x}_{st}$ calculated with NUTS, $f_{DL}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release through Dewey Lake Red Beds to accessible environment for $\mathbf{x}_{st}$ calculated with NUTS, $f_S[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release to land surface due to brine flow up a plugged borehole for $\mathbf{x}_{st}$ calculated with NUTS or PANEL as appropriate, $f_{S-F}(\mathbf{x}_{st,0})$ ~ flow field calculated for $\mathbf{x}_{st,0}$ with SECOFL2D, $f_{N-P}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release to Culebra for $\mathbf{x}_{st}$ calculated with NUTS or PANEL as appropriate, and $f_{S-T}\{\mathbf{x}_{st,0}, f_{S-F}(\mathbf{x}_{st,0}), f_{N-P}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]\}$ ~ ground water transport release through Culebra to accessible environment calculated with SECOTP2D ($\mathbf{x}_{st,0}$ is used as an argument to f_{S-T} because drilling intrusions are assumed to cause no perturbations to the flow field in the Culebra).

The probability space $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ for stochastic uncertainty and the function f indicated in Eq. (10.3) lead to the required CCDF for normalized releases to the accessible environment (Fig. 10.1). In particular, this CCDF can be represented as an integral involving $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ and f (Fig. 10.3). If $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ and f could be unambiguously defined, then the CCDF in Fig. 10.3 could be determined with certainty and compared against the specified boundary line. Unfortunately, such certainty does not exist in the 1996 WIPP PA, which leads to the probability space $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ for subjective uncertainty.

When the elements $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$ are included, the function f in Eq. (10.3) has the form $f(\mathbf{x}_{st}, \mathbf{x}_{su})$. In turn, the expression defining the CCDF in Fig. 10.3 becomes

$$prob(Rel > R \mid \mathbf{x}_{su}) = \int_{\mathcal{S}_{st}} \delta_R[f(\mathbf{x}_{st}, \mathbf{x}_{su})] d_{st}(\mathbf{x}_{st} \mid \mathbf{x}_{su}) dV_{st}, \quad (10.4)$$

where $\delta_R[f(\mathbf{x}_{st}, \mathbf{x}_{su})] = 1$ if $f(\mathbf{x}_{st}, \mathbf{x}_{su}) > R$ and 0 if $f(\mathbf{x}_{st}, \mathbf{x}_{su}) \leq R$. Uncertainty in $\mathbf{x}_{su}$ as characterized by $(\mathcal{S}_{su}, \mathcal{A}_{su}, p_{su})$ then leads to a distribution of CCDFs, with one CCDF resulting for each $\mathbf{x}_{su}$ in $\mathcal{S}_{su}$ (Fig. 10.4).

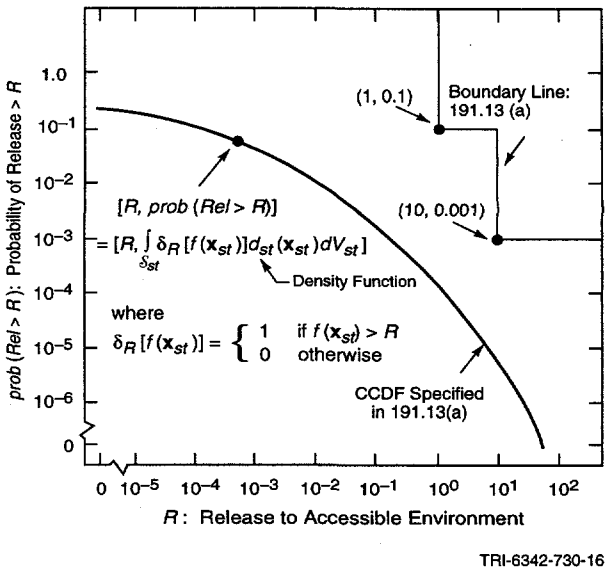


Fig. 10.3. Definition of CCDF specified in 40 CFR 191, Subpart B as an integral involving the probability space $(\mathcal{S}_{st}, \mathcal{A}_{st}, p_{st})$ for stochastic uncertainty and a function f defined on $\mathcal{S}_{st}$ (Fig. 4, Helton et al. 1998b).

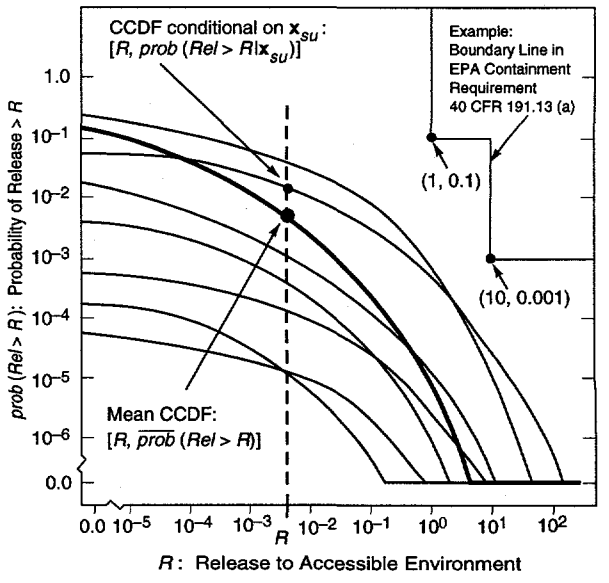


Fig. 10.4. Individual CCDFs conditional on elements $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$ (i.e., CCDFs represented by $[R, prob(Rel > R|\mathbf{x}_{su})]$; see Eq. (10.4)) and associated mean CCDF (i.e., CCDF represented by $[R, \overline{prob}(Rel > R)]$; see Eq. (10.7)).

10.2 Implementation of Analysis

The guidance in 194.34(a) was implemented by developing the probability space ($\mathcal{S}_{st}$, $\mathcal{A}_{st}$, p_{st}), the function $f(\mathbf{x}_{st}, \mathbf{x}_{su})$, and a Monte Carlo procedure based on simple random sampling (Sect. 5.1) for the approximation of the integral, and hence the associated CCDF, in Eq. (10.4). Conditional on an element $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$, the Monte Carlo approximation procedure has the form

$$prob(Rel > R \mid \mathbf{x}_{su}) \doteq \sum_{i=1}^{nS} \delta_R[f(\mathbf{x}_{st,i}, \mathbf{x}_{su})]/nS, \quad (10.5)$$

where $\mathbf{x}_{st,i}$, $i = 1,2,\ldots, nS = 10{,}000$, is a random sample of size nS from ($\mathcal{S}_{st}$, $\mathcal{A}_{st}$, p_{st}). This approximation procedure required evaluating the models in Table 10.1 for a relatively small number of elements of $\mathcal{S}_{st}$ and then using these evaluations to construct $f(\mathbf{x}_{st,i}, \mathbf{x}_{su})$ for the large number of sample elements (i.e., nS = 10,000) used in the summation in Eq. (10.5) (see Helton et al. 1998a for numerical details).

The guidance in 194.34(b) was implemented by developing the probability space ($\mathcal{S}_{su}$, $\mathcal{A}_{su}$, p_{su}). Latin hypercube sampling (Sect. 5.3) was selected as the sampling technique required in 194.34(c) because of the efficient manner in which it stratifies across the range of each sampled variable. For a Latin hypercube or random sample of size n, the requirement in 194.34(c) is equivalent to the inequality

$$1 - 0.99^n > 0.95, \quad (10.6)$$

which results in a minimum value of 298 for n. In consistency with the preceding result, the 1996 WIPP PA used an LHS of size 300 from the probability space ($\mathcal{S}_{su}$, $\mathcal{A}_{su}$, p_{su}) for subjective uncertainty. Actually, as discussed below, three replicated LHSs of size 100 each were used, which resulted in a total sample size of 300 (Sect. 5.6). Further, the requirement in 194.34(d) is met by simply providing plots that contain all the individual CCDFs produced in the analysis (i.e., one CCDF for each LHS element; i.e., plots of the form indicated in Fig. 10.4).

The requirement in 194.34(f) involves the mean of the distribution of CCDFs, with this distribution resulting from subjective uncertainty (Fig. 10.4). In particular, each individual CCDF in Fig. 10.4 is conditional on an element $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$ and is defined by the points [R, $prob(Rel > R|\mathbf{x}_{su})$], with $prob(Rel > R|\mathbf{x}_{su})$ given in Eq. (10.4). Similarly, the mean CCDF is defined by the points [R, $\overline{prob}(Rel > R)$], where

$$\overline{prob}(Rel > R) = \text{mean probability of a release greater than size } R$$

$$= \int_{\mathcal{S}_{su}} prob(Rel > R|\mathbf{x}_{su}) d_{su}(\mathbf{x}_{su}) dV_{su}$$

$$= \int_{\mathcal{S}_{su}} \left[\int_{\mathcal{S}_{st}} \delta_R[f(\mathbf{x}_{st}, \mathbf{x}_{su})] d_{st}(\mathbf{x}_{st} \mid \mathbf{x}_{su}) dV_{st} \right] d_{su}(\mathbf{x}_{su}) dV_{su} \quad (10.7)$$

and $d_{su}(\mathbf{x}_{su})$ is the density function associated with ($\mathcal{S}_{su}$, $\mathcal{A}_{su}$, p_{su}).

The integral over $\mathcal{S}_{su}$ in the definition of $\overline{prob}(Rel > R)$ is too complex to be determined exactly. The EPA anticipated that a sampling-based integration procedure would be used to estimate this integral, with the requirement in 194.34(f) placing a condition on the accuracy of this procedure.

Given that Latin hypercube sampling is to be used to estimate the outer integral in Eq. (10.7), the confidence intervals required in 194.34(f) can be obtained with a replicated sampling technique proposed by Iman (1982). In this technique, the LHS to be used is repeatedly generated with different random seeds. These samples lead to a sequence $\overline{prob}_r(Rel > R)$, r = 1, 2, ..., nR, of estimated mean exceedance probabilities, where $\overline{prob}_r(Rel > R)$ defines the mean CCDF obtained for sample r (i.e., $\overline{prob}_r(Rel > R)$ is the mean probability that a normalized release of size R will be exceeded; see Eq. (10.7)) and nR is the number of independent LHSs generated with different random seeds. Then,

$$\overline{\overline{prob}}(Rel > R) = \sum_{r=1}^{nR} \overline{prob}_r(Rel > R)/nR \quad (10.8)$$

and

$$SE(R) = \left\{ \sum_{r=1}^{nR} \left[\overline{\overline{prob}}(Rel > R) - \overline{prob}_r(Rel > R) \right]^2 / nR(nR-1) \right\}^{1/2} \quad (10.9)$$

provide an additional estimate of the mean CCDF and estimates of the standard errors associated with the individual mean exceedance probabilities. The t-distribution with $nR-1$ degrees of freedom can be used to place confidence intervals around the mean exceedance probabilities for individual R values (i.e., around $\overline{\overline{prob}}(Rel > R)$). Specifically, the $1-\alpha$ confidence interval is given by $\overline{\overline{prob}}(Rel > R) \pm t_{1-\alpha/2}\, SE(R)$, where $t_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the t-distribution with $nR-1$ degrees of freedom (e.g., $t_{1-\alpha/2} = 4.303$ for $\alpha = 0.05$ and $nR = 3$). The same procedure can also be used to place pointwise confidence intervals around percentile curves. The implementation of this procedure is the reason for the three replicated LHSs indicated in Sect. 5.6.

At the beginning of the computational implementation of the 1996 WIPP PA, only the 31 variables in Table 3.1 that are used as input to BRAGFLO had been fully specified (i.e., their distributions D_j had been unambiguously defined); the remaining variables that would be incorporated into the definition of $\mathbf{x}_{su}$ were still under development. To allow the calculations with BRAGFLO to proceed, the LHSs indicated in Sect. 5.6 were actually generated from $nX = 75$ variables, with the first 31 variables being the then specified inputs to BRAGFLO and the remaining 44 variables being assigned uniform distributions on [0, 1]. Later, when the additional variables were fully specified, the uniformly distributed variables were used to generate sampled values from them consistent with their assigned distributions. This procedure allowed the analysis to go forward while maintaining the integrity of the Latin hypercube sampling procedure for the overall analysis. As previously indicated, 26 additional variables were eventually defined, with the result that the elements $\mathbf{x}_{su}$ of $\mathcal{S}_{su}$ had an effective dimension of $nX = 57$.

10.3 Uncertainty and Sensitivity Analysis Results

The CCDF used in comparisons with the EPA release limits (Figs. 10.1, 10.3) is the most important single result generated in the 1996 WIPP PA. This CCDF arises from stochastic uncertainty. However, because there is subjective uncertainty in quantities used in the generation of this CCDF, its value cannot be known with certainty. The use of Latin hypercube sampling leads to an estimate of the uncertainty in the location of this CCDF (Fig. 10.5), with the individual CCDFs falling substantially to the left of the release limits. The left frame (Fig. 10.5a) shows the individual CCDFs obtained for replicate R1, and the right frame (Fig. 10.5b) shows the mean and selected quantile curves obtained from pooling the three replicates. The mean curve in Fig. 10.5b is formally defined in Eq. (10.7), and the construction procedures used to obtain the individual curves in Fig. 10.5b are described in conjunction with Fig. 7.8.

The replicated samples described in Sect. 5.6 were used to obtain an indication of the stability of results obtained with Latin hypercube sampling. For the total release CCDFs in Fig. 10.5, the results obtained for the three replicates (i.e., R1, R2, R3) were very stable, with little variation in the locations of the mean and quantile curves occurring across replicates (Fig. 10.6a). Indeed, the mean and quantile curves for the individual replicates overlay each other to the extent that they are almost indistinguishable. As a result, the procedure indicated in conjunction with Eqs. (10.8) and (10.9) provides a very tight confidence interval around the estimated mean CCDF (Fig. 10.6b).

The sampling-based approach to uncertainty analysis has created a pairing between the LHS elements in Eq. (5.16) and the individual CCDFs in Fig. 10.5a that can be explored with the previously discussed sensitivity analysis techniques. One possibility for investigating the sources of the uncertainty that give rise to the distribution of CCDFs in Fig. 10.5a is to determine what is giving rise to the variation in exceedance probabilities for individual release values on the abscissa. This variation in exceedance probabilities can be investigated in exactly the same manner as the variation in pressure at individual times was investigated for the pressure curves in Fig. 7.5 and presented in Fig. 8.3. Specifically, PRCCs, SRRCs, or some other measure of sensitivity can be calculated for the exceedance probabilities associated with individual release values. This measure for different sampled variables can be plotted above the corresponding release values on the abscissa and then connected to obtain a representation for how sensitivity changes for changing values on the abscissa. For the CCDFs in Fig. 10.5a, this analysis approach shows that the exceedance probabilities for individual release values are primarily influenced by *WMICDFLG* and *WTAUFAIL* (shear strength of waste), with the exceedance probabilities tending to increase as *WMICDFLG* increases and tending to decrease as *WTAUFAIL* increases (Fig. 10.7).

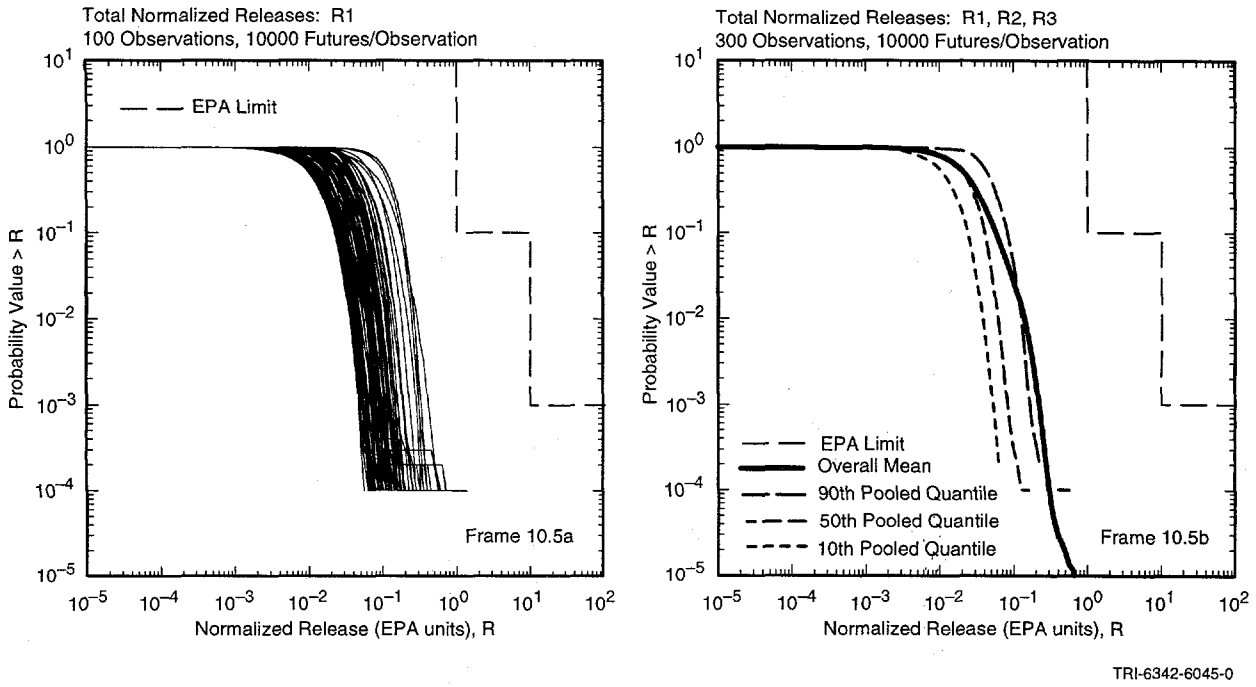


Fig. 10.5. Distribution of CCDFs for total normalized release to the accessible environment over 10,000 yr: (a) 100 individual CCDFs for replicate R1, and (b) mean and percentile curves estimated from 300 CCDFs obtained by pooling replicates R1, R2 and R3 (Figs. 6, 7, Helton et al. 1998b).

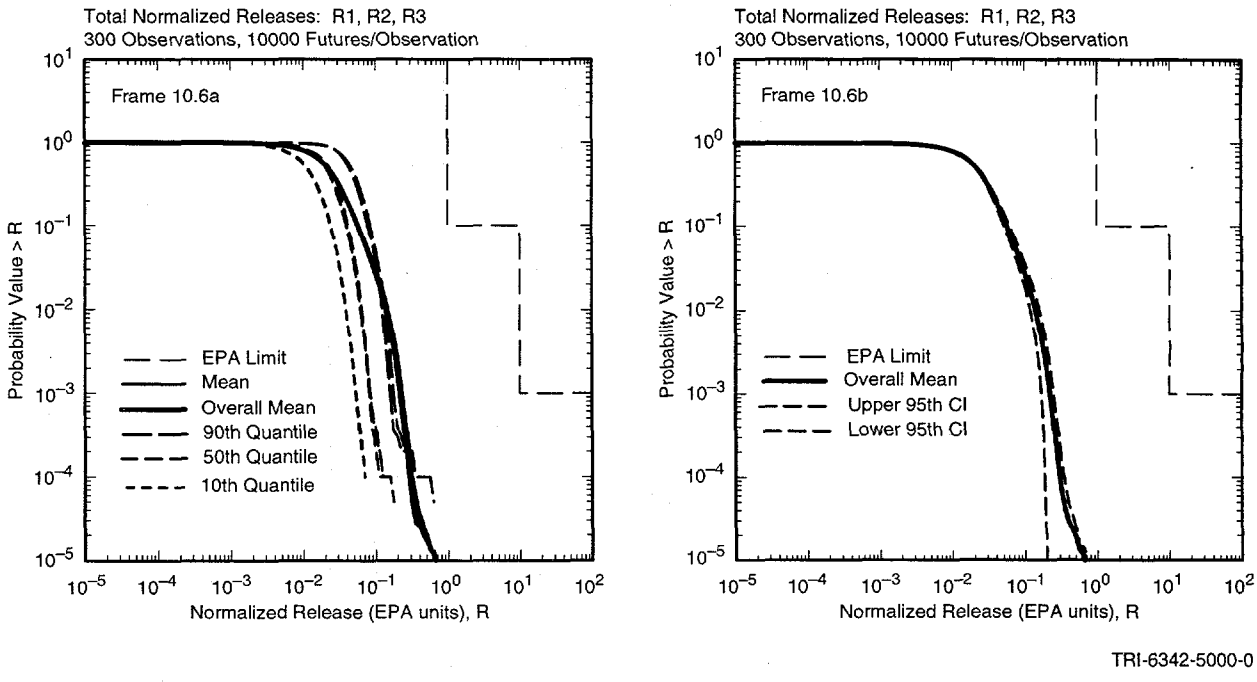


Fig. 10.6. Stability of estimated distribution of CCDFs for normalized release to the accessible environment: (a) mean and quantile curves for individual replicates, and (b) confidence interval around mean CCDF obtained by pooling the three individual replicates (Fig. 8, Helton et al. 1998b).

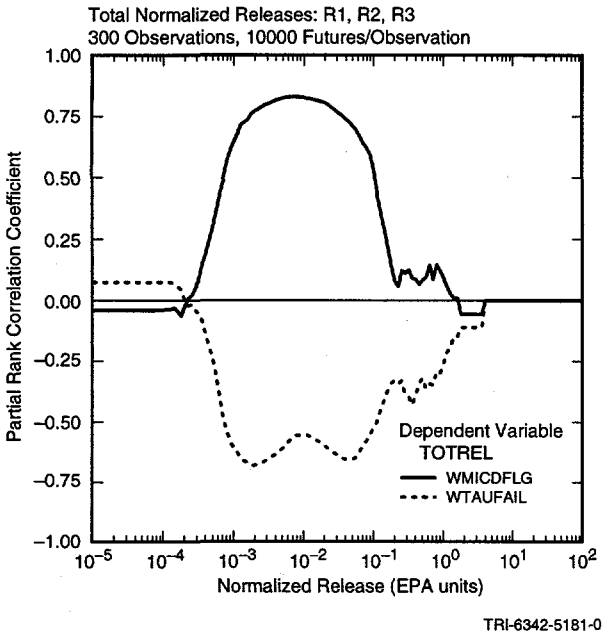


Fig. 10.7. Sensitivity analysis based on PRCCs for CCDFs for normalized release to the accessible environment (Fig. 14, Helton 1999).

Another possibility is to reduce the individual CCDFs to expected values over stochastic uncertainty and then to perform a sensitivity analysis on the resultant expected values. In the context of the CCDF representation in Eq. (10.4), this expected value can be formally defined by

$$E(R \mid \mathbf{x}_{su}) = \int_{S_{st}} f(\mathbf{x}_{st}, \mathbf{x}_{su}) d_{st}(\mathbf{x}_{st} \mid \mathbf{x}_{su}) dV_{st}. \quad (10.10)$$

The LHS in Eq. (5.16) then results in a sequence of values $E(R \mid \mathbf{x}_{su,k})$, $k = 1,2, \ldots, nLHS$, that can be explored with the previously discussed sensitivity analysis procedures. For example, stepwise regression analysis shows that *WMICDFLG* and *WTAUFAIL* are the dominant variables with respect to the uncertainty in $E(R \mid \mathbf{x}_{su})$, but with lesser effects due to a number of additional variables (Table 10.2).

This section briefly describes the 1996 WIPP PA and illustrates uncertainty and sensitivity analysis in the context of this PA. Additional details are available in other presentations (Helton et al. 1998a, Helton 1999, Helton et al. 1999).

Table 10.2. Stepwise Regression Analysis with Rank-Transformed Data for Expected Normalized Release Associated with Individual CCDFs for Total Release Due to Cuttings and Cavings, Spallings and Direct Brine Release (Table 5, Helton 1999)

	Expected Normalized Release		
Step[a]	Variable[b]	SRRC[c]	R^{2d}
1	*WMICDFLG*	0.60	0.40
2	*WTAUFAIL*	−0.39	0.55
3	*WGRCOR*	0.21	0.59
4	*WPRTDIAM*	−0.19	0.63
5	*HALPOR*	0.17	0.65
6	*BHPRM*	−0.17	0.68
7	*HALPRM*	0.16	0.71
8	*WASTWICK*	0.11	0.72
9	*ANHPRM*	0.09	0.73

[a] Steps in stepwise regression analysis.
[b] Variables listed in order of selection in regression analysis with *ANHCOMP* and *HALCOMP* excluded from entry into regression model.
[c] Standardized rank regression coefficients (SRRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.

11.0 Summary

Sampling-based methods for uncertainty and sensitivity analysis have a number of desirable properties, including (i) conceptual simplicity, (ii) ease and flexibility in adaptation to specific analysis situations, (iii) stratification over the range of each uncertain variable, (iv) direct estimation of distribution functions to characterize the uncertainty in model predictions, and (v) availability of a variety of sensitivity analysis techniques.

The results of sampling-based uncertainty and sensitivity analyses are conditional on the distributions assigned to the uncertain (i.e., sampled) variables. Thus, care must be used in assigning these distributions. A possibility is to carry out multiple iterations of an analysis. The first iteration could be performed with rather crude distribution assumptions to determine the most important variables. Then, additional resources could be focused on characterizing the uncertainty in these variables before a second iteration of the analysis is carried out.

A number of techniques for sensitivity analysis have been described. However, many additional techniques for analyzing multivariate data exist that could be productively applied in a sampling-based sensitivity analysis. In particular, there are undoubtedly many pattern recognition techniques that could be successfully adapted for use in sensitivity analysis.

Sampling-based uncertainty and sensitivity analyses are usually performed for two reasons: (i) to determine the uncertainty in model predictions (e.g., to ascertain if model predictions fall within some region of concern), and (ii) to determine the dominant variables that give rise to the uncertainty in model predictions (e.g., to identify the variables on which limited research funds should be concentrated). However, there is also a third reason to perform a sampling-based uncertainty and sensitivity analysis: (iii) to verify that the model under consideration is operating correctly. Due to the concurrent variation of many model inputs and the efficacy of sensitivity analysis procedures in identifying the effects of model inputs on individual model predictions, sampling-based uncertainty and sensitivity analysis procedures provide a powerful tool for model and analysis quality assurance.

Sampling-based sensitivity analysis procedures are based on identifying patterns in a mapping between model inputs and predictions. Different procedures are predicated on the identification of different types of patterns. Thus, a procedure will not perform well if the mapping under consideration does not contain the type of pattern that that particular procedure seeks to identify. As a result, a good sensitivity analysis strategy is to use several different procedures that seek to identify different types of patterns. With this approach, there is a reasonable chance that each important model input will be identified by at least one of the procedures.

Sensitivity analysis provides a way to identify the model inputs that most affect the uncertainty in model predictions. However, sensitivity analysis does not provide an explanation for such effects. This explanation must come from the analysts involved and, of course, be based on the mathematical properties of the model under consideration. An inability to develop a suitable explanation for the effects of a particular model input is often indicative of an error in the development of the model or the implementation analysis.

References

Allen, B.C., T.R. Covington, and H.J. Clewell. 1996. "Investigation of the Impact of Pharmacokinetic Variability and Uncertainty on Risks Predicted with a Pharmacokinetic Model for Chloroform," *Toxicology*. Vol. 111, no. 1-3, 289-303.

Allen, D.M. 1971. *The Prediction Sum of Squares as a Criterion for Selecting Predictor Variables*. Report No. 23. Lexington, KY: University of Kentucky, Department of Statistics.

Anderson, T.W. 1984. *An Introduction to Multivariate Statistical Analysis*. 2nd ed. New York, NY: John Wiley & Sons.

Apostolakis, G. 1990. "The Concept of Probability in Safety Assessments of Technological Systems," *Science*. Vol. 250, no. 4986, 1359-1364.

Barry, T.M. 1996. "Recommendations on the Testing and Use of Pseudo-Random Number Generators Used in Monte Carlo Analysis for Risk Assessment," *Risk Analysis*. Vol. 16, no. 1, 93-105.

Bean, J.E., M.E. Lord, D.A. McArthur, R.J. MacKinnon, J.D. Miller, and J.D. Schreiber. 1996. "Analysis Package for the Salado Flow Calculations (Task 1) of the Performance Assessment Analysis Supporting the Compliance Certification Application." Albuquerque, NM: Sandia National Laboratories. (Copy on file in the Sandia WIPP Central Files, Sandia National Laboratories, Albuquerque, NM as WPO #40514.)

Belsley, D.A., E. Kuh, and R.E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley & Sons.

Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analysis*. New York, NY: Springer-Verlag.

Berglund, J.W. 1992. *Mechanisms Governing the Direct Removal of Wastes from the Waste Isolation Pilot Plant Repository Caused by Exploratory Drilling*. SAND92-7295. Albuquerque, NM: Sandia National Laboratories.

Berglund, J.W. 1996. "Analysis Package for the Cuttings and Spallings Calculations (Tasks 5 and 6) of the Performance Assessment Calculation Supporting the Compliance Certification Application." Albuquerque, NM: Sandia National Laboratories. (Copy on file in the Sandia WIPP Central Files as WPO #40521.)

Bernstein, P.L. 1996. *Against the Gods: The Remarkable Story of Risk*. New York, NY: John Wiley & Sons.

Blower, S.M., and H. Dowlatabadi. 1994. "Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: an HIV Model, As an Example," *International Statistical Review*. Vol. 62, no. 2, 229-243.

Bonano, E.J., S.C. Hora, R.L. Keeney, and D. von Winterfeld. 1990. *Elicitation and Use of Expert Judgment in Performance Assessment for High-Level Radioactive Waste Repositories*. NUREG/CR-5411, SAND89-1821. Albuquerque, NM: Sandia National Laboratories.

Bonano, E.J., and G.E. Apostolakis. 1991. "Theoretical Foundations and Practical Issues for Using Expert Judgments in Uncertainty Analysis of High-Level Radioactive Waste Disposal," *Radioactive Waste Management and the Nuclear Fuel Cycle*. Vol. 16, no. 2, 137-159.

Breeding, R.J., C.N. Amos, T.D. Brown, E.D. Gorham, J.J. Gregory, F.T. Harper, W. Murfin, and A.C. Payne. 1992. *Evaluation of Severe Accident Risks: Quantification of Major Input Parameters: Experts' Determination of Structural Response Issues*. NUREG/CR-4551, SAND86-1309. Vol. 2, Part 3, Rev. 1. Albuquerque, NM: Sandia National Laboratories.

Breshears, D.D., T.B. Kirchner, and F.W. Whicker. 1992. "Contaminant Transport Through Agroecosystems: Assessing Relative Importance of Environmental, Physiological, and Management Factors," *Ecological Applications*. Vol. 2, no. 3, 285-297.

Campolongo, F., A. Saltelli, T. Sørensen, and S. Tarantola. 2000. In press. "Hitchhiker's Guide to Sensitivity Analysis," *Mathematical and Statistical Methods for Sensitivity Analysis*. Eds. A. Saltelli, K. Chan, and E.M. Scott. New York, NY: John Wiley & Sons.

Chan, M.S. 1996. "The Consequences of Uncertainty for the Prediction of the Effects of Schistosomiasis Control Programmes," *Epidemiology and Infection*. Vol. 117, no. 3, 537-550.

Clemen, R.T., and R.L. Winkler. 1999. "Combining Probability Distributions from Experts in Risk Analysis," *Risk Analysis*. Vol. 19, no. 2, 187-203.

Conover, W.J. 1980. *Practical Nonparametric Statistics*. 2nd ed. New York, NY: John Wiley & Sons.

Conover, W.J., and R.L. Iman. 1981. "Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics," *American Statistician*. Vol. 35, no. 3, 124-129.

Cook, R.D., and S. Weisberg. 1982. *Residuals and Influence in Regression*. New York, NY: Chapman and Hall.

Cook, I., and S.D. Unwin. 1986. "Controlling Principles for Prior Probability Assignments in Nuclear Risk Assessment," *Nuclear Science and Engineering*. Vol. 94, no. 2, 107-119.

Cooke, R.M. 1991. *Experts in Uncertainty, Opinion and Subjective Probability in Science*. New York, NY: Oxford University Press.

Cullen, A.C., and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. New York, NY: Plenum Press.

Daniel, C., and F.S. Wood. 1980. *Fitting Equations to Data: Computer Analysis of Multifactor Data*. 2nd ed. New York, NY: Wiley-Interscience.

David, H.A. 1970. *Order Statistics*. New York, NY: John Wiley & Sons.

Draper, N.R., and H. Smith. 1981. *Applied Regression Analysis*. 2nd ed. New York, NY: John Wiley & Sons.

Efron, B., and R.J. Tibshirani. 1993. *Introduction to the Bootstrap*. New York, NY: Chapman & Hall.

Eisenhart, C. 1964. "The Meaning of 'Least' in Least Squares," *Journal of the Washington Academy of Science*. Vol. 54, 24-33.

EPRI (Electric Power Research Institute). 1989. *Probabilistic Seismic Hazard Evaluations at Nuclear Plant Sites in the Central and Eastern United States: Resolution of the Charleston Earthquake Issue*. NP-6395D. Palo Alto, CA: Electric Power Research Institute.

Feller, W. 1971. *An Introduction to Probability Theory and Its Applications*. Volume II. 2nd ed. New York, NY: John Wiley & Sons.

Fishman, G.S. 1996. *Monte Carlo: Concepts, Algorithms, and Applications*. New York, NY: Springer-Verlag New York, Inc.

Golub, G., and C.F. van Loan. 1983. *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press.

Gwo, J.P., L.E. Toran, M.D. Morris, and G.V. Wilson. 1996. "Subsurface Stormflow Modeling with Sensitivity Analysis Using a Latin-Hypercube Sampling Technique," *Ground Water*. Vol. 34, no. 5, 811-818.

Hacking, I. 1975. *The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference*. New York, NY; London, England: Cambridge University Press.

Hamby, D.M. 1994. "A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models," *Environmental Monitoring and Assessment*. Vol. 32, no. 2, 135-154.

Hamby, D.M. 1995. "A Comparison of Sensitivity Analysis Techniques," *Health Physics*. Vol. 68, no. 2, 195-204.

Harper, F.T., R. J. Breeding, T.D. Brown, J.J. Gregory, A.C. Payne, E.D. Gorham, and C.N. Amos. 1990. *Evaluation of Severe Accident Risks: Quantification of Major Input Parameters. Expert Opinion Elicitation on In-Vessel Issues*. NUREG/CR-4551, SAND86-1309. Vol. 2, Part 1, Rev. 1. Albuquerque, NM: Sandia National Laboratories.

Harper, F.T., A.C. Payne, R.J. Breeding, E.D. Gorham, T.D. Brown, G.S. Rightley, J.J. Gregory, W. Murfin, and C.N. Amos. 1991. *Evaluation of Severe Accident Risks: Quantification of Major Input Parameters. Experts' Determination of Containment Loads and Molten Core Containment Interaction Issues*. NUREG/CR-4551, SAND86-1309. Vol. 2, Part 2, Rev. 1. Albuquerque, NM: Sandia National Laboratories.

Harper, F.T., R.J. Breeding, T.D. Brown, J.J. Gregory, H.-N, Jow, A.C. Payne, E.D. Gorham, C.N. Amos, J.C. Helton, and G. Boyd.. 1992. *Evaluation of Severe Accident Risks: Quantification of Major Input Parameters. Experts' Determination of Source Term Issues.*

NUREG/CR-4551, SAND86-1309. Vol. 2, Part 4, Rev. 1. Albuquerque, NM: Sandia National Laboratories.

Harter, H.L. 1983. "Least Squares," *Encyclopedia of Statistical Sciences*. Vol. 4: Icing the Tails to Limit Theorems. Eds. S. Kotz, N.L. Johnson, and C.B. Read.. New York, NY: John Wiley & Sons. Vol. 4, 593-598.

Helton, J.C. 1993a. "Risk, Uncertainty in Risk, and the EPA Release Limits for Radioactive Waste Disposal," *Nuclear Technology*. Vol. 101, no. 1, 18-39.

Helton, J.C. 1993b. "Uncertainty and Sensitivity Analysis Techniques for Use in Performance Assessment for Radioactive Waste Disposal," *Reliability Engineering and System Safety*. Vol. 42, no. 2-3, 327-367.

Helton, J.C. 1994. "Treatment of Uncertainty in Performance Assessments for Complex Systems," *Risk Analysis*. Vol. 14, no. 4, 483–511.

Helton, J.C. 1996. "Probability, Conditional Probability and Complementary Cumulative Distribution Functions in Performance Assessment for Radioactive Waste Disposal," *Reliability Engineering and System Safety*. Vol. 54, no. 2-3, 145-163.

Helton, J.C. 1997. "Uncertainty and Sensitivity Analysis in the Presence of Stochastic and Subjective Uncertainty," *Journal of Statistical Computation and Simulation*. Vol. 57, no. 1-4, 3-76.

Helton, J.C. 1999. "Uncertainty and Sensitivity Analysis in Performance Assessment for the Waste Isolation Pilot Plant," *Computer Physics Communications*. Vol. 117, no. 1-2, 156-180.

Helton, J.C., and R.J. Breeding. 1993. "Calculation of Reactor Accident Safety Goals," *Reliability Engineering and System Safety*. Vol. 39, no. 2, 129-158.

Helton, J.C., and D.E. Burmaster. 1996. "Guest Editorial: Treatment of Aleatory and Epistemic Uncertainty in Performance Assessments for Complex Systems," *Reliability Engineering and System Safety*, Vol. 54, no. 2-3, 91-94.

Helton, J.C., and H.J. Iuzzolino. 1993. "Construction of Complementary Cumulative Distribution Functions for Comparison with the EPA Release Limits for Radioactive Waste Disposal," *Reliability Engineering and System Safety*. Vol. 40, no. 3, 277-293.

Helton, J.C., R.L. Iman, J.D. Johnson, and C.D. Leigh. 1989. "Uncertainty and Sensitivity Analysis of a Dry Containment Test Problem for the MAEROS Aerosol Model," *Nuclear Science and Engineering*. Vol. 102, no. 1, 22-42.

Helton, J.C., J.D. Johnson, M.D. McKay, A.W. Shiver, and J.L. Sprung. 1995a. "Robustness of an Uncertainty and Sensitivity Analysis of Early Exposure Results with the MACCS Reactor Accident Consequence Model," *Reliability Engineering and System Safety*. Vol. 48, no. 2, 129-148.

Helton, J.C., J.D. Johnson, A.W. Shiver, and J.L. Sprung. 1995b. "Uncertainty and Sensitivity Analysis of Early Exposure Results with the MACCS Reactor Accident Consequence Model," *Reliability Engineering and System Safety*. Vol. 48, no. 2, 91-127.

Helton, J.C., J.E. Bean, B.M. Butcher, J.W. Garner, J.D. Schreiber, P.N. Swift, and P. Vaughn. 1996. "Uncertainty and Sensitivity Analysis for Gas and Brine Migration at the Waste Isolation Pilot Plant: Fully Consolidated Shaft," *Nuclear Science and Engineering*. Vol. 122, no. 1, 1-31.

Helton, J.C., J.E. Bean, J.W. Berglund, F.J. Davis, K. Economy, J.W. Garner, J.D. Johnson, R.J. MacKinnon, J. Miller, D.G. O'Brien, J.L. Ramsey, J.D. Schreiber, A. Shinta, L.N. Smith, D.M. Stoelzel, C. Stockman, and P. Vaughn. 1998a. *Uncertainty and Sensitivity Analysis Results Obtained in the 1996 Performance Assessment for the Waste Isolation Pilot Plant*. SAND98-0365. Albuquerque, NM: Sandia National Laboratories.

Helton, J.C., J.D. Johnson, H.-N. Jow, R.D. McCurley, and L.J. Rahal. 1998b. "Stochastic and Subjective Uncertainty in the Assessment of Radiation Exposure at the Waste Isolation Pilot Plant," *Human and Ecological Risk Assessment*. Vol. 4, no. 2, 469-526.

Helton, J.C., D.R. Anderson, H.-N. Jow, M.G. Marietta, and G. Basabilvazo. 1999. "Performance Assessment in Support of the 1996 Compliance Certification Application for the Waste Isolation Pilot Plant," *Risk Analysis*. Vol. 19, no. 5, 959-986.

Hoffman, F.O., and J.S. Hammonds. 1994. "Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability," *Risk Analysis*. Vol. 14, no. 5, 707-712.

Hora, S.C., and R.L. Iman. 1989. "Expert Opinion in Risk Analysis: The NUREG-1150 Methodology," *Nuclear Science and Engineering*. Vol. 102, no. 4, 323-31.

Ibrekk, H., and M.G. Morgan. 1987. "Graphical Communication of Uncertain Quantities to Nontechnical People," *Risk Analysis*. Vol. 7, no. 4, 519-529.

Iman, R.L. 1992. "Uncertainty and Sensitivity Analysis for Computer Modeling Applications," *Reliability Technology - 1992, The Winter Annual Meeting of the American Society of Mechanical Engineers, Anaheim, California, November 8-13, 1992*. Ed. T.A. Cruse. AD-Vol. 28. New York, NY: American Society of Mechanical Engineers, Aerospace Division. 153-168.

Iman, R.L., and W.J. Conover. 1979. "The Use of the Rank Transform in Regression," *Technometrics*. Vol. 21, no. 4, 499-509.

Iman, R.L., and W.J. Conover. 1982. "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics: Simulation and Computation*. Vol. B11, no. 3, 311-334.

Iman, R.L., and J.M. Davenport. 1980. *Rank Correlation Plots for Use with Correlated Input Variables in Simulation Studies*. SAND80-1903. Albuquerque, NM: Sandia National Laboratories.

Iman, R.L., and J.M. Davenport. 1982. "Rank Correlation Plots for Use with Correlated Input Variables," *Communications in Statistics: Simulation and Computation*. Vol. B11, no. 3, 335-360.

Iman, R.L., J.M. Davenport, E.L. Frost, and M.J. Shortencarier. 1980. *Stepwise Regression with PRESS and Rank Regression (Program and User's Guide)*. SAND79-1472. Albuquerque, NM: Sandia National Laboratories.

Iman, R.L., and J.C. Helton. 1988. "An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models," *Risk Analysis*. Vol. 8, no. 1, 71-90.

Iman R.L., and J.C. Helton. 1991. "The Repeatability of Uncertainty and Sensitivity Analyses for Complex Probabilistic Risk Assessments," *Risk Analysis*. Vol. 11, no. 4, 591-606.

Iman, R.L., J.C. Helton, and J.E. Campbell. 1981a. "An Approach to Sensitivity Analysis of Computer Models, Part 1. Introduction, Input Variable Selection and Preliminary Variable Assessment," *Journal of Quality Technology*. Vol. 13, no. 3, 174-183.

Iman, R.L., J.C. Helton, and J.E. Campbell. 1981b. "An Approach to Sensitivity Analysis of Computer Models, Part 2. Ranking of Input Variables, Response Surface Validation, Distribution Effect and Technique Synopsis," *Journal of Quality Technology*. Vol. 13, no. 4, 232-240.

Iman, R.L. and M.J. Shortencarier. 1984. *A Fortran 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Models*. NUREG/CR-3624, SAND83-2365. Albuquerque, NM: Sandia National Laboratories.

Iman, R.L., M.J. Shortencarier, and J.D. Johnson. 1985. *A FORTRAN 77 Program and User's Guide for the Calculation of Partial Correlation and Standardized Regression Coefficients*. NUREG/CR-4122, SAND85-0044. Albuquerque, NM: Sandia National Laboratories.

Keeney, R.L., and D. von Winterfeldt. 1991. "Eliciting Probabilities from Experts in Complex Technical Problems," *IEEE Transactions on Engineering Management*. Vol. 38, no. 3, 191-201.

Kleijnen, J.P.C., and J.C. Helton. 1999a. "Statistical Analyses of Scatterplots to Identify Important Factors in Large-Scale Simulations. 1: Review and Comparison of Techniques," *Reliability Engineering and System Safety*. Vol. 65, no. 2, 147-185.

Kleijnen, J.P.C., and J.C. Helton. 1999b. "Statistical Analyses of Scatterplots to Identify Important Factors in Large-Scale Simulations. 2: Robustness of Techniques," *Reliability Engineering and System Safety*. Vol. 65 , no. 2, 187-197.

Kleijnen, J.P.C., and J.C. Helton. 1999c. *Statistical Analyses of Scatterplots to Identify Important Factors in Large-Scale Simulations*. SAND98-2202. Albuquerque, NM: Sandia National Laboratories.

L'Ecuyer, P. 1998. "Random Number Generation," *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*. Ed. J. Banks. New York, NY: John Wiley & Sons. 93-137.

LaVenue, A.M. 1996. "Analysis of the Generation of Transmissivity Fields for the Culebra Dolomite." Albuquerque, NM: Sandia National Laboratories. (Copy

on file in the Sandia WIPP Central Files as WPO #40517.)

LaVenue, A.M., and B.S. RamaRao. 1992. *A Modeling Approach To Address Spatial Variability Within the Culebra Dolomite Transmissivity Field.* SAND92-7306. Albuquerque, NM: Sandia National Laboratories.

Ma, J.Z., and E. Ackerman. 1993. "Parameter Sensitivity of a Model of Viral Epidemics Simulated with Monte Carlo Techniques. II. Durations and Peaks," *International Journal of Biomedical Computing.* Vol. 32, no. 3-4, 255-268.

Ma, J.Z., E. Ackerman, and J.-J. Yang. 1993. "Parameter Sensitivity of a Model of Viral Epidemics Simulated with Monte Carlo Techniques. I. Illness Attack Rates," *International Journal of Biomedical Computing.* Vol. 32, no. 3-4, 237-253.

MacDonald, R.C., and J.E. Campbell. 1986. "Valuation of Supplemental and Enhanced Oil Recovery Projects with Risk Analysis," *Journal of Petroleum Technology.* Vol. 38, no. 1, 57-69.

McKay, M.D., R.J. Beckman, and W.J. Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics.* Vol. 21, no. 2, 239-245.

McKone, T.E. 1994. "Uncertainty and Variability in Human Exposures to Soil Contaminants Through Home-Grown Food: A Monte Carlo Assessment," *Risk Analysis.* Vol. 14, no. 4, 449-463.

Meyer, M.A., and J.M. Booker. 1991. *Eliciting and Analyzing Expert Judgment: A Practical Guide.* New York, NY: Academic Press.

Mosleh, A., V.M. Bier, and G. Apostolakis. 1988. "A Critique of Current Practice for the Use of Expert Opinions in Probabilistic Risk Assessments," *Reliability Engineering and System Safety.* Vol. 20, no. 1, 63-85.

Myers, R.H. 1990. *Classical and Modern Regression with Applications.* 2nd ed.. Boston, MA: PWS-Kent Publishing Co.

Neter, J., and W. Wasserman. 1974. *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs.* Homewood, IL: Richard D. Irwin.

NRC (National Research Council). 1983. *Risk Assessment in the Federal Government: Managing the Process.* Committee on the Institutional Means for Assessment of Risks to Public Health, Commission on Life Sciences. Washington, DC: National Academy Press.

NRC (National Research Council). 1992. *Combining Information: Statistical Issues and Opportunities for Research.* Panel on Statistical Issues and Opportunities for Research in the Combination of Information, Committee on Applied and Theoretical Statistics, Board on Mathematical Sciences, Commission on Physical Sciences, Mathematics, and Applications. Washington, DC: National Academy Press.

NRC (National Research Council). 1996. *The Waste Isolation Pilot Plant: A Potential Solution for the Disposal of Transuranic Waste.* Committee on the Waste Isolation Pilot Plant, Board on Radioactive Waste Management, Commission on Geosciences, Environment, and Resources. Washington, DC: National Academy Press.

Ortiz, N.R., T.A. Wheeler, R.J. Breeding, S.C. Hora, M.A. Meyer and R.L. Keeney. 1991. "Use of Expert Judgment in NUREG-1150," *Nuclear Engineering and Design.* Vol. 126, no. 3, 313-331.

Owen, A.B. 1992. "A Central Limit Theorem for Latin Hypercube Sampling," *Journal of the Royal Statistical Society. Series B. Methodological.* Vol. 54, no. 2, 541-551.

Paté-Cornell, M.E. 1996. "Uncertainties in Risk Analysis: Six Levels of Treatment," *Reliability Engineering and-System Safety.* Vol. 54, no. 2-3, 95-111.

Payne, A.C., S.L. Daniel, D.W. Whitehead, T.T. Sype, S.E. Dingman, and C.J. Shaffer. 1992. *Analysis of the LaSalle Unit 2 Nuclear Power Plant: Risk Methods Integration and Evaluation Program (RMIEP).* NUREG/CR-4832, SAND92-0537. Albuquerque, NM: Sandia National Laboratories. Vols. 1-3.

PLG (Pickard, Lowe, and Garrick, Inc., Westinghouse Electric Corporation, and Fauske & Associates, Inc.). 1982. *Indian Point Probabilistic Safety Study.* Prepared for the Power Authority of the State of New York and Consolidated Edison Company of New York, Inc. Irvine, CA: Pickard, Lowe, and Garrick, Inc.

PLG (Pickard, Lowe, and Garrick, Inc.). 1983. *Seabrook Station Probabilistic Safety Assessment.* PLG-0300. Prepared for Public Service Company of New

Hampshire and Yankee Atomic Electric Company. Irvine, CA: Pickard, Lowe, and Garrick, Inc. Vols. 1-6, Summary Report.

Press, W.H., S.A. Teukolsky, W.T. Vetterlings, and B.P. Flannery. 1992. *Numerical Recipes in FORTRAN: The Art of Scientific Computing.* 2nd ed. New York, NY: Cambridge University Press.

Price, P.S., S.H. Su, J.R. Harrington, and R.E. Keenan. 1996. "Uncertainty and Variation of Indirect Exposure Assessments: An Analysis of Exposure to Tetrachlorodibenzene-p Dioxin from a Beef Consumption Pathway," *Risk Analysis.* Vol. 16, no. 2, 263-277.

Quade, D. 1989. "Partial Correlation," *Encyclopedia of Statistical Sciences.* Eds. S. Kotz, N.L. Johnson, and C.B. Read. New York, NY: John Wiley & Sons. Supplement Vol., 117-120.

Raj, D. 1968. *Sampling Theory.* New York, NY: McGraw-Hill.

Ramsey, J.L., and M.G. Wallace. 1996. "Analysis Package for the Culebra Flow and Transport Calculations (Task 3) of the Performance Assessment Calculations Supporting the Compliance Certification Application." Albuquerque, NM: Sandia National Laboratories. (Copy on file in the Sandia WIPP Central Files as WPO #40516.)

Rechard, R.P. 1999. *Historical Background on Assessing the Performance of the Waste Isolation Pilot Plant.* SAND98-2708. Albuquerque, NM: Sandia National Laboratories.

Risk Assessment Forum. 1997. *Guiding Principles for Monte Carlo Analysis.* M. Firestone, P. Fenner-Crisp, T. Barry, D. Bennett, S. Chang, M. Callahan, A.-M. Burke, M. Olsen, J.M. Chand, P. Cirone, D. Barne, W.P. Wood, and S.M. Knott. EPA/630/R-97/001. Washington, DC: U.S. Environmental Protection Agency.

Roache, P.J. 1993. "The SECO Suite of Codes for Site Performance Assessment," *High Level Radioactive Waste Management, Proceedings of the Fourth Annual International Conference, Las Vegas, NV, April 26-30, 1993..* La Grange Park, IL: American Nuclear Society; New York, NY: American Society of Civil Engineers. Vol. 2, 1586-1594.

Saltelli, A., and J. Marivoet. 1990. "Non-Parametric Statistics in Sensitivity Analysis for Model Output: A Comparison of Selected Techniques," *Reliability Engineering and System Safety.* Vol. 28, no. 2, 229-253.

Saltelli, A., and I.M. Sobol'. 1995. "About the Use of Rank Transformation in Sensitivity Analysis of Model Output," *Reliability Engineering and System Safety.* Vol. 50, no. 3, 225-239.

Sanchez, M.A., and S.M. Blower. 1997. "Uncertainty and Sensitivity Analysis of the Basic Reproductive Rate. Tuberculosis as an Example," *American Journal of Epidemiology.* Vol. 145, no. 12, 1127-1137.

Seber, G.A.F. 1977. *Linear Regression Analysis.* New York, NY: John Wiley & Sons.

Silverman, B.W. 1986. *Density Estimation for Statistics and Data Analysis.* New York, NY: Chapman & Hall.

Stein, M. 1987. "Large Sample Properties of Simulations Using Latin Hypercube Sampling," *Technometrics.* Vol. 29, no. 2, 143-151.

Steinberg, H.A. 1963. "Generalized Quota Sampling," *Nuclear Science and Engineering.* Vol. 15, 142-145.

Stockman, C., A. Shinta, and J.W. Garner. 1996. "Analysis Package for the Salado Transport Calculations (Task 2) of the Performance Assessment Analysis Supporting the Compliance Certification Application." Albuquerque, NM: Sandia National Laboratories. (Copy on file in the Sandia WIPP Central Files as WPO #40515.)

Stoelzel, D.M., and D.G. O'Brien. 1996. "Analysis Package for the BRAGFLO Direct Release Calculations (Task 4) of the Performance Assessment Calculations Supporting the Compliance Certification Application." Albuquerque, NM: Sandia National Laboratories. (Copy on file in the Sandia WIPP Central Files as WPO #40520.)

Stone, C.M. 1997a. *SANTOS - A Two-Dimensional Finite Element Program for the Quasistatic, Large Deformation, Inelastic Response of Solids.* SAND90-0543. Albuquerque, NM: Sandia National Laboratories.

Stone, C.M. 1997b. *Final Disposal Room Structural Response Calculations.* SAND97-0795. Albuquerque, NM: Sandia National Laboratories.

Thompson, K.M., and J.D. Graham. 1996. "Going Beyond the Single Number: Using Probabilistic Risk Assessment to Improve Risk Management," *Human and Ecological Risk Assessment.* Vol. 2, no. 4, 1008-1034.

Thorne, M.C. 1993. "The Use of Expert Opinion in Formulating Conceptual Models of Underground Disposal Systems and the Treatment of Associated Bias," *Reliability Engineering and System Safety.* Vol. 42, no. 2-3, 161-180.

U.S. DOE (U.S. Department of Energy). 1996. *Title 40 CFR Part 191 Compliance Certification Application for the Waste Isolation Pilot Plant.* DOE/CAO-1996-2184. Carlsbad, NM: U.S. Department of Energy, Waste Isolation Pilot Plant, Carlsbad Area Office. Vols. I-XXI.

U.S. EPA (Environmental Protection Agency). 1985. "40 CFR Part 191: Environmental Standards for the Management and Disposal of Spent Nuclear Fuel, High-Level and Transuranic Radioactive Wastes; Final Rule," *Federal Register.* Vol. 50, no. 182, 38066-38089.

U.S. EPA (Environmental Protection Agency). 1993. "40 CFR Part 191: Environmental Radiation Protection Standards for the Management and Disposal of Spent Nuclear Fuel, High-Level and Transuranic Radioactive Wastes; Final Rule," *Federal Register.* Vol. 58, no. 242, 66398-66416.

U.S. EPA (Environmental Protection Agency). 1996. "40 CFR Part 194: Criteria for the Certification and Re-Certification of the Waste Isolation Pilot Plant's Compliance With the 40 CFR Part 191 Disposal Regulations; Final Rule," *Federal Register.* Vol. 61, no. 28, 5224-5245.

U.S. NRC (Nuclear Regulatory Commission). 1990-1991. *Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants.* NUREG-1150. Washington, D.C.: U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Division of Systems Research. Vols. 1-3.

Wagner, H.M. 1995. "Global Sensitivity Analysis," *Operations Research.* Vol. 43, no. 6, 948-969.

Weisberg, S. 1985. *Applied Linear Regression.* 2nd ed. New York, NY: John Wiley & Sons.

Whiting, W.B., T.-M. Tong, and M.E. Reed. 1993. "Effect of Uncertainties in Thermodynamic Data and Model Parameters on Calculated Process Performance," *Industrial and Engineering Chemistry Research.* Vol. 32, no. 7, 1367-1371.

WIPP PA (Performance Assessment). 1996. *BRAGFLO, Version 4.00, User's Manual.* Albuquerque, NM: Sandia National Laboratories. (Copy on file in the Sandia WIPP Central Files as WPO #30703.)

Distribution List
SAND99-2240

Federal Agencies

US Department of Energy (4)
Office of Civilian Radioactive Waste Mgmt.
Attn: Deputy Director, RW-2
Director, RW-56
Office of Human Resources & Admin.
Director, RW-53
Office of Program Mgmt. & Integ.
Director, RW-44
Office of Waste Accept., Stor., & Tran.
Forrestal Building
1000 Independence Avenue, SW
Washington, DC 20585

US Department of Energy
Yucca Mountain Site Characterization Office
Director, RW-3
Office of Quality Assurance
MS 523
P.O.1551 Hillshire Drive, Suite A
Las Vegas, NV 89134

US Department of Energy
Research & Waste Management Division
Attn: Director
200 Administrative Road
Oak Ridge, TN 37831

US Department of Energy (6)
Carlsbad Area Office
Attn: I. Triay
G. T. Basabilvazo
D. Galbraith
M. McFadden
B. Bennington (Acting)
D. Mercer
Mailroom
P.O. Box 3090
Carlsbad, NM 88221-3090

US Department of Energy
Office of Environmental Management
Office of the Deputy Assistant Secretary
For Project Completion
Attn: M. Frei, EM-40
Forrestal Building
1000 Independence Avenue, SW
Washington, DC 20585-0002

US Department of Energy (3)
Office of Environmental Management
Office of the Deputy Assistant Secretary
For Project Completion
Attn: J. Turi, EM-43, Trevion II
19901 Germantown Road
Germantown, MD 20874-1290

US Department of Energy
Office of Environmental Management
Attn: S. Schneider, EM-44, Trevion II
19901 Germantown Road
Germantown, MD 20874-1290

US Department of Energy (2)
Office of Environment, Safety & Health
Attn: C. Borgstrom, EH-42
Office of NEPA Policy and Assistance
1000 Independence Avenue, SW
Washington, DC 20585

US Department of Energy (2)
Idaho Operations Office
Fuel Processing & Waste Mgmt. Division
850 DOE Drive
Idaho Falls, ID 83401

US Environmental Protection Agency (2)
Radiation Protection Programs
Attn: M. Oge
6401A
1200 Pennsylvania Avenue, NW
Washington, DC 20460

ACTA
Tim Hasselman
2790 Skypark Dr, Suite 310
Torrance, CA 90505-5345

Institute for Defense Analysis
Operational Evaluation Div.
Hans Mair
1801 North Beauregard Street
Alexandria, VA 22311-1772

US Department of Energy
Juan Meza
DP 51
1000 Independence Ave., SW
Washington, DC 20585

Timothy M. Barry
Chief, Science – Policy, Planning, and Evaluation
Pm 223X U.S. EPA
Washington, DC 20460

Boards

Defense Nuclear Facilities Safety Board
Attn: Chairman
625 Indiana Ave. NW, Suite 700
Washington, DC 20004

Nuclear Waste Technical Review Board (2)
Attn: Chairman
2300 Clarendon Blvd. Ste 1300
Arlington, VA 22201-3367

State Agencies

Attorney General of New Mexico
P.O. Drawer 1508
Santa Fe, NM 87504-1508

Environmental Evaluation Group (3)
Attn: Library
7007 Wyoming NE
Suite F-2
Albuquerque, NM 87109

NM Environment Department (3)
Secretary of the Environment
1190 St. Francis Drive
Santa Fe, NM 87502-0110

NM Bureau of Mines & Mineral Resources
Socorro, NM 87801

Laboratories/Corporations

Battelle Pacific Northwest Laboratories
P.O. Box 999
Richland, WA 99352

Dr. Pamela Doctor
Battelle Northwest
P.O. Box 999
Richland, WA 99352

Los Alamos National Laboratory (8)
Mail Station 5000
P.O. Box 1663
Los Alamos, NM 87545
Attn: B. Erdal, MS J591, E-ST
M. D. McKay MS F600
S. Doebling MS P946
S. Keller-McNulty MS F600
H. Martz MS F600
K. Hanson MS P940
E. Kelly MS F600
C. Nitta MS L-096

Tech Reps, Inc.
Attn: J. Chapman (1)
Loretta Robledo (2)
5000 Marble NE, Suite 222
Albuquerque, NM 87110

Westinghouse Electric Corporation (5)
Attn: Library
J. Epstein
J. Lee
R. Kehrman
P.O. Box 2078
Carlsbad, NM 88221

S. Cohen & Associates
Attn: Bill Thurber
1355 Beverly Road #250
McLean, VA 22101

Len Schwer
Schwer Engineering & Consulting
6122 Aaron Court
Windsor, CA 95492

Southwest Research Institute
Charles E. Anderson
P.O. Drawer 28510
San Antonio, TX 78284

Naval Research Laboratory
Allen J. Goldberg
Cod 5753,
4555 Overlook Avenue
S.W. Washington D.C. 20375

Naval Research Laboratory
Robert Gover
Cod 5753,
4555 Overlook Avenue
S.W. Washington D.C. 20375

B. John Garrick
Garrick Consulting
923 SouthRiver Road, Suite 204
St. George, UT 84790-6801

Christopher G. Whipple
ICF Kaiser Engineers
1800 Harrison St., 7th Floor
Oakland, CA 94612-3430

Dr. Kenneth T. Bogen
LLNL/ENV SCI DIVL – 453
P.O. Box 808
Livermore, CA 94550

David E. Burmaster
ALCEON Corporation
P.O. Box 382669
Harvard Square Station
Cambridge, MA 02238-2669

Scott Ferson
Applied Biomathematics
100 North Country Road
Setauket, NY 11733

Dr. Robert J. Budnitz
Future Resources Associates
2039 Shattuck Avenue, Suite 402
Berkeley, CA 94704

Dr. Tom Cotton
JK Research Associates
2650 Park Tower Drive, Suite 800
Vienna, VA 22180

Dr. John Kessler
Electronic Power Research Institute
3412 Hillview Avenue
Palo Alto, CA 94304-1395

D. Warner North
Decision Focus Incorporated
650 Castro Street, Suite 300
Mountain View, CA 94041-2055

Duke Engineering and Services (6)
Attn: S. David Sevougian
Joon Lee
Bryan Bullard
Kevin Mons
Ahmed Monib
Rob Howard
CRWMS M&O
1180 Town Center Drive
Las Vegas, NV 89134

Michael B. Gross
Michael Gross Enterprises
21 Tradewind Passage
Corte Madera, CA 94925

Beta Corporation Int.
Attn: E. Bonano
6613 Esther, NE
Albuquerque, NM 87109

Center for Nuclear Waste Regulatory Analysis
(CNWRA)
Southwest Research Institute
Attn: B. Sagar
P.O. Drawer 28510
620 Culebra Road
San Antonio, TX 78284

Duke Engineering and Services (2)
Attn: Banda S. Ramarao
Srikanta Mishra
9111 Research Boulevard.
Austin, TX 78758

Senes Oak Ridge, Inc (2)
Center for Risk Analysis
Attn: Steve Bartell
F. Owen Hoffman
102 Donner Drive
Oak Ridge, TN 37810

National Academy of Sciences
WIPP Panel

Tom Kiess (15)
Staff Study Director
Rm HA456
2101 Constitution Avenue, NW
Washington, DC 20418

Universities

University of New Mexico
Geology Department
Attn: Library
200 Yale Boulevard
Albuquerque, NM 87131

University of Washington
College of Ocean & Fishery Sciences
Attn: G. R. Heath
201 Old Oceanography Bldg.
Seattle, WA 98195

University of Pennsylvania
Department of Systems Engineering
Chun-Hung Chen
220 South 33rd St.
Philadelphia, PA 19104-6315

Vanderbilt University
Department of Civil and Environmental Engineering
Sankaran Mahadevan
Box 6077, Station B
Nashville, TN 37235

New Mexico State University
College of Engineering, MSC 3449
Richard Hills
P.O. Box 30001
Las Cruces, NM 88003

Prof. G. E. Apostolakis
Department of Nuclear Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139-4307

Prof. V. M. Bier
Department of Industrial Engineering
University of Wisconsin
Madison, WI 53706

Prof. M. Elisabeth Paté-Cornell
Department of Industrial Engineering and Management
Stanford University
Stanford, CA 94305

Prof. C. Frey
Department of Civil Engineering
Box 7908/NCSU
Raleigh, NC 27659-7908

Prof. Yacov Y. Haimes
Center for Risk Management of Engineering Systems
D111 Thornton Hall
University of Virginia
Charlottesville, VA 22901

Prof. D. B. Hattis
CENTED
Clark University
950 Main Street
Worcester, MA 01610

Prof. Ali Mosleh
Center for Reliability Engineering
University of Maryland
College Park, MD 20714-2115

Prof. T. G. Theofanous
Department of Chemical and Nuclear Engineering
University of California
Santa Barbara, CA 93106

Prof. Steve Hora
Institute of Business and Economic Studies
University of Hawaii, Hilo
523 W. Lanikaula
Hilo, HI 96720-4091

Prof. Thomas E. McKone
School of Public Health
University of California
Berkeley, CA 94270-7360

Prof. Herschel Rabitz
Princeton University
Department of Chemistry
Princeton, NJ 08544

Prof. Robert L. Winkler
Fuqua School of Business
Duke University
Durham, NC 27708-0120

F. E. Haskin
901 Brazos Place, SE
Albuquerque, NM 87123

Thomas H. Pigford
Department of Nuclear Engineering
4159 Etcheverry Hall
University of California
Berkeley, CA 94720

C. John Mann
Department of Geology
245 Natural History Bldg.
1301 West Green Street
University of Illinois
Urbana, IL 61801

Frank W. Schwartz
Department of Geology and Mineralogy
Ohio State University
Scott Hall
1090 Carmack Road
Columbus, OH 43210

David M. Hamby
University of Michigan
109 Observatory Street
Ann Arbor, MI 48109-2029

Rodney C. Ewing
Nuclear Engineering and Radiological Science
University of Michigan
Ann Arbor, MI 48109-2104

David Okrent
Mechanical and Aerospace Engineering Department
University of California
48-121 Engineering IV Building
Los Angeles, CA 90095-1587

W. E. Kastenberg
Department of Nuclear Engineering
University of California, Berkeley
Berkeley, CA 94720

Adrian E. Raftery
Department of Statistics
University of Washington
Seattle, WA 98195

Bruce Beck
University of Georgia
D.W. Brooks Drive
Athens, GA 30602-2152

Dr. Alison Cullen
University of Washington
Box 353055
208 Parrington Hall
Seattle, WA 98195-3055

U. M. Diwekar
Center for Energy and Environmental Studies
Carnegie Mellon University
Pittsburgh, PA 15213-3890

G. McRae
Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

M. D. Morris
Department of Statistics
Iowa State University
304 Anedecor Hall
Ames, IW 50011-1210

D. E. "Steve" Stevenson
Clemson University
Computer Science Department
442 Edwards Hall – Box 341906

Libraries

Thomas Branigan Memorial Library
200 Picacho Avenue
Las Cruces, NM 88001

Government Publications Department
Zimmerman Library
University of New Mexico
Albuquerque, NM 87131

New Mexico Junior College
Pannell Library
5317 Lovington Highway
Hobbs, NM 88240

New Mexico State Library
Attn: N. McCallan
1209 Camino Carlos Rey
Santa Fe, NM 87505-9860

New Mexico Tech
Martin Speere Memorial Library
801 Leroy Place
Socorro, NM 87801

WIPP Information Center
P.O. Box 3090
Carlsbad, NM 88221

Foreign Addresses

Atomic Energy of Canada, Ltd.
Whiteshell Laboratories
Attn: B. Goodwin
T. Andres
Pinawa, Manitoba, CANADA R0E 1L0

Dr. Arnold Bonne
Acting Head of the Waste Technology Section
Division of Nuclear Fuel Cycle and Waste Management
International Atomic Energy Agency
P.O. Box 100
A-1400 Vienna
AUSTRIA

Claudio Pescatori
AERI/NEA/OECD
LeSeine St. Germain
12 Boulevard des iles
92130 Issy-les-Moulineaux
FRANCE

Francois Chenevier (2)
ANDRA
Parc de la Croix Blanche
1-7 rue Jean Monnet
92298 Chatenay-Malabry Cedex
FRANCE

Claude Sombret
Centre d'Etudes Nucleaires de la Vallee Rhone
CEN/VALRHO
S.D.H.A. B.P. 171
30205 Bagnols-Sur-Ceze
FRANCE

Commissariat a L'Energie Atomique
Attn: D. Alexandre
Centre d'Etudes de Cadarache
13108 Saint Paul Lez Durance Cedex
FRANCE

Ghislain de Marsily
University Pierre et Marie Curie
Laboratorie de Geologie Applique
4, Place Jussieu
T.26 – 5e etage
75252 Paris Cedex 05
FRANCE

Bundesanstalt fur Geowissenschaften und
Rohstoffe
Attn: M. Langer
Postfach 510 153
30631 Hannover
GERMANY

Bundesministerium fur Forschung und
Technologie
Postfach 200 706
5300 Bonn 2
GERMANY

Gesellschaft fur Anlagen und Reaktorsicherheit
(GRS)
Attn: B. Baltes
Schwertnergasse 1
50667 Cologne
GERMANY

Forshunginstitute
GRS (2)
Attn: Eduard Hofer
B. Kryzkacz-Hausmann
Forschungsgelande Nebau 2
85748 Garching
GERMANY

Tamas Turanyi
Eotvos University (ELTE)
P.O. Box 32
1518 Budapest
HUNGARY

Jan Marivoet
Centre d'Etudes de L'Energie
Nucleaire
Boeretang 200
2400 MOL
BELGIUM

Prof. I. Papazoglou
Institie of Nuclear Technology-Radiation
Protection
N.C.S.R. Demokritos
Aghia Papakevi
153-10 Athens
GREECE

European Commission (4)
Attn: Francesca Campolongo
Karen Chan
Stefano Tarantola
Andrea Saltelli
JRC Ispra, ISIS
21020 Ispra
ITALY

Enrico Zio
Politecnico di Milano
Via Ponzia 34/3
20133 Milan
ITALY

Ricardo Bolado
Polytechnical University of Madrid
Jose Gutierrez, Abascal, 2
28006 Madrid
SPAIN

David Rios Insua
University Rey Juan Carlos
ESCET-URJC, C. Humanes 63
28936 Mostoles
SPAIN

Dr.-Ing. Klaus Kuhn
TU Clausthal Institut fur Bergbau
Erzstr. 20
38678 Clausthal-Zellerfeld
GERMANY

Shingo Tashiro
Japan Atomic Energy Research Institute
Tokai-Mura, Ibaraki-Ken, 319-11
JAPAN

Toshimitsu Homma
Nuclear Power Engineering Corporation
3-17-1 Toranomon, Minato-Ku
Tokyo 1015
JAPAN

Netherlands Energy Research Foundation ECN
Attn: J. Prij
P.O. Box 1
1755 ZG Petten
THE NETHERLANDS

Prof. Roger Cooke
Department of Mathematics
Delft University of Technology
P.O. Box 5031 2800 GA Delft
THE NETHERLANDS

Louis Goossens
Safety Science Group
Delft University of Technology
P.O. Box 5031 2800 GA Delft
THE NETHERLANDS

Prof. J.P.C. Kleijnen
Department of Information Systems
Tilburg University
5000 LE Tilburg
THE NETHERLANDS

A. Seebregts
ECN P.O. Box 1
1755 ZG Petten
THE NETHERLANDS

Willem Van Groenendaal
Tilburg University
P.O. Box 90153
5000 LE Tilburg
THE NETHERLANDS

Svensk Karnbransleforsorjning AB
Attn: F. Karlsson
Project KBS (Karnbranslesakerhet)
Box 5864
102 48 Stockholm
SWEDEN

Prof. S. E. Magnusson
Lund University
P.O. Box 118
22100 Lund
SWEDEN

Prof. Christian Ekberg
Chalmers University of Technology
Department of Nuclear Chemistry
41296 Goteborg
SWEDEN

Nationale Genossenschaft fur die Lagerung
Radioaktiver Abfalle (2)
Attn: S. Vomvoris
P. Zuidema
Hardstrasse 73
5430 Wettingen
SWITZERLAND

AEA Technology
Attn: J. H. Rees
D5W/29 Culham Laboratory
Abingdon, Oxfordshire OX14 3ED
UNITED KINGDOM

AEA Technology
Attn: W. R. Rodwell
044/A31 Winfrith Technical Centre
Dorchester, Dorset DT2 8DH
UNITED KINGDOM

Daniel A. Galson
Galson Science Ltd.
35, Market Place
Oakham
Leicestershire LE15 6DT
UNITED KINGDOM

David Draper
University of Bath
Claverton Down
Bath BA2 7AY
UNITED KINGDOM

AEA Technology
Attn: J. E. Tinson
B4244 Harwell Laboratory
Didcot, Oxfordshire OX11 ORA OQJ
UNITED KINGDOM

Prof. Marian Scott
Department of Statistics
University of Glasgow
Glasgow G12 BQW
UNITED KINGDOM

Prof. Simon French
School of Informatics
University of Manchester
Coupland 1
Manchester M13 9pl
UNITED KINGDOM

Arthur Jones
Nat. Radio. Prot. Board
Chilton, Didcot
Oxon OX110RQ
UNITED KINGDOM

Prof. M. Newby
Department of Acturial Sci and Statistics
City University
Northhampton SQ
London EC1V OHB
UNITED KINGDOM

Prof. Russell Cheng
University of Kent at Canterbury
Cornwallis Building
Canterbury, Kent, CT2 7NF
UNITED KINGDOM

B.G.J. Thompson
20 Bonser Road
Twickenham
Middlesex, TW1 4RG
ENGLAND

Ben-Haim Yakov
Department of Mechanical Engineering
Technion-Israel Institute of Technology
Haifa 32000
ISRAEL

Mauro Cicchetti
European Commission Joint Research Centre
Space Applications Institute
P.O. Box 262
Via E. Fermi 1
21020 Ispra (VA)
ITALY

Andrzej Dietrich
Oil and Gas Institute
Lubicz 25 A
31-503 Krakow
POLAND

Suzan Gazioglu
University of Glasgow
Department of Statistics
Mathematics Building
Glasgow G12 8QW
SCOTLAND

Mikhail Iosjpe
Protection Authority
Norwegian Radiation
Grini Naringspark 13
P.O. Box 55
1332 Oesteraas
NORWAY

Thierry Alex Mara
Université de la Réunion
Lab. De Génie Industriel
15, Avenue René Cassin
BP 7151
97715 St. Denis
La Réunion
FRANCE

Pedro Prado
CIEMAT,
Avda Complutense 22,
28040 Madrid
SPAIN

P. Rigney
Andra
1-7 rue Jean Monnet
Parc de la Croix Blanche
92298 – Chatenay Malabry
FRANCE

A.P. Bourgeat
Équipe Analyse Numérique
Faculté des Sciences
42023 St. Entienne Cedex 2
FRANCE

J. Jaffré
INRIA – Roquencourt
B.P. 105
78153 Le Chesnay Cedex
FRANCE

M. J. W. Jansen
Dienst Landbouwkundig
Onerzoek
Centrum voor
Plantenveredelings-en
Reproduktieonderzoek
(CPRO-DLO)
Postbus 16
6700 AA Wageningen
THE NETHERLANDS

Professor A. O'Hagen
Department of Probability and Statistics
University of Sheffield
Hicks Building
Sheffield S3 7RH
UNITED KINGDOM

Tim Belford
TU Delft
Mekelweg 4
2928 CD Delft
THE NETHERLANDS

Bernd Kraan
TU Delft
Mekelweg 4
2628 CD Delft
THE NETHERLANDS

Roberto Pastres
University of Venice
Dorsuduro 2137
30123 Venice
ITALY

Ilya Sobol
Russian Academy of Sciences
Miusskaya Square
125047 Moscow
RUSSIA

Other

Leonard F. Konikow
US Geological Survey
431 National Center
Reston, VA 22092

Dr. Bob Andrews
1280 Town Center Dr.
Las Vegas, NV 89314

Prof. Charles Fairhurst
417 5th Avenue N
South Saint Paul, MN 55075

P.S. Price
129 Oakhurst Road
Cape Elizabeth, ME 04107

Dr. Gareth Parry
19805 Bodmer Ave
Poolesville, MD 200837

Internal Addresses

MS	Org.	
0191	3010	K. W. Larson
0191	3010	J. D. Miller
0417	9800	R. G. Easterling
0434	12334	R. J. Breeding
0557	9133	T. L. Paez
0557	9133	A. Urbina
0701	6100	P. B. Davies
0708	6214	P. S. Veers
0716	6805	C. E. Olson
0716	6805	R. L. Hunter
0716	6805	P. G. Kaplan
0718	6141	L. J. Dotson
0720	6804	K. B. Sorenson
0720	6850	S. M. Howarth
0731	6805	J. T. Schneider (2)
0734	6803	L. D. Bustard
0735	6115	E. Webb
0735	6115	T. F. Corbet, Jr.
0735	6115	B. Holt
0735	6115	L. C. Meigs
0746	6411	R. M. Cranwell
0746	6411	D. J. Anderson
0746	6411	J. E. Campbell
0746	6411	L. P. Swiler
0747	6410	A. Camp
0747	6410	G. D. Wyss
0748	6413	J. J. Gregory
0748	6413	R. D. Waters
0748	6413	D. G. Robinson
0755	6233	M. D. Siegel
0759	5845	M. S. Tierney
0771	6000	W. D. Weart
0771	6800	M. Chu
0771	6800	S. Y. Pickering
0776	6852	E. J. Nowak
0776	6852	C. D. Leigh
0776	6852	H. W. Papenguth
0776	6852	K. M. Economy
0776	6852	R. D. McCurley
0776	6852	D. K. Rudeen
0776	6852	J. D. Shreiber
0776	6852	M. J. Shortencarier
0776	6852	J. L. Ramsey
0776	6852	R. P. Rechard

0776	6852	A. R. Schenker
0776	6852	R. Aguilar
0776	6852	N. D. Francis, Jr.
0776	6852	B. W. Arnold
0776	6852	C. L. Axness
0776	6852	S. P. Kuzio
0778	5355	H. A. Dockery
0778	6850	J. J. Loukota
0778	6851	P. N. Swift
0778	6851	S. G. Bertram
0778	6851	M. G. Wallace
0778	6851	G. E. Barr
0778	6851	J. H. Gauthier
0778	6851	R. J. MacKinnon
0778	6851	M. L. Wilson
0778	6851	C. T. Stockman
0778	6851	R. G. Baca
0778	6851	S. G. Bertram
0779	6805	L. S. Gomez
0779	6821	A. H. Treadway
0779	6821	N. Belcourt
0779	6821	J. Reynolds
0779	6849	J. C. Helton (25)
0779	6848	H. N. Jow
0779	6849	D. R. Anderson
0779	6849	J. E. Bean
0779	6849	F. J. Davis
0779	6849	J. N. Emery
0779	6849	J. D. Johnson
0779	6849	J. A. Rollstin
0779	6849	L. C. Sanchez
0779	6849	M. Williamson
0779	6849	P. J. Chen
0779	6849	A. R. Lappin
0779	6849	R. W. Barnard
0779	6849	W. G. Perkins
0781	5800	J. J. Danneels
0819	9211	T. G. Trucano
0825	9115	L. W. Young
0828	9133	M. Pilch
0828	9133	B. F. Blackwell
0828	9133	K. J. Dowding
0828	9133	W. L. Oberkampf
0828	9133	C. Romero
0828	9133	V. J. Romero
0829	12323	M. L. Abate
0829	12323	B. M. Rutherford
0830	12335	K. V. Diegert
0834	9100	T. Y. Chu
0847	9123	A. F. Fossum
0847	9124	D. R. Martinez
0847	9124	K. F. Alvin
0847	9211	M. S. Eldred
0847	9211	J. R. Red-Horse
0977	6524	S. M. DeLand
1035	7112	A. S. Reiser
1110	8950	L. J. Lehoucq
1137	6535	G. K. Froehlich
1137	6535	H. C. Ogden
1395	6810	P.E. Shoemaker
1395	6821	M. Marietta
1395	6821	B. L. Baker
1395	6821	M-A. Martell
1395	6821	M. E. Fewell
1395	6821	J. W. Garner
1395	6821	T. Hagdu
1395	6821	P. Vaughn
1395	6822	F. D. Hansen
1395	6822	M. K. Knowles
1395	6822	L. H. Brush
1395	6822	R. L. Beauheim
1395	6822	B. A. Howard
1395	6823	Y. Wang
1395	6860	M. Marietta (Acting)
1397	6822	L. H. Brush
1399	6850	A. Orrell
1399	6855	P. E. Sanchez
0731	6805	NWM Library (20)
9018	8940-2	Central Technical Files
0899	9616	Technical Library (2)
0612	9612	Review and Approval Desk, For DOE/OSTI