SANDIA REPORT

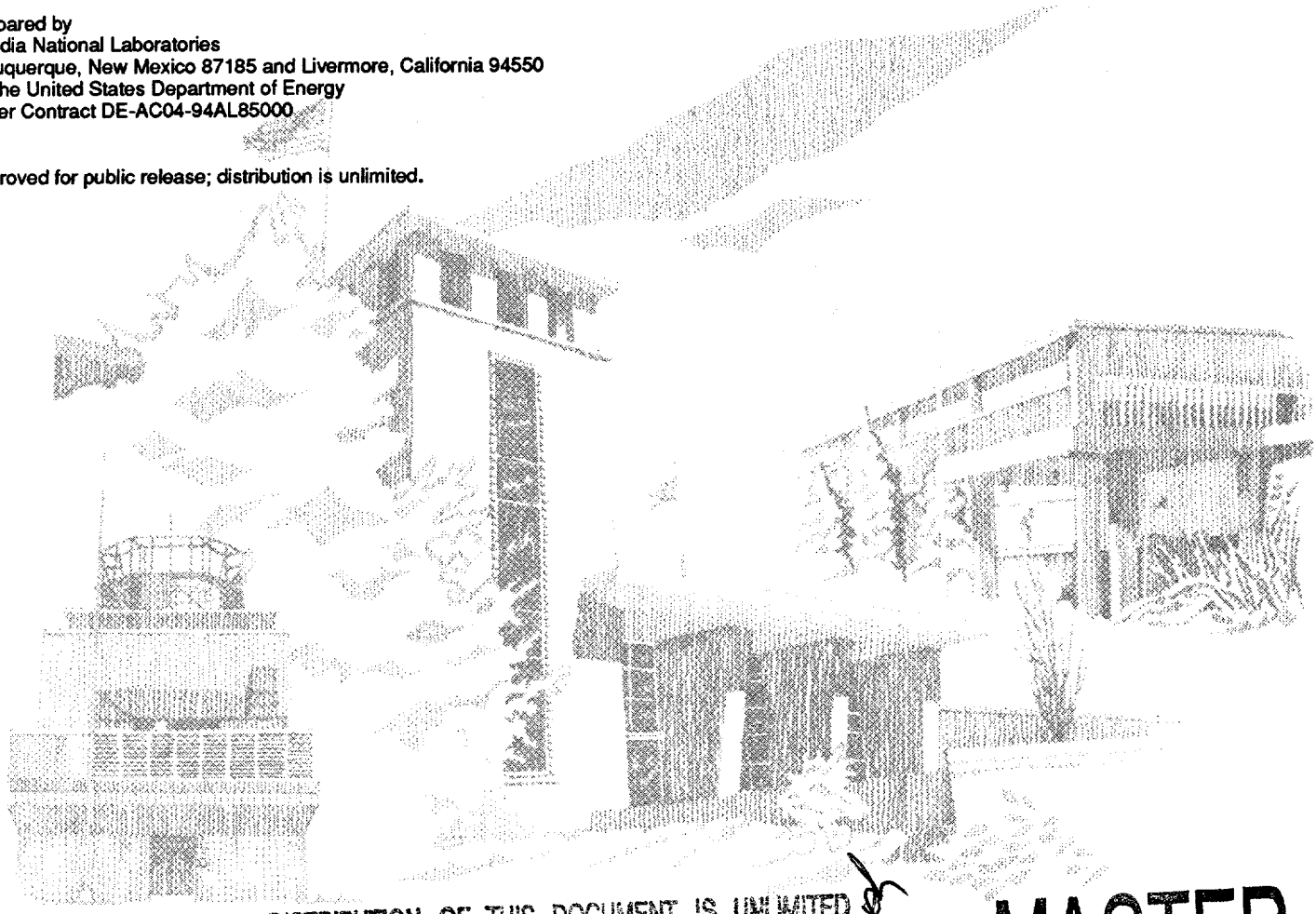
SAND95–2080 • UC–814
Unlimited Release
Printed October 1995

Yucca Mountain Site Characterization Project

Use of Stratigraphic Models as Soft Information to Constrain Stochastic Modeling of Rock Properties: Development of the GSLIB-Lynx Integration Module

M. V. Cromer, C. A. Rautman

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-94AL85000



"Prepared by Yucca Mountain Site Characterization Project (YMSCP) participants as part of the Civilian Radioactive Waste Management Program (CRWM). The YMSCP is managed by the Yucca Mountain Project Office of the U.S. Department of Energy, DOE Field Office, Nevada (DOE/NV). YMSCP work is sponsored by the Office of Geologic Repositories (OGR) of the DOE Office of Civilian Radioactive Waste Management (OCRWM)."

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
PO Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

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

NTIS price codes
Printed copy: A03
Microfiche copy: A01

SAND95-2080
Unlimited Release
Printed October 1995

Distribution
Category UC-814

Use of Stratigraphic Models as Soft Information to Constrain Stochastic Modeling of Rock Properties: Development of the GSLIB-Lynx Integration Module

Marc V. Cromer
Spectra Research Institute
Albuquerque, New Mexico 87106

Christopher A. Rautman
Geohydrology Department
Sandia National Laboratories
Albuquerque, New Mexico 87185

Abstract

Rock properties in volcanic units at Yucca Mountain, Nevada are controlled largely by relatively deterministic geologic processes related to the emplacement, cooling, and alteration history of the tuffaceous lithologic sequence. Differences in the lithologic character of the rocks have been used to subdivide the rock sequence into stratigraphic units, and the deterministic nature of the processes responsible for the character of the different units can be used to infer the rock material properties likely to exist in unsampled regions. This report proposes a quantitative, theoretically justified method of integrating interpretive geometric models, showing the three-dimensional distribution of different stratigraphic units, with numerical stochastic simulation techniques drawn from geostatistics. This integration of soft, constraining geologic information with hard, quantitative measurements of various material properties can produce geologically reasonable, spatially correlated models of rock properties that are free from "stochastic artifacts" for use in subsequent physical-process modeling, such as the numerical representation of groundwater flow and radionuclide transport.

Prototype modeling conducted using the GSLIB-Lynx Integration Module computer program, known as **GLINTMOD**, has successfully demonstrated the proposed integration technique. The method involves the selection of stratigraphic-unit-specific material-property expected values that are then used to constrain the probability function from which a material property of interest at an unsampled location is simulated. In its current Fortran implementation, **GLINTMOD** draws upon soft information from the geologic framework only at those locations for which the geostatistical search algorithm is unable to locate a user-specified number of either measured conditioning data or previously simulated grid nodes within a user-specified neighborhood. Material-property models created using the initial version of **GLINTMOD** reproduce an underlying synthetic framework model to a much greater degree than otherwise identical models simulated using only hard conditioning data.

Acknowledgments

This work was supported by the U.S. Department of Energy, Yucca Mountain Site Characterization Project, under contract DE-AC04-94AL85000. The planning documents that guided this work are Study Plan 8.3.1.4.3.2 and Work Agreement WA-0015.

Contents

Figures

Tables

Use of Stratigraphic Models as Soft Information to Constrain Stochastic Modeling of Rock Properties: Development of the GSLIB-Lynx Integration Module

Introduction

Numerical models of rock material properties will be used to evaluate the behavior of various engineered features and the waste-isolation performance of the potential Yucca Mountain nuclear waste repository site in southern Nevada. Pre- and post-closure performance assessment activities must evaluate not only the effects of the geometric distribution of rock types within the subsurface, but also the impacts of heterogeneities in the distribution of material properties within that geometry.

Nuclear Regulatory Commission licensing requirements to make quantitative predictions of repository-system behavior for extended periods into the future necessitate some type of assessment of the uncertainty associated with those predictions. A widely held assumption is that uncertainty assessment probably will take the approach of some kind of Monte Carlo simulation, in which the material properties and other parameters of a physical-process model are varied in a way that reflects the uncertainty in those values. Variations in the input parameters are propagated through a numerical representation of the physical process under investigation, and this variability is captured in a range of some performance measure—ground-water travel times or cumulative radionuclide releases, for example. Geostatistical simulation has been recognized as a preferred method for developing material-property models for input to these Monte Carlo-style uncertainty assessments, specifically because this type of simulation describes the spatial continuity of properties that may control performance of repository systems. The importance of spatial continuity in performance modeling dictates that material-property models be generated (simulated) and evaluated as a whole, in contrast to the more conventional Monte-Carlo approach that simply samples either uniform or spatially random property values from a univariate distribution.

This report describes an approach to the Monte-Carlo generation of spatially correlated material property models of a complex geologic system in a data-sparse environment for use on the Yucca Mountain Project. The approach is designed to incorporate soft, external geologic knowledge to produce geologically reasonable property models in a relatively data-poor modeling environment.

Geologic and Modeling Framework

Volcanic Geology of Yucca Mountain

Yucca Mountain is located in the southwestern Nevada volcanic field, in the southern basin-and-range province of Nevada. The geology comprises a thick sequence of Tertiary volcanic rocks, consisting largely of a series of variably welded and nonwelded ash-flow tuffs that are separated from one another by intercalated intervals of air-fall tuffs and reworked or "bedded" tuffaceous deposits. A defining characteristic of the entire Tertiary sequence, which extends to depths of more than 6,000 feet in the vicinity of the potential repository (Carr and others, 1986), is stratigraphic and stratiform layering. Layering exists on several scales and originates through the geologic pro-

cesses responsible for the emplacement of these rocks. The layering is thus relatively deterministic in nature, and this deterministic nature can be used in creating material-property models of the Yucca Mountain site.

At the largest scale, layering is induced through the very essence of the volcanic process responsible for formation of the southwestern Nevada volcanic field. Large-volume ash-flow tuffs, which form the dominant rock type present at and near Yucca Mountain, originate through massive, eviscerating eruptions of pyroclastic material from large magma chambers in the subsurface. These massive eruptions generally are associated with collapse and subsidence of rock units overlying the magma chamber into a resulting caldera. Although much of the erupted magmatic material appears to fall back into the actively subsiding caldera, many cubic kilometers of particulate material is deposited outward from the caldera as pyroclastic flows. These flows produce first-order, stratigraphic layering of eruptive products that may be separated widely in time, composition, volume, and other characteristics that determine the rock properties associated with a particular ash-flow sheet.

Other geologic processes operated within the thick, large-volume ash-flow tuffs to produce second-order, mostly stratiform, layering. Following large-volume ash-flow eruptions, the deposited material cools by loss of heat both to the underlying, former topographic surface and to the exposed, land-surface environment. Internally, the still-hot mass of glass shards and other debris compacts to varying degrees under its own weight, forming a welded tuff. Magmatic gasses still contained within the glassy, fragmented material exsolve to form a free vapor phase; these gasses act to alter both primary volcanic glass and early devitrification mineral assemblages to secondary assemblages that exhibit different material properties from their unaltered precursors. These compaction, cooling, and alteration phenomena are controlled largely by temperature and pressure gradients within the laterally extensive, tabular deposits. These gradients are, in gross form, perpendicular to the stratigraphic top and bottom of these sheet-like deposits, although marked divergence from this general geometry may occur locally. Thus, the distribution of these alteration zones, which are generally parallel to isograds, is generally stratiform. These welding and alteration zones, however, are the result of secondary phenomena superimposed on a primary stratigraphic layering, which may cause what appears to be a stratabound feature within a limited geographic area actually to cut across time-stratigraphic layering at a low angle when observed on a larger scale.

Additionally, time-stratigraphic layering is produced when the upper parts of thick pyroclastic-flow deposits are reworked by sedimentary processes. These processes winnow fines from the volcanic debris, deposit clays, and otherwise impact the nature of the material and affect the resulting material properties. Small-volume volcanic eruptions produce well-sorted air-fall deposits of coarse pumice or fine ash. The sorting of these time-stratigraphic volcanic deposits induces material properties quite different from those associated with the more-catastrophically deposited ash flows.

Finally, following complete cooling of the deposits of the southwest Nevada volcanic field and even following the cessation of volcanic activity in the region, widespread geologic processes that affected the distribution of material properties continued to operate. Surficial weathering continued to alter the near-surface portions of the different rock units. The presence of certain ground-water conditions in the subsurface may have altered any remaining volcanic glass to zeolitic materials (Bish and Aronson, 1993). Zeolitic and lesser development of clay minerals may alter hydrologic flow properties independently of the bulk

properties of the rock, because these minerals are effectively grown in-place, the result of partial dissolution of pre-existing glass and the precipitation of new mineral forms. These types of late-stage alteration zones are typically even less stratabound than early-stage cooling-related phenomena.

Stratigraphic Control of Material Properties

The distribution and variation in rock material properties within the volcanic sequence at Yucca Mountain is the result of the complex interplay of a number of different, but relatively well defined, geologic processes. The resulting features of the rock are deterministic to the extent that their observation at one location virtually assures the existence and proximity of similar and/or related features in roughly predictable locations elsewhere in the same general region. Recognition of the control of rock characteristics or material properties by stratigraphy is not new. In a qualitative sense, stratigraphic control of rock properties is the fundamental basis for recognizing a layered, genetic geologic system and for subdividing the rock column into geologic units. Major differences in rock type, phenocryst assemblages, chemical composition, and post-emplacement alteration have been used to subdivide the volcanic rocks of the southwestern Nevada volcanic field (table 1) for many decades (Christiansen and Lipman, 1965; Lipman and McKay, 1965; Lipman and others, 1966; Byers and others, 1976; Scott and Bonk, 1984; Spengler and Fox, 1989; Sawyer and others, 1994; Geslin and Moyer, 1995).

More quantitative efforts to use stratigraphy as a basis for predicting material properties are not new, either. An early attempt relevant to the Yucca Mountain Project was work by Scott and others (1983), who recognized a distinction between conventional geologic units and material-property units, and who provided some generalized, average material-property values for this non-conventional subdivision of the volcanic section. A similar, but more formal, classification was work by Ortiz and others (1965), who subdivided the volcanic rocks at Yucca Mountain into units that exhibited similar thermal, mechanical, and hydrologic characteristics. Effectively, this thermal/mechanical stratigraphic subdivision (table 1) ignored genetic considerations, and emphasized the distinction between welded and nonwelded rock types as the first-order control on material properties (figure 1). This distinction based upon degree of welding functionally translates to a first-order material-property subdivision based upon porosity. Additional second-order subdivisions related to post-emplacement alteration were also recognized by Ortiz and her coworkers. So successful was this classification at capturing material-property variability at the Yucca Mountain site (despite other weaknesses), that this "thermal/mechanical stratigraphy" has formed the basis for most subsequent large-scale performance modeling (Dudley and others, 1988; Barnard and Dockery, 1991; Barnard and others, 1992; Wilson and others, 1994).

More recently, outcrop-transect sampling studies at Yucca Mountain (Rautman and others, 1991; 1993; 1995; Rautman and Flint, 1992; Istok and others, 1994; Flint and others, 1996) have identified smaller-scale variations in material-property distributions that appear related to "microstratigraphic" or informal zonal units (figure 2). These units are also inferred to originate in the geologic processes responsible for formation of the tuffs at Yucca Mountain). For example, the histograms shown in figure 2(a) illustrate several distinctly separate populations of porosity values associated with four of the zones, all within the welded part of the Topopah Spring Tuff, recognized by Scott and Bonk (1984) in their geologic mapping of Yucca Mountain. These different, but closely related, welded zonal

Table 1: Comparison of several stratigraphic subdivisions of volcanic rocks at Yucca Mountain and encountered on the Yucca Mountain Site Characterization Project.

Geologic Unit (from Sawyer and others, 1994)		Older hydrologic zonation (modified after Scott and Bonk, 1984)			Proposed zonation of Buesch and others (USGS, written comm., 1994)	Thermal/mechanical unit (Ortiz and others, 1985)
Paintbrush Group	Tiva Canyon Tuff	Paintbrush Tuff	Tiva Canyon Member	ccr - caprock	Tpcrv	TCw
				cuc - upper cliff	Tpcn	
				cul - upper lithophysal	Tpcrl	
					Tpcpul	
				cks - clinkstone	Tpcpmn	
				cll - lower lithophysal	Tpcpll	
				ch - hackly	Tpcplnh	
				cc - columnar	Tpcplnc	
					Tpcpv3	
				ccs - shardy base	Tpcpv2	PTn
					Tpcpv1	
	Yucca Mtn. Tuff		Yucca Mtn. Mbr.			
	Pah Cyn. Tuff		Pah Cyn. Mbr.			
	Topopah Spring Tuff		Topopah Spring Member	upper nonwelded	Tptrv3	
					Tptrv2	
				tc - caprock	Tptrv1	TSw1
				tr - rounded	Tptrn	
				tul - upper lithophysal	Tptrl	
					Tptpul	
				tn - nonlithophysal	Tptpmn	TSw2
				tll - lower lithophysal	Tptpll	
				tm - mottled	Tptpln	
				tv - basal vitrophyre	Tptpv3	TSw3
				nonwelded base	Tptpv2	CHn1
					Tptpv1	
Calico Hills Formation		Tuffaceous Beds of Calico Hills		(not subdivided)	Unit 5 Unit 4 Unit 3 Unit 2 Unit 1	
					bedded tuff unit basal sandstone unit	CHn2
Crater Flat Group	Prow Pass Tuff	Crater Flat Tuff	Prow Pass Member	Not subdivided	Unit 4	CHn3
					Unit 3	PPw
					Unit 2 Unit 1	
	"bedded tuff"		"bedded tuff"		bedded tuff unit	CFUn
	Bullfrog Tuff		Bullfrog Member		Not subdivided (?)	BFw
						CFMn1
	"bedded tuff"		"bedded tuff"			CFMn2
						CFMn3
	Tram Tuff		Tram Member			TRw
						Not Recognized

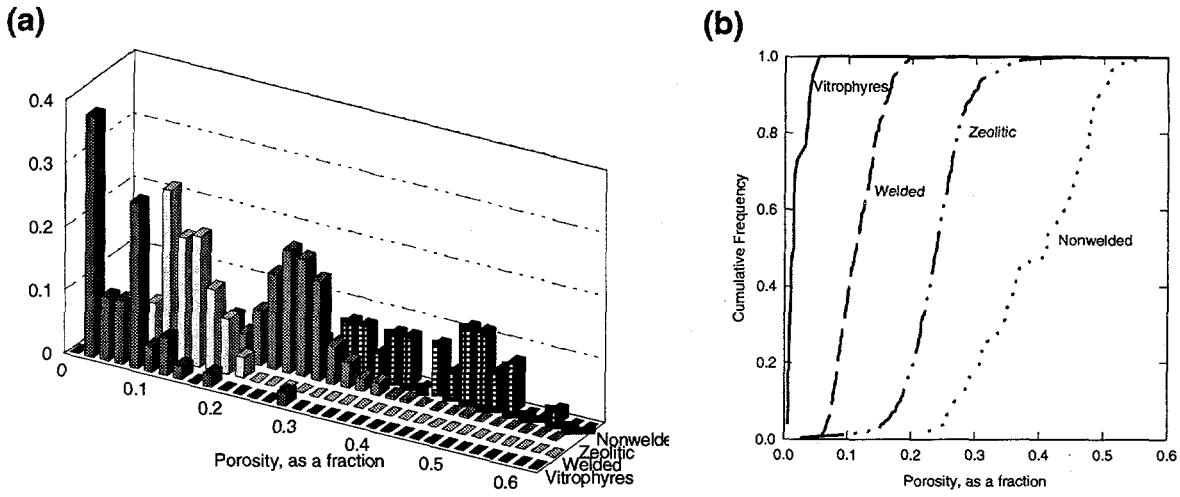


Figure 1. Porosity distributions for different major rock types at Yucca Mountain in (a) histogram format; (b) cumulative distribution function (*cdf*) format. *Cdf*s are used in the balance of this report because of their easier conceptual tie to probability distributions. Porosity data are from drill hole USW SD-9; measurements provided by L.E. Flint (U.S. Geological Survey).

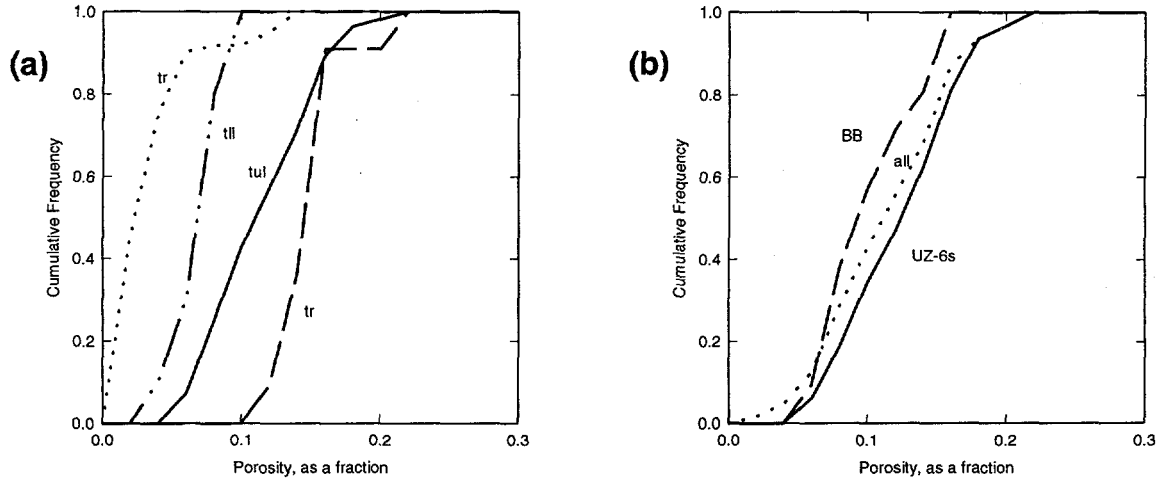


Figure 2. Histograms showing porosity values obtained by outcrop sampling of (a) four welded zones within the Topopah Spring Tuff as subdivided by Scott and Bonk (1984); and (b) separate geographic locations for the upper lithophysal zone (tul). Unit codes: tc–caprock zone; tr–rounded zone; tul–upper lithophysal zone; tll–lower lithophysal zone; all zones belong to the Topopah Spring Tuff (table 1). Location codes: UZ-6s–transect west of drill hole USW UZ-6s; BB–Busted Butte; all–all locations aggregated. Data are from Rautman and others (1993) and Flint and others (1996).

units are characterized by distinctly different modal porosity values, the ranges of measured values clearly overlap one another, and the units are distinctly heterogeneous. Selecting the porosity data for one of the four zones (the upper lithophysal zone of the Topopah Spring Tuff) and subdividing the data by geographic location [figure 2(b)] indicates that the within-unit variability is generally less than the variation between different welded zones. These same outcrop studies have also demonstrated that the internal material properties are spatially correlated (Rautman and others, 1993; Istok and others, 1994; McKenna and Rautman, 1995), although the range of spatial correlation observed depends to some extent upon the subdivision considered.

Level of Stratigraphic Subdivision

The stratigraphic-comparison chart of table 1 indicates that there are many different ways to subdivide a complex volcanic sequence, such as that present at Yucca Mountain. Given that a number of different approaches to stratigraphic classification are possible, an important issue is to determine what level of subdivision is appropriate for material-properties modeling? As the number of small-scale layers increases, additional spatial resolution is possible. However, the modeling effort increases as well, and the number of sample data per unit decreases leading to reduced statistical confidence. Note that the modeling effort includes not only the computational requirements to create the individual material-properties models, but also the "bookkeeping" required to put the various individual models back into a single property model for physical-process (e.g., flow) modeling. For example, if there are k distinct material-property layers to be represented in a material-property model, and subsequent Monte Carlo evaluation requires that n different simulations be evaluated, the total number of material property models that must be constructed is $k \times n$. The number of Monte Carlo runs involved in a typical uncertainty assessment may easily exceed one hundred. Even if the number of conditioning data is adequate, if k becomes as large as indicated by some of the columns in table 1 (30 or more), the overall rock-property modeling effort may become intractable. This is, perhaps, the most serious difficulty involved in creating detailed, deterministically layered, stochastic rock-property models.

Another major difficulty with the discrete representation of many separate layers in a material-property model is that the geometric position of the various units also must be represented. This geometric modeling requirement raises two separate issues, both involving the question of certainty in the spatial positions of these contacts. First, modeling layers as separate entities implies that there is little or no uncertainty in the location of the contacts and thicknesses of these units in three-dimensional space. If the spatial geometry of the different material-property layers is being developed from sparse subsurface data, or largely from outcrop mapping projected into the subsurface, the positions and/or thicknesses of the different units at various locations may be interpretive at best and speculative at worst. Typically, however, there is very little attempt to quantify the uncertainty associated with such geologic interpretations, and almost never is there an effort to quantify the effects of such geometric uncertainty on the higher-level physical-process. Some performance assessment modeling of Yucca Mountain (Kaplan, 1993) has shown that ground-water travel times (for example) may be quite sensitive *in general* to the thickness of different geologic units possessing markedly different hydrologic properties. However, most performance analyses have emphasized uncertainty in (uncorrelated) material properties and have accepted the geologic framework as a deterministic given (for example: Barnard and Dockery, 1991, Barnard

and others, 1992; Wilson and others, 1994). In a notable exception to this focus on material-property uncertainty, Wilson and Robey (1994) demonstrated that it is possible for cumulative radionuclide releases over a 10,000-year period (another higher-level performance measure) from one-dimensional flow columns to vary over several orders of magnitude when the thicknesses of geologic units composing the columns were modeled stochastically.

The second issue of geometric uncertainty involves the nature of the contact between one unit and its neighbor. Many, in fact most, of the finer-scale subdivisions in table 1 involve the distinction of one type or degree of roughly stratiform alteration from another, or from essentially unaltered materials. These types of "contacts" are gradational. Whereas the rock type exposed in one outcrop may be quite different from that in another outcrop some distance away, and a contact may easily be drawn between the positions of the two observations (this is the principle underlying classical field mapping), the actual change in rock type from one alteration zone to another may, in fact, be quite gradational in the subsurface. Gradations in rock type or in the nature and degree of alteration add to the uncertainty in describing and modeling the positions of these contacts, and somewhat arbitrary criteria may be required to define a contact. Unless such criteria are quite explicit and readily applied, inconsistent application of those criteria by different investigators describing the rocks at different locations may add yet another degree of uncertainty in the location of specific geologic contacts. Uncertainty in the locations of contacts between rock units propagates throughout a three-dimensional model, particularly if that model is constructed using relatively automated, computer-based algorithms.

The issues of geometric uncertainty are only complicated in the presence of faulting or other post-depositional deformation of the rock mass. The discussion thus far has involved only *stratigraphic* issues related to the modeling of a particular contact or contacts at a given spatial position given incomplete information or a gradational change in rock type. The question of spatial position related to *structural* deformation of the rock units can increase greatly the uncertainty associated with the three-dimensional position of a given contact. The more arbitrary the definition of a contact, the more difficult it is to reconstruct the effects of faulting and folding. Confounding palinspastic reconstruction at Yucca Mountain is the fact that there almost certainly have been multiple episodes of fault displacement, some of which may have taken place prior to development of some of the alteration phenomena that are invoked to define stratigraphic "units" in the more detailed classifications (table 1).

A mostly separate issue involved in gradational, and hence arbitrary, contacts is that the physical behavior modeled numerically at a sharp contact between two units of markedly different character may be quite different from the actual physical behavior occurring in real rocks exhibiting a gradational change in properties. Generally, as the magnitude of the discontinuity in material properties between two adjacent modeling cells increases, the time steps necessary to solve the partial differential equations involved in the numerical process model become progressively smaller, requiring longer execution times. Thus, a side benefit to a more realistic representation of gradational material-property transitions may be enhanced performance of physical-process modeling codes. This enhanced performance would be especially beneficial in a Monte-Carlo modeling exercise involving many hundreds of flow or flow-and-transport simulations.

Material-Property Modeling of Yucca Mountain and a Proposed Modeling Technique

A regulatory requirement exists to model the physical system of Yucca Mountain, and many different types of models have been constructed. The data available for use in constructing material-property models of Yucca Mountain vary greatly in their nature, quantitative rigor, number, and geographic distribution. For example, qualitative, descriptive information obtained from areally extensive geologic mapping, small-area outcrop studies, and linear drill-hole observations of various types (core, cuttings, geophysical log traces) are typically used to construct geologic (geometric) models. These different types of descriptive observations are generally reduced to three-dimensional spatial coordinates for various desired contact "picks," and these picks are then connected via some modeling method (interpolation). This modeling method may be nothing more sophisticated than the drawing of pencil lines on paper between two spatially positioned contacts using a straightedge (for example, the cross-sectional model of Scott and Bonk, 1984). Alternatively, one of several computer-based algorithms may be used to connect a series of spatially distributed points and define a complex surface (Ortiz and others, 1985). The interpolation method may be applied in a relatively mechanistic and automatic fashion, or a significant amount of geologic judgement and interpretation may enter into the final geologic model (Buesch and others, 1993; Fridrich and others, 1994). This interpretive information presumably makes use of what is known generally about relevant geologic processes (see, for example "Volcanic Geology of Yucca Mountain" on page 1). This type of geologic modeling in a data-sparse environment is invariably rather interpretive.

At the opposite end of the modeling spectrum, numerical models of material properties are generally constructed from quantitative laboratory or other measurements of specific physical properties (porosity, hydraulic conductivity, compressive strength). Typically, these measurements involve samples of subsurface materials from drill holes or underground excavations. Drill-hole measurements are frequently spatially biased, with abundant measurements available along a drill-hole trace; the number of individual drill holes may be relatively limited. Although definition of the process by which quantitative material properties will be modeled also involves a significant amount of geologic experience and judgement, once the modeling parameters are defined, the modeling activity itself is generally highly automated and mathematical in execution. In this respect, numerical properties modeling is frequently considered "objective."

Geostatistical modeling comprises one set of mathematical algorithms for modeling material-property values from spatially distributed quantitative observations. What distinguishes geostatistical methods from other algorithmic property-modeling approaches is the use of a calculated measure of spatial correlation, the spatial covariance function, which is more generally referred to as the *variogram.* Although geostatistical modeling methods attempt to use measures of spatial continuity developed from the specific data being modeled, there are limits to the ability of the methods to produce geologically reasonable models from sparse data. Beyond the range of spatial correlation, there is little information contained in a set of data that bears directly on the local values present or likely to exist at a given location, especially if "units" containing different properties have been aggregated. Geostatistical algorithms are unable to interpret general geologic principles to supply additional information not available in the data provided to the modeling algorithm (Rautman and Robey, 1994).

A potential solution to difficulties encountered in modeling material properties

quantitatively in a data-sparse environment is to combine interpretive geologic (geometric) modeling methods that can incorporate documented quasi-deterministic trends and features with geostatistical modeling techniques that provide a relatively objective means of integrating information from different sources and of quantifying the resulting uncertainty. A three-dimensional, computer-based geologic model of Yucca Mountain is used as soft information to constrain the material-property values that are allowed to be simulated within that domain. Such a model contains a geologist's understanding of the overall geologic environment, and it allows interpretation of the geologic unit present at every point within the modeled domain. The geologic information is "soft," in that there is no one-to-one correspondence of geologic unit with some "representative" (or average) material property. Additionally, the control exerted by the interpretive geologic model is soft, in that the presence of "hard," quantitative measurements of a particular material property will override the soft information at the locations of that hard data. The degree of influence of measured data on the simulation of unsampled locations not corresponding to a hard datum is described by the model of spatial continuity used in the modeling process.

More specifically, the incorporation of soft, constraining data into geostatistically simulated material-property models should enhance reproduction of geologically reasonable spatial continuity. Lateral variability will be well-described with little uncertainty near conditioning observations. Where there are few or no hard conditioning data, uncertainty will be greater (i.e., there will be increased variation among individual simulations), but the persistence of the material-property unit and the value about which the simulated values vary will be constrained by the expected value obtained from the framework model. Contacts of all types will be well defined where constrained by data. Sharp changes in material properties will be fairly abrupt and conditioned to the hard measurements. Gradational transitions will, in fact, be gradational. Away from actual measurements, however, both types of contacts will be relatively uncertain, reflecting the increased uncertainty that is associated with a lack of hard data. In summary, models generated using this integrated approach should be spatially correlated, stochastic representations that are consistent with both known properties data and geologic interpretations.

In the sections of this report that follow, we describe a specific implementation of this approach of using stratigraphic models as soft information to constrain stochastic modeling of rock material properties. The stratigraphic framework model consists of a geologic interpretation of data from the Yucca Mountain site similar to that described by Buesch and others (1993). The model has been developed using the Lynx Geotechnical Modeling System (GMS), a comprehensive geologic-modeling software package marketed by Lynx Geosystems, Incorporated, of Vancouver, British Columbia.[†] The geostatistical material-property modeling algorithm is taken principally from the program **SGSIM**, in the GSLIB library of geostatistical algorithms and ancillary programs developed at Stanford University (Deutsch and Journel, 1992). GSLIB programs are public-domain software. The software code that implements this integration of GSLIB geostatistical modeling with soft information taken from a Lynx geologic model is termed the ***G***SLIB-***L***ynx ***Int***egration ***Mod***ule, or program **GLINTMOD**.

[†]The use of trade, product, industry, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Theoretical Foundation of Integration

Review of Geostatistical Simulation

Geostatistical simulation is a modeling technique used to produce an attribute field that honors the spatial variability and global character of the sampled values of a variable. Simulation is relevant in situations where extreme (high or low) values of a variable—and particularly the connectivity of those extreme values—may strongly influence the operation of a physical process, such as hydrologic flow or radionuclide transport. Simulation is also relevant in situations in which it is necessary to assess characterization uncertainty, the type of uncertainty that results from less-than-exhaustive description and sampling of a geologic site.

A number of different simulation algorithms have been described historically. We have adopted the sequential simulation approach first described by Journel and Alabert (1989), and implemented in published algorithms by Gomez-Hernandez and Srivastava (1990) and Deutsch and Journel (1992). To construct a simulated model using the sequential approach, a random path is defined through a discretized (gridded) domain that will visit each grid node once and only once. At each location, a search is conducted for nearby measured data (or previously simulated values), and a local, *conditional,* cumulative-distribution function (*ccdf*) is estimated. The *ccdf* is interpreted as a probability distribution function, and a value is sampled randomly from that probability distribution. The generated value is assigned to that grid node, and the simulation process moves to the next location along the random path. Note that simulated values are added *sequentially* to the collection of values that are used in the estimation of the *ccdf* at subsequent grid nodes. Values simulated early in the process influence the simulation of later locations, and thus contribute to the propagation of spatial correlation structure within the resulting model.

Because geostatistical simulation is based on sampling from the appropriate local, conditional cumulative-distribution function of the random variable being modeled, it is necessary to describe the expectation and form of that *ccdf*. The expected value of the local *ccdf* is estimated through kriging data values *that have been transformed to represent the position of each measured datum on the population cumulative probability distribution.* To understand this type of transformation, consider the sample population portrayed in figure 3(a). Each measured value is represented by a value along the x-axis. If the data are rank ordered, each measurement also corresponds to a value measured along the y-axis, representing the cumulative probability value associated with that measurement. For example, in figure 3(a), a value of 0.1 (assume 10 percent porosity) corresponds approximately to the 80th percentile of the cumulative distribution function, or 0.80. Furthermore, as figure 3(b) indicates, each measurement can be converted through this quantile-preserving process to an equivalent value of a computationally-friendly distribution, such as the standard normal (or any other valid distribution). In the example, the original measurement value of 0.1 corresponds to a standard normal (μ=0, σ^2=1) transformed value of 0.85; the relationship is that both of these values represent the 80th percentile of their respective cumulative distribution functions (*cdf*s).

Use of the normal-score transformed values offers several computational benefits. Any kriged estimate represents the theoretical expectation of the relevant spatially distributed random variable (Journel, 1983; Isaaks and Srivastava, 1989). If we can assume a *multivariate* Gaussian random variable, then the shape or form of the distribution is fixed, and the probability distribution of that random variable is specified completely by its mean,

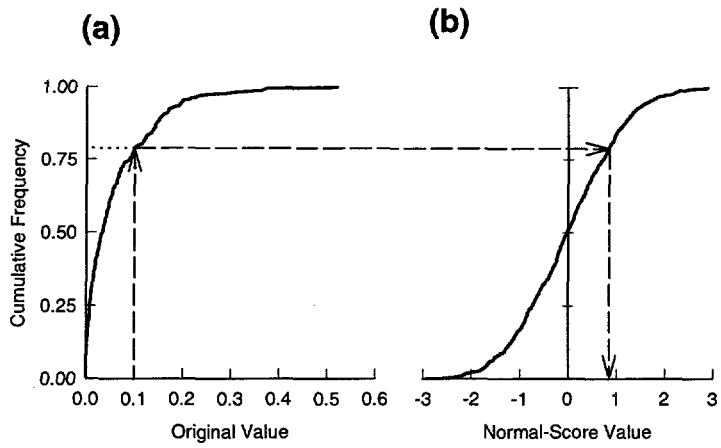


Figure 3. Graphical representation of transformation from the space of (a) a real-world variable to (b) normal-score space. The transformation preserves the quantile rankings of the data (redrawn after Isaaks and Srivastava, 1989).

its variance, and its spatial covariance. The covariance enters into the calculation through the kriging process and the use of the specified variogram model of spatial continuity. If, for convenience, we work with a *standardized* normal distribution, the variance is one by definition, and we have now specified the single remaining parameter, the expected value, through kriging the nearby transformed data. We have thus completely specified the local, conditional cumulative-distribution function required for generation of the simulated value simply through the process of kriging the transformed, surrounding data

This estimation process is represented conceptually in figure 4. The left side of the diagram represents measured values located by a search of nearby data for two different neighborhoods surrounding two unsampled locations to be simulated. These several values are transformed to their normal-score equivalents, as shown graphically in the figure (center), and the expected normal-score value for the specific location considered is computed through kriging. For simplicity, the values shown on the figure assume an isotropic spatial continuity structure and equal distances from the measured values to the point being simulated (leading to equal weighting of the data). For neighborhood number 1, all of the surrounding measurements represent low values of porosity; accordingly, the associated normal scores are low, and the *ccdf* at this location has a relatively low expected normal-score value of -1.15. For neighborhood 2, the nearby porosity measurements are much higher, the normal-score transformed values are high, and the expectation at this unsampled location is for a porosity with a normal-score value of 1.65.

The remainder of the simulation process is illustrated conceptually in figure 5. A random number, uniformly distributed in the inclusive interval from zero to one ([0,1]) is generated, and the cumulative-probability position of this random number is projected to the corresponding cumulative-probability value of the locally conditioned Gaussian *ccdf*. In figure 5, the illustrated ccdf represents that determined for neighborhood number 1 of figure 4 ($E\{Z(\mathbf{u})\} = -1.15$). In the example, a ran-

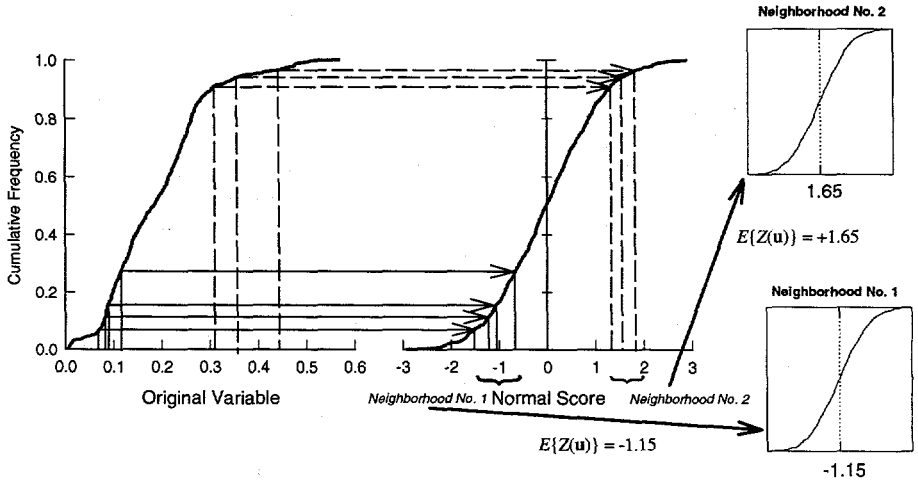


Figure 4. Conceptual representation of the process of determining the local, conditional cumulative distribution function for a spatially distributed random variable through kriging. Arrows indicate schematic projection from one distribution to another.

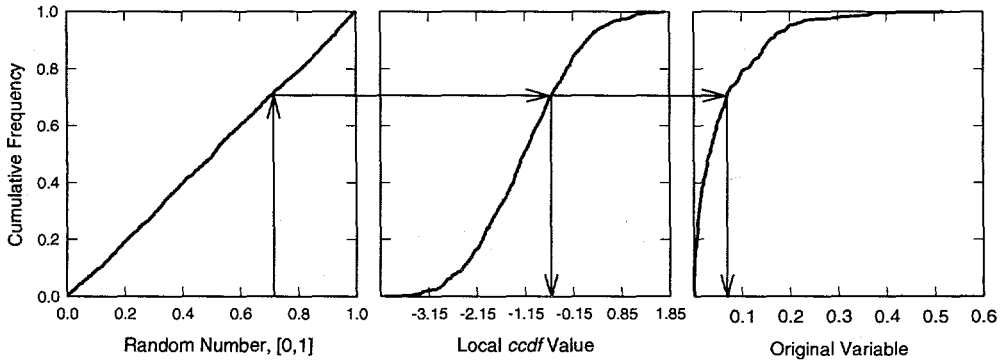


Figure 5. Conceptual representation of the simulation process. A random number uniformly distributed in the interval [0,1] is drawn and projected onto a local *ccdf* with an expected value computed by kriging nearby, standard-normal-transformed, measured values. Ultimately, this simulated normal-score value is projected back onto the standard-normal *cdf* of the original variable and the simulation process is complete.

dom value of 0.71 is generated, which translates through the projection process to a normal-score value of approximately –0.60. This value of –0.60 is added to the set of information available to condition as-yet unsimulated grid nodes, and it will influence the normal-score values assigned to those nodes in association with all other measured data and previously simulated values as weighted by the spatial covariance structure. When the simulation is complete, the value of –0.60 will be back-transformed to the space of the original variable, as suggested by the right-hand portion of figure 5, corresponding to a value of approximately 0.066 (6.6 percent porosity).

Note that a variable that exhibits a non-Gaussian histogram may still possess multivariate Gaussian spatial behavior. In fact, as illustrated by the example of figure 3, virtually any distribution may be transformed so that it exhibits a "normal" histogram without changing the spatial covariance structure of the data. This *univariate* transformation of a population, which preserves the quantile relationships of the original data, does not alter the spatial correlation structure exhibited by the variable. The variogram of the transformed values is virtually identical to the variogram of the original values, because the variogram [$2\gamma(\mathbf{h})$] is a function of the *differences* between values at two locations (Journel and Huijbregts, 1978):

$$2\gamma(\mathbf{h}) = \frac{1}{N}\sum [Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2, \quad \text{eq. 1}$$

where $Z(\mathbf{u})$ is the value of the measured variable at spatial location $\mathbf{u}$, and $Z(\mathbf{u} + \mathbf{h})$ is the value of the variable at another location a vector distance $\mathbf{h}$ from $\mathbf{u}$. N is simply the number of such pairs considered. Thus, although the absolute magnitudes of the original and transformed variables are different, for each separation distance ($\mathbf{h}$), the difference term involves values of essentially the same quantile units for a spatially correlated variable. If, for $\mathbf{h}=\mathbf{h}_1$, $[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]$ is a small value on average relative to that corresponding to $\mathbf{h}=\mathbf{h}_2$, it will be a small value whether computed in the space of the original variable or computed in the space of the transformed variable. The range of spatial correlation, defined as the separation distance at which the variogram value reaches an essentially constant level equal to the univariate variance of the data, remains constant, and the quasi-constant sill value is rescaled by the variance of the original distribution.

A practical test that the variable of interest exhibits multivariate Gaussian spatial behavior is to show that the *bi*variate cumulative distribution function (or variogram) is approximately normal. Journel (1983) states that if a random variable is bivariate normal, the sills of the indicator variograms for any particular threshold corresponding to the univariate, cumulative distribution function value, $F(z)$, can be computed as:

$$s^2 = F(z) - F^2(z) \quad , \qquad \text{eq. 2}$$

where s^2 is the magnitude of the indicator variogram sill and $F(z)$ is expressed as a decimal in the range [0,1]. For the median indicator threshold [$F(z)$=0.5], the sill should be a maximum at $0.5 - 0.5^2 = 0.25$, and the sill should decrease symmetrically for quantiles $F(z) = 1 - F(z)$. Highly asymmetric indicator variogram sills for equivalent high and low quantile thresholds (first and third quartile or first and ninth decile, for example) are fairly good evidence that a variable is not completely multivariate Gaussian in nature.

If the variable does not exhibit an acceptable degree of multivariate Gaussian spatial behavior, determining the form of the probability function is more difficult, and the *ccdf* must be estimated in a more brute-force manner using nonparametric techniques, such as indicator coding of the histogram (Journel, 1983). Indicator simulation would then be required to estimate the appropriate *ccdf* at each location to be simulated (Gomez-Hernandez and Srivastava, 1990; Deutsch and Journel, 1992). A full description of indicator geostatistics is beyond the scope of this discussion.

Simple and Ordinary Kriging

The proposed method for incorporating soft information, such as that contained in the Lynx geometric model, builds upon the basic linear-regression algorithm known as simple kriging (Journel and Huijbregts, 1978; Clark, 1979; Deutsch and Journel, 1992). In its most general form, simple kriging attempts to estimate the deviation of a spatially distributed

random function at a given location from the local expected value of that function based on a weighted linear combination of the deviations of nearby measured values from their appropriate expectations. Thus,

$$Z^*_{SK}[(\mathbf{u}) - m(\mathbf{u})] = \sum_{\alpha=1} \lambda_\alpha(\mathbf{u})\,[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)] \quad , \quad \text{eq. 3}$$

where $Z^*_{SK}(\mathbf{u})$ is the simple kriging (*SK*) estimator at an unsampled location, $\mathbf{u}$; $Z(\mathbf{u})$ is the random function model at location $\mathbf{u}$; the $Z(\mathbf{u}_\alpha)$ are the n data at locations $\mathbf{u}_\alpha$, α=1, 2,...,n, being used to construct the estimate; and $m(\mathbf{u})=E\{Z(\mathbf{u})\}$ is the *location-dependent* expected value of the variable $Z(\mathbf{u})$. The weights (λ_α) applied to the data-value deviations to obtain the simple kriging estimate are determined from the appropriate two-point covariance matrices, involving both the data-to-data covariance, $C(\mathbf{u}_\beta,\mathbf{u}_\alpha)$, α=1, 2,...,n; β=1, 2,...,n, and the data-to-unknown-location covariance, $C(\mathbf{u},\mathbf{u}_\alpha)$, α=1, 2,...,n.

This generalized formulation of the simple kriging algorithm requires prior knowledge (i.e., knowledge external to the estimation problem) of the n+1 location-specific expected values, $m(\mathbf{u})$ and $m(\mathbf{u}_\alpha)$, α=1, 2,...,n. Because of this restriction, practical implementation of simple kriging as an estimation procedure typically requires a prior decision of stationarity of the random function $Z(\mathbf{u})$, such that $E\{Z(\mathbf{u})\}=m(\mathbf{u})=m(\mathbf{u}_\alpha)=m$ is constant across the model domain. This simplification allows the simple kriging estimator to be reduced to its stationary form:

$$Z^*_{SK}(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_a(\mathbf{u})\,Z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n} \lambda_\alpha(\mathbf{u})\right] m(\mathbf{u}). \quad \text{eq. 4}$$

The covariance function necessary to calculate the various weights, λ_α, is similarly simplified through the assumption of stationarity to depend only upon the separation distance between two locations, $(\mathbf{u}-\mathbf{u}_\alpha)$, and not upon the actual positions, $\mathbf{u}$ and $\mathbf{u}_\alpha$.

Equation 4 offers us an opportunity to solve the problem of modeling material properties in a probabilistic manner throughout a domain that is locally data-poor. If we can estimate an expected value, $E^*\{Z(\mathbf{u})\}=m^*(\mathbf{u})$, at all locations within the model domain rather than simply assuming a *global* expectation, m, this value can be substituted in eq. 4 and used to constrain the expected value of the probability distribution used for simulation, even at those locations for which no "nearby" data or previously simulated values can be found.

A properly constructed three-dimensional geologic model (one without overlaps or gaps in the various volume components) shows the geologic unit *interpreted* to be present at every physical location within the modeled volume. The interpretation may or may not be correct in actual fact, but a well-constructed geologic model will provide a plausible and logically consistent geometric arrangement of rock types and other geologic units that can incorporate the full extent of a geologist's understanding of that and other similar physical systems. If there are a reasonable number of measured values for a desired material property that can be tied to the geologic units used in the geometric modeling process, it should be possible to develop an estimated expected value, $E^*\{Z(\mathbf{u})\}=m^*(\mathbf{u})$, for that material property corresponding to each appropriate aggregation of modeled geologic units, $(\mathbf{u})$. Note that here $\mathbf{u}$ corresponds to both a spatial location $[(x,y,z,)]$ and the geologic unit present at that spatial location. This development of a material-property expectation follows directly from the logic shown by the histograms of figure 2(a).

It is important that the expected values (m^*) be as unbiased as possible, because any

bias in this term of will be propagated globally throughout the model. If portions of the model domain are particularly data poor, the expectation may, in fact, be the principal source of information used in the simulation process. Over- or undersampling of different geographic regions or geologic zones may contribute to development of biased expected values. Clustering of sampling in geologically "interesting" or conveniently accessible regions is nearly inevitable in most site-characterization investigations; however, it is geostatistically undesirable and should be eliminated or compensated for in developing material-property expectations (Journel, 1983; Deutsch, 1989).

As to precedent for the practice of substituting an estimate, m^*, for the true-but-unknown expectation, $E\{Z(\mathbf{u})\}$ in eq. 3, and the degree of error that is introduced through that substitution, consider another, alternative method of estimating the random function $Z(\mathbf{u})$. This modeling method is also based on a linear combination of surrounding data, but it involves eliminating the second term of eq. 4 by requiring that the weights sum to one. This alternative method of estimation has been termed ordinary kriging (*OK*) and the estimator is obtained by simplifying eq. 4, and substituting a different set of weighting factors, υ_α, for the λ_α, which are constrained such that $\sum_{\alpha=1}^{n} \upsilon_\alpha(\mathbf{u}) = 1$. Thus, we can write:

$$Z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n} \upsilon_\alpha(\mathbf{u}) Z(\mathbf{u}_\alpha), \qquad \text{eq. 5}$$

where $Z^*_{OK}(\mathbf{u})$ is the ordinary kriging estimator at location, $\mathbf{u}$. The covariance function for the ordinary kriging system is identical to that defined for the stationary simple kriging system, with the addition of matrix entries necessary to assure that the weights sum to one.

If there are many data involved in an estimation problem, the computations involved in solving large matrix equations generally restrict the number of data values that can be considered in computing the necessary weights from the applicable covariance function. For this reason, kriging is generally applied only to the n nearby data, where the neighborhood of nearby data for consideration moves along with the location $\mathbf{u}$ being estimated.

Now if we estimate the random function $Z(\mathbf{u})$ using, for practical computational reasons, only a limited number of nearby data values in eq. 5 (n is chosen smaller than the total number of data available, N), it is apparent that we are giving greater weight to the n nearby data than to the remaining, $N–n$, more-distant data (in *OK*) or to the global, stationary mean, m (in *SK*). Indeed, in both eq. 4 and 5, the overall weight assigned to the global mean is explicitly zero. This amounts to implicitly re-estimating that prior expected value for each such neighborhood of nearby data. If we conduct a separate data search for each location $\mathbf{u}$ being estimated, we are estimating m as a function of $\mathbf{u}$, thus we have a location-specific estimate, $m^*(\mathbf{u})$.

Although $m^*(\mathbf{u})$ is only an *estimate* of the true (but unknown) location-specific expected value, m, it seems more satisfying intuitively to hold that local estimates would be reflected more accurately by deviations about a local mean (in contrast to deviations about a global and constant mean), *assuming there are sufficient data to provide a good estimate of that local mean.*

Substituting our local, estimated expectation, m^*, into eq. 4, we can write

$$Z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n} \upsilon_a(\mathbf{u}) Z(\mathbf{u}_\alpha) = \sum_{a=1}^{n} \lambda_a(\mathbf{u}) Z(\mathbf{u}_\alpha) + \left[1 - \sum_{a=1}^{n} \lambda_a(\mathbf{u})\right] m^*(\mathbf{u}), \qquad \text{eq. 6}$$

which demonstrates that the ordinary kriging

estimator is effectively a simple kriging estimate incorporating a location-dependent *estimate* of the prior expected value, $m^*(\mathbf{u}) = E\{Z(\mathbf{u})\}$. It is this implicit re-estimation of the local mean at every point that accounts for the well-known robust nature of the ordinary kriging estimator (Armstrong and Boufassa, 1988; Boufassa and Armstrong, 1989; Englund, 1990).

Were there sufficient measured material-property data to use ordinary kriging in the estimation of the local *ccdf*s in a simulation problem, there would be no need for an external source of soft data. It is precisely because we are forced to model in a data-sparse environment that the ability to add that external information, the E[Z(u)}=m* in eq. 4 and 6, can greatly improve the geologic reasonableness and robustness of the resulting models.

GSLIB-Lynx Integration

Development of a piece of computer software necessarily becomes an undertaking that emphasizes specifics. The development of **GLINTMOD** is no exception. However, the algorithm underlying the integration of GSLIB simulation subroutines with geologic framework models constructed using the Lynx GMS is completely general. The critical feature is the ability to extract from a geometric framework model the identity of the geologic unit inferred to exist at an arbitrary, unsampled location within the model domain, at which a simulated material-property value is required. Beyond the specific coding necessary to read and write common file formats and to execute the necessary translations and rotations of potentially different model-coordinate systems, neither the GSLIB subroutines nor the Lynx GMS are mandatory components of the integration process. The methodology is closely allied with the concept of sequential Gaussian simulation, because of the simplicity and general applicability of this approach for modeling continuous variables (Deutsch and Journel, 1992). However, the GSLIB subroutine **SGSIM** is only one possible implementation of the sequential simulation concept. In similar manner, there are many three-dimensional geometric modeling packages available. Virtually any other program that produces three-dimensional geometric models of geology could be substituted for the Lynx GMS as a source of soft information in the **GLINTMOD** integration approach.

Conceptual Development

Conceptually, the integration process incorporated into the initial implementation of **GLINTMOD** is fairly straightforward. The process is essentially nothing more than a check of the results of the standard **SGSIM** data search, followed by a call to the Lynx model export file if appropriate. The logic of the Lynx-GSLIB relationship is illustrated in the flow diagram of figure 6. The major part of the conventional **SGSIM** simulation procedure is located on the left-hand side of the flow diagram (part ① of figure 6). Additions to the simulation logic are indicated by boxes and logical connections shown in the heavier lineweight items mostly on the right-hand side of the figure (②).

The sequential Gaussian simulation process is initiated in the conventional manner (Deutsch and Journel, 1992) by mapping the hard conditioning data into model coordinates and defining the random simulation path through the model domain. At each node to be simulated, **SGSIM** searches a user-specified neighborhood surrounding that location for hard conditioning data. These data presumably provide the most reliable, local conditioning information, and thus they are examined first. The algorithm then searches for previously simulated grid nodes. **SGSIM** allows the user to specify both a minimum and a maximum number of original data to use in simulating a vacant grid node, as well as to specify the

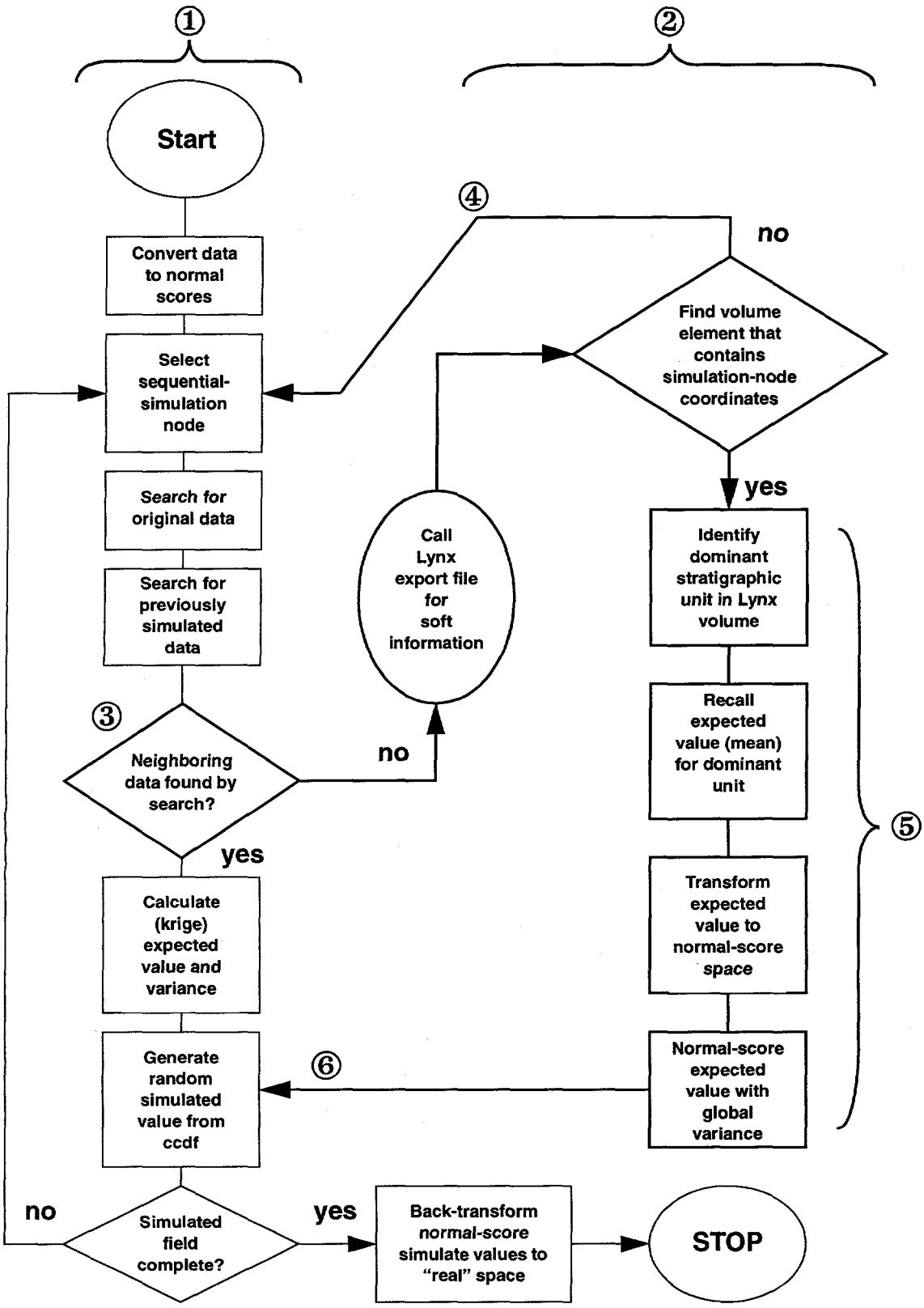


Figure 6. Flow chart illustrating the logic of the GSLIB-Lynx integration algorithm. Circled numbers are keyed to discussion in text.

maximum number of previously simulated values for consideration.

The number of previously simulated values that can be located by the search strategy during the early stages of simulation is small. In a data-sparse modeling environment, it is also likely that the search process will not find any of the original conditioning data. The unmodified SGSIM algorithm handles this lack of local information either by leaving the grid node unsimulated or by selecting a value at random from the global cumulative distribution function. However, as indicated by the first decision point in the logic diagram of figure 6 (③), **GLINTMOD** recognizes that inadequate conditioning information has been found within the search neighborhood and issues a call to the Lynx export file containing the stratigraphic-unit information for soft data.

GLINTMOD then reconciles the **SGSIM** model coordinate system with the Lynx model coordinate system and uses the proper three-dimensional spatial description to locate the proper Lynx-model cell containing the grid node to be simulated. **GLINTMOD** assumes that the simulation grid forms a subset of the host Lynx model. If the current simulation node falls outside the Lynx model domain, **GLINTMOD** returns a null value to the simulation algorithm, effectively skipping simulation at that location (④). This procedure restricts simulation to regions of interest defined by the Lynx geologic model. The principal reasoning behind this procedure is to allow simulation to be limited to "meaningful" Lynx volume elements, such as below the topographic surface or to a specific geologic unit.

Once **GLINTMOD** has identified the appropriate Lynx-model cell, the algorithm queries the Lynx export file and obtains the identity of the dominant geologic unit contained within that cell. The expected material-property value corresponding to that geologic unit is then identified, transformed to normal-score space using the same normal-score transform used for the hard conditioning data, and passed back to the main **SGSIM** simulation program (⑤). This process effectively repositions the central tendency of the cumulative distribution function, conditioning it to the subglobal, unit-specific expected value. The actual material-property value assigned to the current grid node is sampled randomly from this new unit-specific *ccdf* (⑥), and the likelihood is that the simulated value will be consistent with the material-property values that are generally associated with that unit. However, uncertainty considerations dictate that there is a finite probability that the simulated material property may be quite different. The newly simulated node is added to the set of "previously simulated values," and the simulation process continues until the entire simulated grid is completed.

Gridded Model Representation

The Lynx GMS currently has dimension limitations that allow a maximum of 200 uniformly spaced grid cells in each principal direction. This limitation makes it likely that the Lynx grid, in any particular model, will need to be constructed in such a way that will maximize its areal coverage and minimize heterogeneity of unit classifications within each grid cell.

The two modeling packages must use a consistent method to transfer information between the Lynx geometric model and the GSLIB simulation routine. The most direct method of passing consistent location information between the two software packages is to use "real-world" coordinates. On the Yucca Mountain Project, these consist of Nevada state plane coordinates (east, north, and elevation). The Lynx GMS package simultaneously maintains internal references to grid blocks both in terms of real-world, global coordinates and of a user-specified model-grid coordinate system. The user-specified grid may be

defined in any desired orientation and consists of volume elements that are identified by their centroid coordinate location. GSLIB modeling algorithms, however, rely solely upon an integer-indexed orthogonal grid that is referenced to a user-specified origin.

If both the Lynx geometric model and the GSLIB simulated material properties model use the same directions as their principal coordinate axes [e.g., (x,y,z) = east, north, elevation], conversion between coordinate systems is relatively straightforward, and consists simply of multiplying the relevant index by the grid-block size and adding offsets related to local grid origin. However, if the geology being modeled requires use of an off-angle local grid, translation between model coordinates and global coordinates will require both rotation and translation computations.

Programming Considerations

GLINTMOD was developed using structured Fortran-77 programming consistent with that of the GSLIB software library. A single main subroutine executes the integration of the two models by calling both new and existing subroutines and functions from the **SGSIM** simulation program of GSLIB. Most of the major steps in the integration process (figure 6) are written as independent subroutines. This programming approach imparts a modular nature to the overall integration. However, the approach also requires consistency in coding between the pre-existing GSLIB subroutines and the new subroutines that provide the link between the two main packages. Each subroutine is independent of its calling program. As such, each program segment could be developed separately and tested against manual calculations for verification of proper function.

Prototype Model Demonstration

The completed **GLINTMOD** interface was tested in a prototype mode to demonstrate the functionality of the module. Prototype, synthetic models were devised that tested specific aspects of the desired integration function. One prototype modeling exercise consisted of a simple, two-dimensional model grid in which the coordinate axes of the Lynx "model" file and the **SGSIM** model were parallel to each other. Another test exercise consisted of a simulated model for which the axes were rotated with respect to the global coordinate system used by the constraining geometric model.

Parallel-Grids Example

The parallel-grids prototype modeling exercise was designed to test performance of **GLINTMOD** in a situation that was known to produce geologically unrealistic results when soft information was not incorporated. This type of model would be similar to some of those produced by Rautman and Robey (1994; Wilson and others, 1994) in three-dimensions, in which conditional simulations were constructed in regions extending significantly beyond the range of available conditioning data. These earlier models, while statistically similar to one another and to the hard, conditioning measurements, exhibited the presence of stochastic artifacts that appear as rock units in locations incompatible with geologic understanding (figure 7).

A hypothetical model of layered stratigraphy was defined consisting of four geologic units, as shown in figure 8. This conceptualized geologic model is broadly similar to the layering exhibited by ash-flow deposits at Yucca Mountain. The overall layering was defined as horizontal and parallel to the orthogonal simulation grid used for **SGSIM**. The hypothetical model was then discretized to simulate the format of an export file from a typical Lynx geometric block model. Soft information, consisting of an "expected" porosity value appropriate to each different geologic unit, was assigned to each discretized block location.

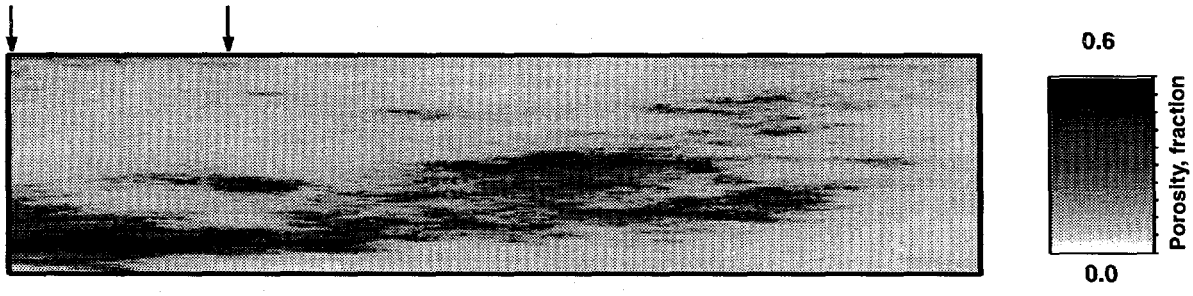


Figure 7. Cross section showing simulated stochastic artifacts that appear as high-porosity rock units in geologically unreasonable positions in the right-hand three-quarters of the figure. Geology consists of horizontal, alternating low-high-low porosity units as represented on the left-hand portion of the figure. Arrows indicate positions of conditioning drill holes. From information in Rautman and Robey, 1993.

A set of artificial drill-hole records (figure 9) was fabricated based on the conceptual geologic model to serve as hard, conditioning data for the simulator. These data were generated in a manner intended to be consistent with the general nature and availability of field data encountered during previous modeling exercises at Yucca Mountain. The data are irregularly spaced in the horizontal dimension, in a manner intended to represent somewhat-clustered drill-hole locations. The vertical spacing is much closer than the horizontal spacing, again consistent with the actual spacing of samples along drill holes. The vertical extent of data along any particular drill-hole trace varies as well; some drill holes contain intervals of missing data, much as might be encountered in the field as a lost-core zone. A statistical summary of the hypothetical conditioning data is presented in table 2.

The drill hole records have been positioned within the model domain to demonstrate how the addition of soft, constraining information will improve the simulation of porosity values consistent with the underlying geologic model, specifically in those portions of the domain that are beyond the range of spatial correlation captured by the variogram. Drill holes 1 and 2 are relatively close to each other. Together, these conditioning data plus the range of spatial correlation provided as input to the simulator will constrain stochastic modeling in this portion of the model. Note that these holes do not extend vertically through the entire model domain. Drill hole 3, on the other hand contains a long record that provides information on every unit. However, there are no supporting data within the range of correlation (arbitrarily set equal to 1,000 feet or one quarter of the domain width) that will allow this information to propagate beyond the immediately adjoining region, as required by the underlying conceptual geologic model. This is the modeling situation that has, in the past, allowed unconstrained generation of stochastic artifacts inconsistent with the conceptual geologic framework (figure 7).

Geostatistical simulation of porosity was performed on a two-dimensional grid discretizing the domain shown in figure 9. The simulation grid consists of 40 cells horizontally and 30 cells vertically; dimensions of these cells are assumed to be 100 feet long and 50 feet high. The geometric model (figure 8) was discretized to represent a Lynx block model comprised of blocks 200 feet long by 100 feet high covering the same model extent. This difference in scale of resolution is

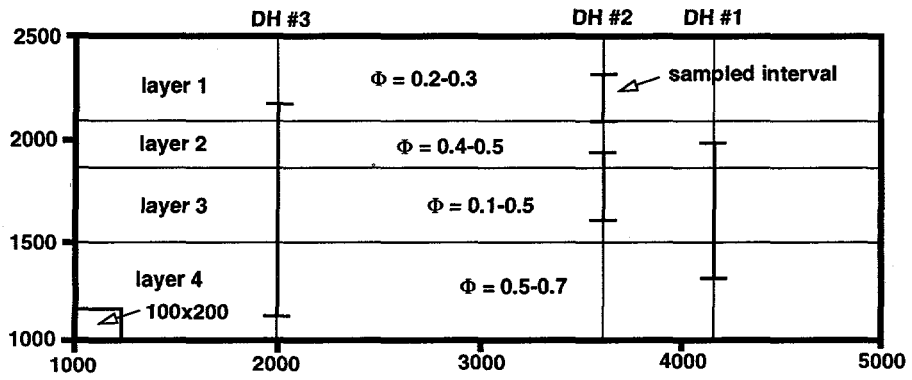


Figure 8. Hypothetical, four-layer geologic model used as the basis of the parallel-grids prototype model. Locations of hypothetical drill holes indicated, as are sampled intervals within drill holes. Discretization of Lynx model uses 100x200-ft cells (illustrated in lower left corner). Dimensions in feet.

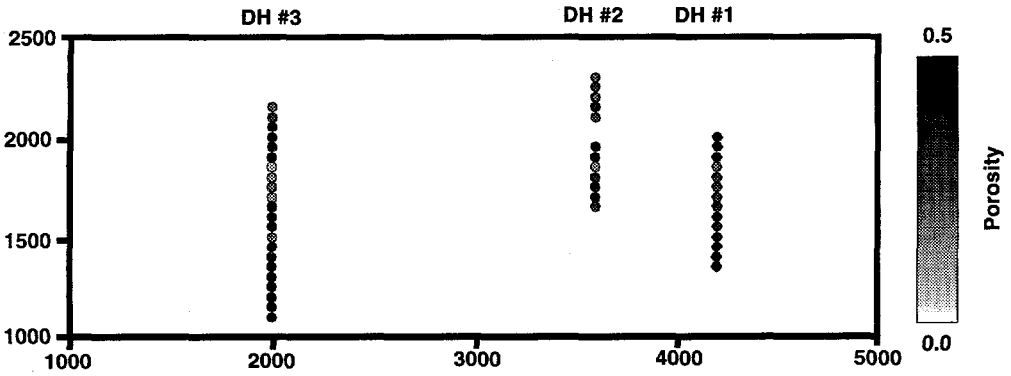


Figure 9. Three fabricated drill-hole porosity records generated based on the hypothetical, four-layer geologic model of figure 8. Dimensions in feet; porosity as a fraction.

Table 2: Statistical Summary of the Four Hypothetical Geologic Units from the Parallel-Grids Prototype Model

[Values are assumed porosity as a fraction; the coefficient of variation is defined as the standard deviation divided by the mean]

	Layer 1	Layer 2	Layer 3	Layer 4	Composite
Mean	0.23	0.42	0.27	0.60	0.37
Standard Deviation	0.04	0.05	0.10	0.07	0.16
Coefficient of Variation	17%	12%	37%	12%	44%
Minimum	0.20	0.35	0.10	0.50	0.10
Maximum	0.30	0.50	0.90	0.75	0.75
No. of Data	7	9	21	11	48

intended to capture an anticipated emphasis on detailed resolution of material properties for subsequent physical-process (i.e., flow) modeling while geometric information probably will be extracted from a framework model of the entire Yucca Mountain site, for which the block size is controlled by the 200-cell limitation (see page 18).

Modeling parameters for the simulation were selected specifically to demonstrate the impact of soft information on portions of the model not adequately constrained by hard drill-hole data. A simple variogram model was constructed using a single spherical structure with a horizontal range of spatial correlation equal to one quarter of the model domain (1,000 feet). This range of correlation is necessarily a composite range, such as that discussed by Rautman and Flint (1992). The vertical range of correlation was provided through assignment of a vertical-to-horizontal anisotropy ratio of 0.15; thus the vertical range of correlation is 150 feet. A small nugget effect of one-tenth the total sill was added to reflect added variability of values at very short separation distances.

The search neighborhood for the simulation algorithm was specified as 300 feet in the horizontal dimension, with an anisotropy ratio of 0.15. This value is smaller than might be used in a real modeling exercise of similar size and extent. The restricted search neighborhood was imposed to force **GLINTMOD** to refer to the Lynx export file for soft information. Recall that the initial version of **GLINTMOD** branches to the Lynx model "on demand" only at those grid nodes where no hard data or previously simulated values are found within the search neighborhood.

Figure 10 shows a single stochastically simulated model of porosity values for the model domain produced using only the hard drill hole data (no **GLINTMOD**). This model honors the hard data at the location of those data (by construction), and the statistical character of the overall model is quite similar to that of the input data (figure 11). The model is somewhat horizontally layered as specified by the input variogram model and its associated anisotropy ratio, and these layers, which are identifiable as lenses of similar grey level, can be observed to persist for distances of up to about 1,000 feet (one-quarter of the 4,000-foot model domain). Correlation within individual lenses is stronger at short distances, and the degree of that correlation progressively diminishes as separation distances approach 1,000 feet.

However, the layers of figure 10 are markedly more lenticular and discontinuous than the desired, underlying conceptual geologic model (figure 8). Closer observation indicates that the concept of a four-layer model is captured to a slight degree. Generally, there is a tendency for higher porosity values (darker greys) to occur in the lower one-half of the model and again somewhat in the middle part of the upper-half. A weak band of lower porosity values (lighter greys) occurs approximately in the middle of the figure and again, although to a lesser extent, near the top. Nevertheless, it is clear that this model is only poorly conditioned by the data. This effect of inadequate conditioning data has been accentuated by the choice of a limited search neighborhood, which allows data values to be located only within about one-third of the actual range of spatial correlation.

Figure 10 serves as a basis of comparison for a property model incorporating the soft information from a Lynx-style geometric model, which is shown in figure 12. During simulation of this porosity model, if no conditioning data (or previously simulated values that presumably already reflect the influence of hard conditioning information) were located within the 300 by 45-foot, elliptical search neighborhood, the **GLINTMOD** main program issued a call for soft information from the

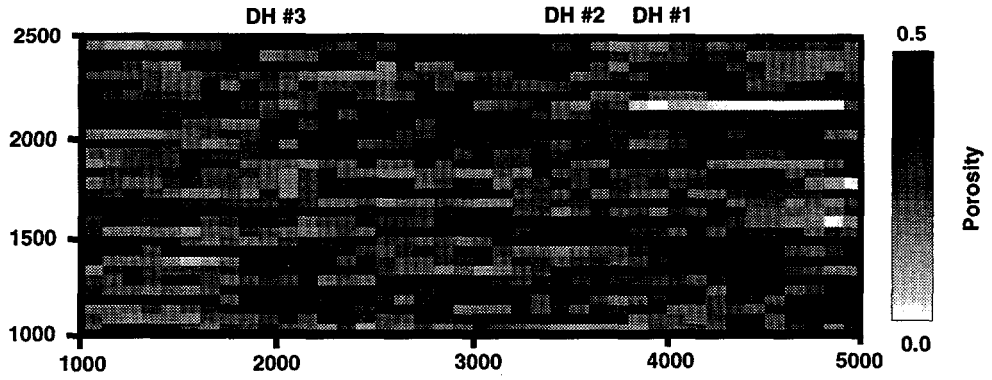


Figure 10. Simulated porosity model generated using only the hard conditioning drill-hole data of figure 9. Dimensions in feet, porosity as a fraction.

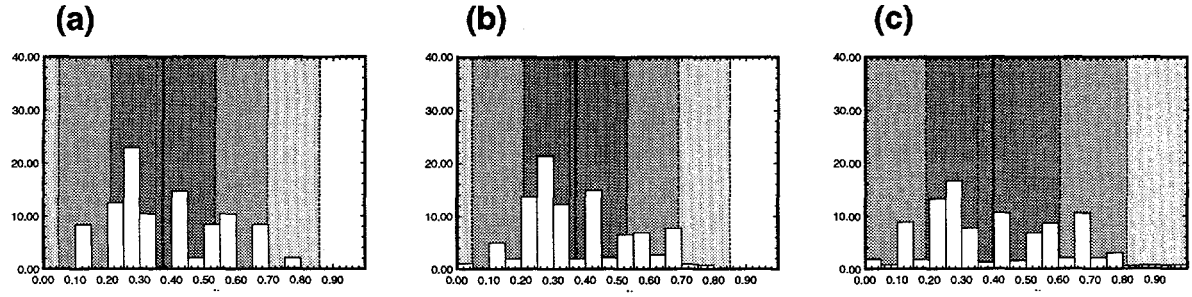


Figure 11. Histograms comparing (a) drill-hole data from figure 9; (b) porosity model simulated using only hard conditioning data; and (c) porosity model simulated using hard conditioning data and soft information from conceptual model of figure 8. Vertically shaded regions each represent one standard deviation.

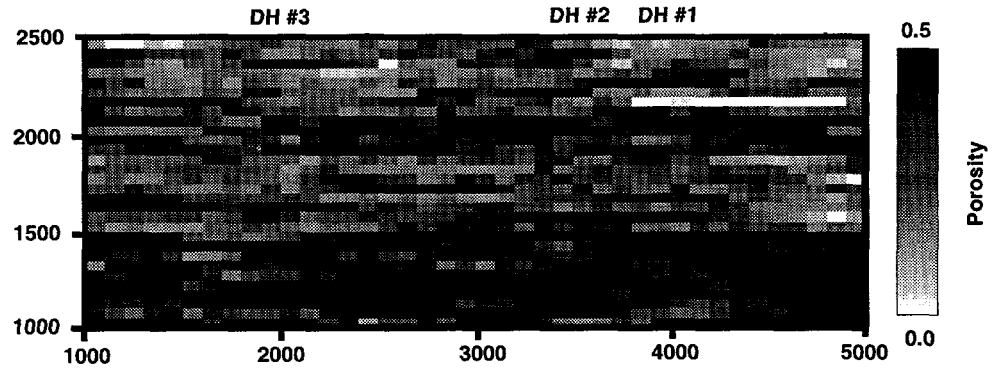


Figure 12. Simulated porosity model generated using both the hard conditioning drill-hole data of figure 9 and material-property expectations from conceptual geologic model of figure 8. Dimensions in feet, porosity as a fraction.

Lynx export file (figure 6, ③). The soft information represents an *a priori* expectation of the material property, given that the otherwise unconditioned grid node belongs to a specified geologic unit. After a specific value for the simulation node is sampled from the unit-specific distribution having that *a priori* expected value, the simulation process moves to the next node along the random path, and the simulated value enters the set of grid nodes that will be searched and incorporated into subsequent simulated nodes as appropriate.

The differences between figures 10 and 12 are obvious. Porosity values are distinctly more continuous across the entire model domain in the simulation produced using soft information extracted from the geometric model via **GLINTMOD**. In particular, reproduction of the lowermost, high-porosity unit, which was conditioned by only the one drill hole (hole number 3; figure 8) is much improved over the model of figure 10. The existence and persistence of this particular unit at distances of more than about one-half the spatial correlation length away from the drill hole location is due solely to the soft information incorporated into this model. The persistence of the alternating low- and high-porosity units in the upper portion of the model is also much improved over the model that uses only the hard data.

Similar to the model shown in figure 10, the **GLINTMOD** version of figure 12 honors the conditioning data at the drill hole locations. The variogram is honored as well; compare, for example, the horizontal and vertical extent of the various distinct-color lenses within the four generalized geologic units. However, as indicated by the histogram of the **GLINTMOD**-produced model in figure 11, the statistical character of the input data is not strictly reproduced. Additional information not contained in the histogram has been added to the model, only in this case, that additional information consists of a geologist's conceptual representation vis-a-vis the Lynx export file. It is unclear exactly to what degree the statistics of the hard conditioning data should be reproduced. If the available drill hole data are representative of the overall domain to be modeled, presumably the two different types of information are consistent, and the statistics of the simulated model will resemble those of the input hard data. If the sampled hard data are biased with respect to the overall geologic model, it is only reasonable to expect that the statistical properties of that overall domain will depart from those of the limited and nonrepresentative hard data. In the present case, note that the lowermost high-porosity unit of figure 8 is represented only poorly by one drill hole (compare figure 9). The histograms of figure 11 reflect the influence of this undersampling of a laterally extensive, high-porosity layer.

Note that although the lateral continuity of the four conceptual geologic units is much improved in the **GLINTMOD**-generated version of figure 12, the exact position of the "contacts" between units and the exact magnitude of porosity values at any specific location is uncertain. This uncertainty is as it should be for real geologic units sampled by only three drill holes extending through only part of the vertical extent of those units. Although soft, interpretive geologic information, such as would be represented by a real Lynx geometric model, has been added, that addition does not result in a single-valued, deterministic representation of material properties.

Figures 13 and 14 illustrate aspects of this geologic uncertainty. Figure 13 shows an expected-value profile for the model generated using only hard drill hole data. Figure 14 is the equivalent expected-value representation of the model generated using **GLINTMOD** and the Lynx export file. Both illustrations were prepared by simulating ten different stochastic realizations, all of which are equally likely and depend only on the random number seed used to initiate the simulation process. The ten sto-

chastic images were then averaged, pixel by pixel, and the resulting mean (expected) value displayed on the cross sections. As the number of simulations combined into an expected-value summary becomes large, the result typically approaches the value modeled through the geostatistical algorithm known as ordinary kriging (Deutsch and Journel, 1992; Rautman and Istok, in press).

The resemblance of figure 14 to the conceptual geologic model of figure 8 is particularly noticeable. All four layers are reproduced quite well, although there is a modest amount of internal variability that may be related principally to the small number of simulations (10) averaged to produce the expected-value representation. Some variability is also caused by the different values contained in the drill hole data. The expected-value profile of figure 13 does a significantly poorer job of reproducing the expected, conceptual geologic model. This failure is attributed principally to the unrealistically small search neighborhood specified for the modeling exercise, a neighborhood that was selected specifically to push the limits of hard-data simulation.

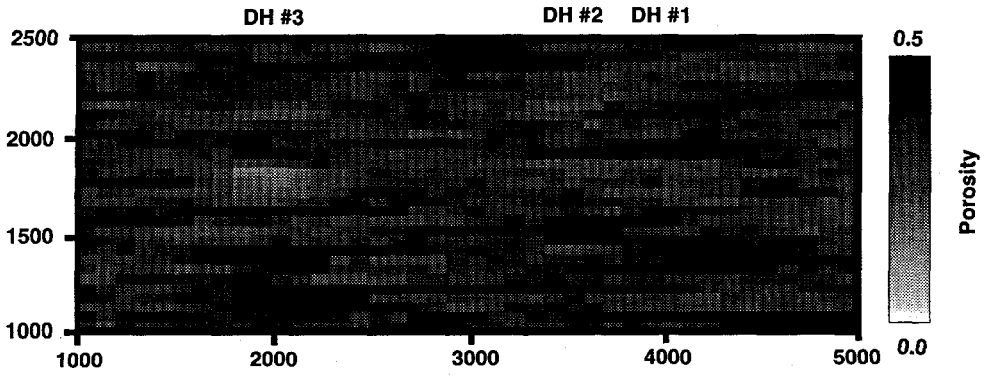


Figure 13. Expected-value profile produced using only hard drill-hole data from figure 9 to condition ten individual stochastic realizations of porosity. Compare to single realization of figure 10. Dimensions in feet, porosity as a fraction.

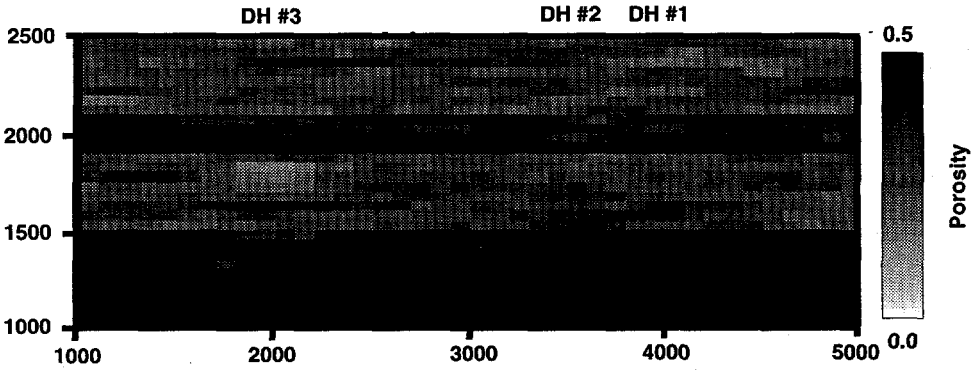


Figure 14. Expected-value profile produced using both hard drill-hole data from figure 9 and soft geometric information from figure 8 to condition ten individual stochastic realizations of porosity. Compare to single realization of figure 12. Dimensions in feet, porosity as a fraction.

Rotated-Grids Example

A second prototype modeling exercise was developed to test the software in situations where the simulation grid and the soft-information grid are not aligned. In this exercise, the hypothetical geometric model consisted of four geologic units, essentially identical to those of the parallel-grid prototype; however, the strata were rotated 6 degrees in the x-z plane to represent dipping stratigraphic units (figure 15). The geometric block model used to discretized this dipping geology (pseudo-Lynx export file) was oriented parallel to the dip. The hard, drill-hole porosity data were not rotated, but were positioned as vertical drill holes located in the same three positions. The objective of this modeling exercise was to examine the performance of **GLINTMOD** in a situation requiring rotation of two completely independent grid systems.

All other modeling parameters and grid specification details were identical to those used in the parallel-grids example. The simulation consisted of 40 cells horizontally by 30 cells vertically; each cell was 100 feet by 50 feet. The geometric model consisted of cells double the dimensions of the simulation grid: 200 feet by 100 feet. The range of spatial correlation was set arbitrarily at 1,000 feet in the direction of maximum continuity, dipping 6 degrees from the horizontal toward the "east," with an anisotropy ratio of 0.15. The search radius was also set to dip 6 degrees from horizontal, with a range of 300 feet in the direction of maximum continuity and with a 0.15 anisotropy ratio. Because the two local coordinate systems are aligned in different directions, a portion of the "horizontal" simulation grid lies outside the physical volume represented by the constraining, "dipping" geometric model. Thus there is no soft information in the geometric model to constrain the simulation of certain material-property nodes. Although a number of alternative actions could have been defined in the case where **GLINTMOD** issued a call to the non-existent portion of the geometric model, the simplest solution, not simulating the node at all, was adopted for this prototype modeling exercise. This alternative might correspond to interpreting the upper limit of the dipping geometric model as surface topography and "missing values" in the corresponding portion of the simulation grid to represent "air."

Expected-value profiles representing ten separate stochastic simulations of this rotated-grid model are presented in figures 16 and 17. Figure 16 represents the base case in which only the hard drill-hole data are used to condition the simulations, and figure 17 represents the case where **GLINTMOD** accessed the geometric model at those grid nodes for which no conditioning data or previously simulated values were located within the search ellipse.

The general impression from these rotated-grid expected-value profiles is much the same as that gained from examination of the parallel-grid examples in figures 13 and 14. The model created using only the limited hard data values is relatively unstructured and does a poor job of reproducing a well-layered geologic system of contrasting porosity units. The conditioning drill-hole data are propagated away from the drill-hole locations in the proper manner; a definite dip of lenticular units toward the east is apparent in the figure. However, without the ability to constrain simulated nodes to the appropriate unit-specific expected material property or a measured value (an effect purposely produced by the short search distance), widely varying porosity values are generated at the same position in different simulations and then propagated into adjoining regions. When these lenses of widely varying porosity values are averaged across the suite of ten realizations to create the expected-value profile, they largely "cancel" each other and yield the "visually bland," average porosity cross section of figure 16.

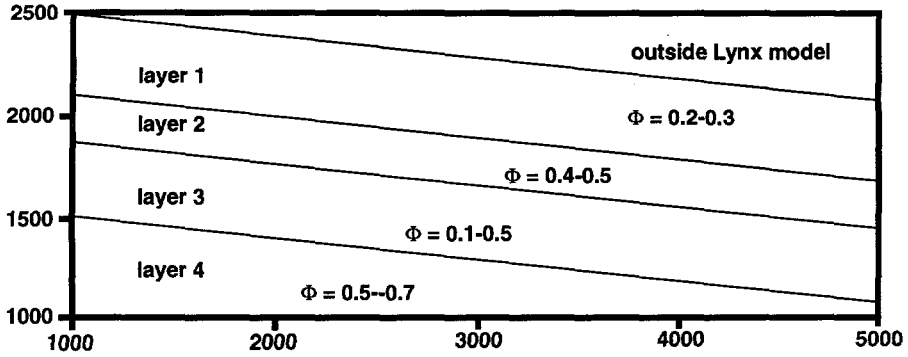


Figure 15. Hypothetical geometric model corresponding to the rotated-grid example. Four geologic units dip at 6 degrees to the east. Drill-hole locations shown in figure 8. Dimensions in feet, porosity as a fraction.

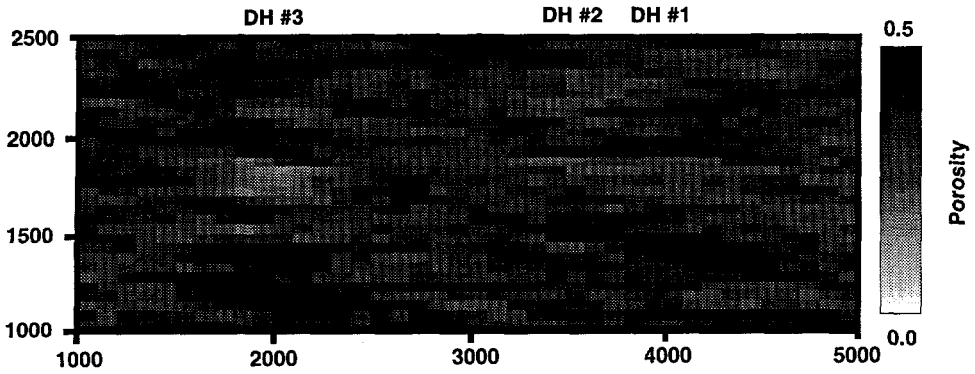


Figure 16. Expected-value profile for the rotated-grid model created using only the hard drill-hole data shown in figure 9. Direction of maximum spatial continuity has been rotated 6 degrees from horizontal toward the right. Dimensions in feet, porosity as a fraction.

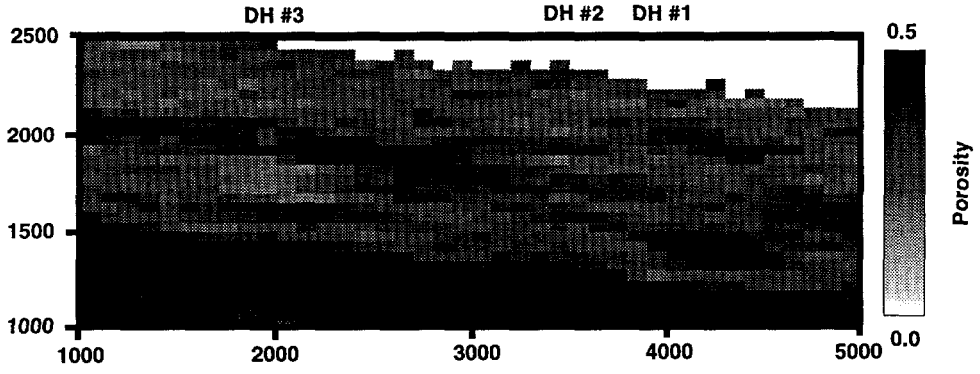


Figure 17. Expected-value profile for the rotated-grid model created using calls to **GLINTMOD** and the conceptual framework model of figure 15. Dimensions in feet, porosity as a fraction.

The **GLINTMOD**-generated expected-value profile of figure 17, however, faithfully reproduces the four-layer conceptual geometric model, complete with dipping strata. The white region in the upper right-hand corner of the model domain corresponds to the region that is undefined in the rotated Lynx-model export file. Although distorted by the 100x50-foot grid-cell discretization, this "topographic" surface between low-porosity "rock" and "sky" dips at the desired 6-degree angle, as do the various geologic layers. Note that even though the four geologic units are relatively well defined and distinct from one another, there is definite internal heterogeneity represented in the expected-value profile.

An interesting feature can be observed in geologic layer 2 near the right-hand side of the model. This region is strongly influenced by drill holes 1 and 2 (figure 8), as the two holes are close together geologically (well within the range of spatial correlation), and close enough physically that simulated grid nodes between the two holes are influenced by the presence of data from both holes within the search neighborhood. Layer 2 is relatively well developed on the left-hand side of the model. The layer extends as a discrete entity across the model, dipping at about 6 degrees, until it reaches the vicinity of the two drill holes. Here the layer breaks-up, almost appears to bifurcate and then reforms immediately adjacent to the right-hand boundary of the model. This "modeled confusion" is a consequence of inconsistency between the hard, drill-hole-based information and the soft (interpretive), geometric information.

The drill hole porosity data (figure 9) were fabricated to represent a horizontally layered geologic conceptual model. When these data were applied to the rotated-grid example case, the spatial "intercepts" of high- and low-porosity samples in the drill holes were not adjusted in keeping with the newly imposed concept of dipping layers (mostly for the sake of simplicity). Therefore, **GLINTMOD** calls at grid nodes located far from the hard conditioning data generated simulated values consistent with a dipping higher-porosity unit throughout the left-hand two-thirds of the model domain and also adjacent to the right-hand boundary. The laterally extensive unit to the west almost certainly was conditioned to measured porosity values in drill hole number 3. However, grid nodes within about 300 feet on either side of the two closely spaced drill holes were simulated in light of both an externally imposed dipping correlation structure and hard data indicating equivalent porosity values at essentially the same vertical positions, indicating a near-horizontal correlation structure. The "breakdown" of the conceptual model in the presence of actual hard data conflicting with that conceptual model is interpreted as a good indicator that soft information from the geometric model does not overly constrain the simulation process. **GLINTMOD** was developed to provide a numerical-simulation algorithm with guidance in cases where actual knowledge is effectively absent: a *model* is then substituted for *observations*. If sufficient observations are present, the data take priority and drive the simulation process.

Discussion

The prototype test modeling conducted using **GLINTMOD** indicates that the integration of soft, constraining information from a geometric geologic model is feasible, and that it results in simulated numerical material-property models that do reflect the underlying model conceptualized in the geologic-model output files. Turning the **GLINTMOD** interface "off" destroys this reproduction of the conceptual model and results in effectively unconstrained property models that would be judged geologically unreasonable in an actual modeling application.

Whether the prototype version of **GLINTMOD** is producing truly geologically real-

istic models in an actual application is another question. Most numerical modeling algorithms involve the specification of a large number of modeling parameters, some of which have immediate and obvious influences on the modeling output and others of which have more subtle or obscure influences. **GLINTMOD** is no exception. Early application of **GLINTMOD** to modeling of Yucca Mountain geology as part of ground-water travel time modeling exercises for 1995 (GWTT-95) has indicated that there are a number of these secondary, subtle-influence parameters that require modeling judgement. This section discusses a few of the more notable observations resulting from early application of **GLINTMOD** in attempting to generate material-property distributions for several cross sections of Yucca Mountain that are compatible with the general conceptual model of the site. Although resolution of these issues is beyond the scope of the prototype model development and testing exercise described in this report, these issues and their associated work-around or modifications suggest that additional development and testing activities with respect to **GLINTMOD** will be required.

When to Use Soft Information?

As described in the section on "Conceptual Development" on page 16, the prototype version of **GLINTMOD** calls for a material-property expectation from the Lynx geologic model only when no neighboring data or, more generally, inadequate data, are found (figure 6, ③). What, then, determines an "adequate" quantity of conditioning information? Originally, the intent was to call **GLINTMOD** only in the case where the number of hard data or previously simulated values located by the neighborhood search was equal to zero. The examples shown in this report were produced in this manner (figures 8 through 17).

In applying **GLINTMOD** to actual modeling of Yucca Mountain, it quickly became apparent that models produced using this approach appeared more random than suggested by the conceptual geologic model. Experimentation with the triggering flag for issuing the subroutine call to the Lynx export file (figure 6) suggested that setting the test value to eight (or fewer) conditioning points produced simulations that followed the input geologic model reasonably well. Examination of the behavior of this flag during execution of an individual simulation indicates that the Lynx framework model is accessed frequently during the earliest stages of the simulation process (figure 18). However, because the sequential path through the model domain is random, the soft information extracted from the framework model is spatially distributed throughout that domain and quickly constrains the entire simulated property model so that the number of external calls plateaus rapidly at approximately 5 to 10 percent of the total number of nodes in the domain.

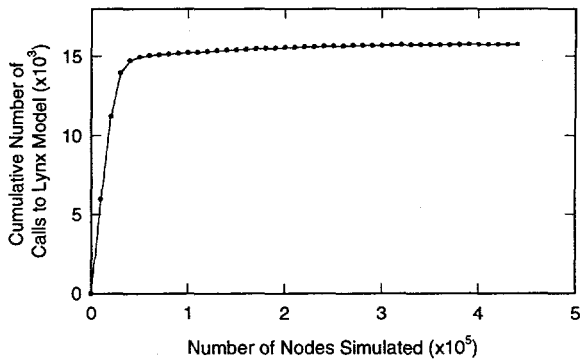


Figure 18. Graph showing the cumulative number of **GLINTMOD** calls to the framework Lynx model as a function of the total number of nodes simulated.

Use of Unit-Specific Variances

Another parametric value involved in the geologic reasonableness issue involves the variance of the *ccdf* from which the simulated value is generated (figure 5). The unmodified **SGSIM** sequential Gaussian simulation algorithm (Deutsch and Journel, 1992) uses two alternative values in this simulation process.

First, if sufficient data are located by the search algorithm, **SGSIM** computes both a kriged value and a kriging variance, which are taken as the expectation and variance (respectively) of the location-specific *cdf*. This follows from multivariate Gaussian spatial theory: the kriging variance is equal to zero at a data location, and it increases to a limit of the variogram sill value (= 1.0) at large distances from data. Second, if sufficient data (original plus previously simulated values) cannot be located by the spatial search, the algorithm assumes a standardized normal Gaussian distribution with $\sigma^2=1$ as the basis of this locally ***un***conditioned *cdf*. This, too, follows from spatial theory: if no additional information is available, the uncertainty associated with an unsampled location ought to approach the variability associated with the data taken without regard to spatial position. The difficulty with these rules is that additional information *is* being incorporated at unsampled locations through the **GLINTMOD** approach. However, that information is "soft" and consists of the rock unit inferred (but not known) to exist at an unsampled location.

Note that although *globally* the normal-score variance of the population is, in fact, equal to one, the variances of the subpopulations within individual geologic units are not. If the identity of the subpopulation (the rock unit) is known, the global variance of one probably overestimates the uncertainty at any specific (local) unsampled grid node. To understand this difference between global and local statistics, refer to the illustration of figure 4. In this conceptual diagram, the variance of the overall population (left-hand side distribution) is transformed so that it equals one for the standard normal distribution (right-hand side of the main figure). However, for neighborhoods similar to neighborhood number 1 (which could be considered to represent low-porosity welded rocks), or separately for those similar to neighborhood number 2 (high-porosity nonwelded rocks), the variances of values from these local neighborhoods are demonstrably less than one. When data are present (the case actually shown in figure 4), this variability is captured through the simple kriging variance. The problem remains to determine an appropriate variance for use when data are not present but when the likely rock-unit identity of the simulation node is known.

An obvious solution for determining the proper "width" of the probability density function is to tie the variance of the empirical *cdf* at all unsampled-but-soft-conditioned locations to the variance of the hard data used to estimate the unit-specific expected value. Thus, instead of looking up only an expectation to use in the case of "inadequate conditioning data found," the call to the Lynx export file would return both an expected value and an "expected" variance. This alternative is logically appealing, and it probably will be the alternative ultimately implemented, should development of **GLINTMOD** continue. Unit-specific variances will need to be estimated carefully, as estimates of the second moment of any distribution are significantly more sensitive to bias in the underlying data than is estimation of the first moment (mean or expectation). Outliers resulting from measurement error, misclassification of samples, local departures from true multiGaussian spatial behavior, or simply locally anomalous physical samples exert a disproportionate influence on the second moment through the squared-deviation term in the computational formula for a variance.

Because there has not yet been adequate time to develop and evaluate unit-specific variances for the prototype version of **GLINTMOD** and to implement the necessary software calls after identification of the issue, a reasonable heuristic device has been implemented for preliminary modeling. After progressively reducing the variance by

experimentation and comparing the resulting models in a semiquantitative manner, the within-unit variance was set to a value of 0.2 [relative to the standard normal (σ^2=1) distribution]. The rationale underlying this apparently arbitrary value is based on the coefficients of variation of porosity data available for Yucca Mountain. The coefficient of variation is a standardized measure of variability, and it is defined as the standard deviation divided by the mean. The observed coefficients of variation for porosity are essentially constant across many different geologic units at Yucca Mountain (see also table 2). Preliminary analyses indicate that the standard deviations of unit-specific porosity data were roughly 40–45 percent of the standard deviation of the combined porosity data. Therefore, if the standard deviation of the normal-score transformed data is one by definition and construction, the appropriate standard deviation of values for an individual geologic unit is approximately 0.40 to 0.45. Converting these values to variances (by squaring) suggests that the variance of the appropriate *ccdf* should be on the order of 0.16 to 0.20. We selected the larger value of 0.20, corresponding to a somewhat more variable probability function.

Resolution of Thin, Hydrologically Significant Units

A final, preliminary issue involved in modeling a real-world cross section of Yucca Mountain involves the coarseness of resolution available through the Lynx GMS. A Lynx model is currently limited to a maximum of 200 uniformly spaced grid cells in each principal direction. This absolute limit and the requirement that the cells be evenly spaced in a particular dimension means that it is difficult to achieve fine-scale resolution of thin, potentially very significant hydrologic units in a thick vertical sequence that also contains large vertical intervals of relatively similar material. Although the Lynx export file contains a complete listing of *all* model components contained within each cell, **GLINTMOD** by default simply reports the volumetrically dominant geologic component as "the" unit present within each cell.

Because the Tiva Canyon and Topopah Spring Tuffs at Yucca Mountain contain quite thin, but also laterally extensive low-porosity vitrophyre units near their upper and lower margins (units Tpcpv3, Tptrv1, and Tptpv3 in table 1) that are thought to influence groundwater flow and transport disproportionately to their volumetric representation in the Lynx model, a modification to the Lynx export-file query subroutine has been developed. Specifically, the look-up routine was modified so that if one of the vitrophyres composes more than approximately 10 percent of the volume of the Lynx grid cell corresponding to a simulation node requiring soft information, the subroutine returns "vitrophyre" as the rock type regardless of the actual, volumetrically dominant unit present in that grid block.

Conclusions

Prototype modeling using **GLINTMOD** has successfully demonstrated the viability of a proposed approach for integrating soft, interpretive information contained in geometric framework geologic models into stochastic numerical-properties modeling using a geostatistical simulation methodology. The method involves selection of a unit-specific expected value to control the distribution of potential values from which the material property of interest at an unsampled location is randomly selected. In the current Fortran implementation, **GLINTMOD** refers to the underlying geologic framework model only at those locations for which the standard simulation search algorithm is unable to locate either hard, measured conditioning data or previously simulated grid nodes within a user-specified search neighborhood. The concept is that if there are no relevant observations close to

the location to be simulated, the best prior estimate of the material property likely to exist at that point is the expected (mean) property of that overall rock type.

References

Armstrong, M., and Boufassa, A., 1988, Comparing the robustness of ordinary kriging and log-normal kriging: Outlier resistance: *Mathematical Geology*, v. 20, p. 447-457.

Barnard, R.W., and Dockery, H.A., eds., 1991, Technical summary of the performance assessment calculational exercises for 1990 (PACE-90), Volume 1: "Nominal configuration" hydrologic parameters and calculational results: *Sandia Report SAND90-2726*, Sandia National Laboratories, Albuquerque, N. Mex., 192 p.

Barnard, R.W., Wilson, M.L., Dockery, H.A., Gauthier, J.H., Kaplan, P.G., Eaton, R.R., Bingham, F.W., and Robey, T.H., 1992, TSPA-1991: An initial total-system performance assessment for Yucca Mountain: *Sandia Report SAND91-2795*, Sandia National Laboratories, Albuquerque, N. Mex., 376 p.

Bish, D.L., and Aronson, J.L., 1993, Paleogeothermal and paleohydrologic conditions in silicic tuff from Yucca Mountain, Nevada: *Clays and Clay Minerals*, v. 41, p. 148-161.

Boufassa, A., and Armstrong, M., 1989, Comparison between different kriging estimators: *Mathematical Geology*, v. 21, p. 331-345.

Buesch, D.C., Nelson, J.E., Dickerson, R.P., and Spengler, R.W., 1993, Development of 3-D lithostratigraphic and confidence models at Yucca Mountain, Nevada, in *High Level Radioactive Waste Management, Proceedings of the Fourth Annual International Conference*, April 26-30, 1993, American Nuclear Society, La Grange Park, Ill., p. 943-948.

Byers, F.M., Jr., Carr, W.J., Orkild, P.P., Quinlivan, W.D., and Sargent, K.A., 1976, Volcanic suites and related cauldrons of the Timber Mountain-Oasis Valley caldera complex, southern Nevada: *U. S. Geological Survey Professional Paper 919*, 70 p.

Carr, M.D., Waddell, S.J., Vick, G.S., Stock, J.M., Monsen, S.A., Harris, A.G., Cork, B.W., and Byers, F.M., Jr., 1986 Geology of drill hole UE25p#1: A test hole into pre-Tertiary rocks near Yucca Mountain, southern Nevada: *U. S. Geological Survey Open-File Report 86-175*, 87 p.

Christiansen, R.L., and Lipman, P.W., 1965, Geologic map of the Topopah Spring NW quadrangle, Nye County, Nevada: *U. S. Geological Survey Geological Quadrangle Map GQ-444.*

Clark, I., 1979, *Practical Geostatistics*, New York: Elsevier Applied Science Publishers, 129 p.

Deutsch, C., 1989, DECLUS: A Fortran 77 program for determining optimal spatial declustering weights: *Computers and Geosciences*, v. 15, p. 325-332.

Deutsch, C.V., and Journel, A.G., 1992, *GSLIB Geostatistical Software Library and User's Guide*, New York: Oxford University Press, 304 p.

Dudley, A.L., Peters, R.R., Gauthier, J.H., Wilson, M.L., Tierney, M.S., and Klavetter, E.A., 1988, Total system performance assessment code (TOSPAC) volume I: PhysicaL and mathematical bases: *Sandia Report SAND85-0002*, Sandia National Laboratories, Albuquerque, N. Mex., 204 p.

Englund, E.J., 1990, A variance of geostatisticians: *Mathematical Geology*, v. 22, p. 417-455.

Flint, L.E., Flint, A.L., Rautman, C.A., and Istok, J.D., 1996, Physical and hydrologic properties of rock outcrop samples at Yucca Mountain, Nevada: *U. S. Geological Survey Open-File Report 95-280.*

Fridrich, C J., Dudley, W.W., Jr., and Stuckless, J.S., 1994, Hydrogeologic analysis of the saturated-zone ground-water system, under Yucca Mountain, Nevada: *Journal of Hydrology*, v. 154, p. 133-168.

Geslin, J.K., and Moyer, T.C., 1995, Summary of lithologic logging of new and existing boreholes at Yucca Mountain, Nevada, March 1994 to June 1994: *U. S. Geological Survey Open-File Report 94-451*, 16 p.

Gomez-Hernandez, J.J., and Srivastava, R.M., 1990, ISIM3D: An ANSI-C three-dimensional multiple indicator conditional simulation program: *Computers and Geosciences*, v. 16, p. 395-440.

Isaaks, E.H., and Srivastava, R.M., 1989, *Applied Geostatistics*, New York: Oxford University Press, 561 p.

Istok, J.D., Rautman, C.A., Flint, L.E., and Flint, A.L., 1994, Spatial variability in hydrologic properties of a volcanic tuff: *Ground Water*, v. 32, p. 751-760.

Journel, A.G., 1983, Nonparametric estimation of spatial distributions: *Mathematical Geology*, v. 15, p. 445-468.

Journel, A.G., and Huijbregts, Ch.J., 1978, *Mining Geostatistics*, New York: Academic Press, 600 p.

Journel, A.G., and Alabert, F., 1989, Non-Gaussian data expansion in the earth sciences: *Terra Nova*, v. 1, p. 123-134.

Kaplan, P.G., 1993, Pre-waste-emplacement ground-water travel time sensitivity and uncertainty analyses for Yucca Mountain, Nevada: *Sandia Report SAND92-0461*, Sandia National Laboratories, Albuquerque, N. Mex., 121 p.

Lipman, P.W., and McKay, E.J., 1965, Geologic map of the Topopah Spring SW quadrangle, Nye County, Nevada: *U. S. Geological Survey Geological Quadrangle Map GQ-439.*

Lipman, P.W., Christiansen, R.L., and O'Connor, J.T., 1966, A compositionally zoned ash-flow sheet in southern Nevada: *U. S. Geological Survey Professional Paper 524-F*, p. F1-F47.

McKenna, S.A., and Rautman, C.A., 1995, Summary evaluation of Yucca Mountain surface transects with implications for downhole sampling, *Sandia Report SAND94-2038*, Sandia National Laboratories, Albuquerque, N. Mex., 45 p.

Ortiz, T.S., Williams, R.L. Nimick, F.B., Whittet, B.C., and South, D.L., 1985, a three-dimensional model of reference thermal/mechanical and hydrologic stratigraphy at Yucca Mountain, southern Nevada: *Sandia Report SAND84-1076*, Sandia National Laboratories, Albuquerque, N. Mex., 79 p.

Rautman, C.A., and Flint, A.L., Chornack, M.P., and McGraw, M., 1991, Microstratigraphic units and spatial correlation of hydrologic properties in tuff, Yucca Mountain, Nevada: *Geological Society of America Abstracts with Programs*, v. 23, no. 5, p. A119-120.

Rautman, C.A., and Flint, A.L., 1992, Deterministic geologic processes and stochastic modeling, in *High Level Radioactive Waste Management, Proceedings of the Third Annual International Conference*, April 12-16, 1992, American Nuclear Society, La Grange Park, Ill., v. 1617-1624.

Rautman, C.A., and Robey, T.H., 1993, Recent developments in stochastic modeling and upscaling of hydrologic properties in tuff: in *High Level Radioactive Waste Management, Proceedings of the Fourth Annual International Conference*, April 26-30, 1993, American Nuclear Society, La Grange Park, Ill., p. 1437-1445.

Rautman, C.A., Istok, J.D., Flint, L.E., Flint, A.L., and Chornack, M.P., 1993, Influence of deterministic geologic trends on spatial variability of hydrologic properties in volcanic tuff, in *High Level Radioactive Waste Management, Proceedings of the Fourth Annual International Conference*, April 26-30, 1993, American Nuclear Society, La Grange Park, Ill., P. 921-929.

Rautman, C.A., and Robey, T.H., 1994, Development of stochastic indicator models of lithology, Yucca Mountain, Nevada: in *High Level Radioactive Waste Management, Proceedings of the Fifth Annual International Conference*,

May 23-27, 1994, American Nuclear Society, La Grange Park, Ill., p. 2510-2519.

Rautman, C.A., Flint, L.E., Flint, A.L., and Istok, J.D., 1995, Physical and hydrologic properties of outcrop samples from a nonwelded to welded stuff transition, Yucca Mountain, Nevada: *U. S. Geological Survey Water Resources Investigations Report 95-4061*, 28 p.

Rautman, C.A., and Istok, J.D., (in press), Probabilistic assessment of ground-water contamination: 1. Geostatistical framework, accepted for publication in *Ground Water.*

Sawyer, D.A., Fleck, R.J., Lanphere, M.A., Warren, R.G., Broxton, D.E., and Hudson, M.R., 1994, Episodic caldera volcanism in the Miocene southwestern Nevada volcanic field: Revised stratigraphic framework, $^{40}Ar/^{39}Ar$ geochronology, and implications for magmatism and extension: *Geological Society of America Bulletin*, v. 106, p. 1304-1318.

Scott, R.B., Spengler, R.W., Diehl, S., Lappin, A.R., and Chornack, M P., 1983, Geologic character of tuffs in the unsaturated zone at Yucca Mountain, southern Nevada, in *Role of the unsaturated zone in radioactive and hazardous waste disposal*, Ann Arbor Science, Ann Arbor, Mich., p. 289-335.

Scott, R.B. and Bonk, J., 1984, Preliminary geologic map of Yucca Mountain, Nye County, Nevada, with geologic sections: *U. S. Geological Survey Open-File Report 84-494.*

Spengler, R.W., and Fox, K.F., Jr., 1989, Stratigraphic and structural framework of Yucca Mountain, Nevada: *Radioactive Waste Management and the Nuclear Fuel Cycle*, v. 13, p. 21-36.

Wilson, M.L., and Robey, T.H., 1994, The effect of stratigraphic uncertainty on repository performance, in *Proceedings of the International Topical Meeting on Nuclear and Hazardous Waste Management, Spectrum '94*, August 14-18, 1994, American Nuclear Society, La Grange Park, Ill., p. 1895-1900.

Wilson, M.L., Gauthier, J.H., Barnard, R.W., Barr, G.E., Dockery, H.A., Dunn, E., Eaton, R.R., Guerin, D.C., Lu, N., Martinez, M.J., Nilson, R., Rautman, C.A., Robey, T.H., Ross, B., Ryder, E.E., Schenker, A.R., Shannon, S.A., Skinner, L.H., Halsey, W.G., Gansemer, J., Lewis, L.C., Lamont, A.D., Triay, I.R., Meijer, A., and Morris, D.E., 1994, Total-system performance assessment for Yucca Mountain, SNL second iteration (TSPA-1993): *Sandia Report SAND93-2675*, Sandia National Laboratories, Albuquerque, N. Mex., 376 p.

YUCCA MOUNTAIN SITE CHARACTERIZATION PROJECT
SAND95-2080 - DISTRIBUTION LIST

1 D. A. Dreyfus (RW-1)
Director
OCRWM
US Department of Energy
1000 Independence Avenue SW
Washington, DC 20585

1 L. H. Barrett (RW-2)
Acting Deputy Director
OCRWM
US Department of Energy
1000 Independence Avenue SW
Washington, DC 20585

1 S. Rousso (RW-40)
Office of Storage and Transportation
OCRWM
US Department of Energy
1000 Independence Avenue SW
Washington, DC 20585

1 R. A. Milner (RW-30)
Office of Program Management
and Integration
OCRWM
US Department of Energy
1000 Independence Avenue SW
Washington, DC 20585

1 D. R. Elle, Director
Environmental Protection Division
DOE Nevada Field Office
US Department of Energy
P.O. Box 98518
Las Vegas, NV 89193-8518

1 T. Wood (RW-14)
Contract Management Division
OCRWM
US Department of Energy
1000 Independence Avenue SW
Washington, DC 20585

4 Victoria F. Reich, Librarian
Nuclear Waste Technical Review Board
1100 Wilson Blvd., Suite 910
Arlington, VA 22209

1 Wesley Barnes, Project Manager
Yucca Mountain Site Characterization
Office
US Department of Energy
P.O. Box 98608--MS 523
Las Vegas, NV 89193-8608

1 Director, Public Affairs Office
c/o Technical Information Resource Center
DOE Nevada Operations Office
US Department of Energy
P.O. Box 98518
Las Vegas, NV 89193-8518

8 Technical Information Officer
DOE Nevada Operations Office
US Department of Energy
P.O. Box 98518
Las Vegas, NV 89193-8518

1 J. R. Dyer, Deputy Project Manager
Yucca Mountain Site Characterization Office
US Department of Energy
P.O. Box 98608 -- MS 523
Las Vegas, NV 89193-88608

1 M. C. Brady
Laboratory Lead for YMP
M&O/Sandia National Laboratories
1261 Town Center Drive
Bldg. 4, Room 421A
Las Vegas, NV 89134

1 J. A. Canepa
Laboratory Lead for YMP
EES-13, Mail Stop J521
M&O/Los Alamos National Laboratory
P.O. Box 1663
Los Alamos, NM 87545

1 Repository Licensing & Quality
Assurance
Project Directorate
Division of Waste Management, MS T7J-9
US NRC
Washington, DC 20555

1 Senior Project Manager for Yucca
Mountain
Repository Project Branch
Division of Waste Management, MS T7J-9
US NRC
Washington, DC 20555

1 NRC Document Control Desk
Division of Waste Management, MS T7J-9
US NRC
Washington, DC 20555

1 Chad Glenn
NRC Site Representative
301 E Stewart Avenue, Room 203
Las Vegas, NV 89101

1 Center for Nuclear Waste
Regulatory Analyses
Southwest Research Institute
6220 Culebra Road
Drawer 28510
San Antonio, TX 78284

2 W. L. Clarke
Laboratory Lead for YMP
M&O/ Lawrence Livermore Nat'l Lab
P.O. Box 808 (L-51)
Livermore, CA 94550

1 Robert W. Craig
Acting Technical Project Officer/YMP
US Geological Survey
101 Convention Center Drive, Suite P-110
Las Vegas, NV 89109

1 J. S. Stuckless, Chief
Geologic Studies Program
MS 425
Yucca Mountain Project Branch
US Geological Survey
P.O. Box 25046
Denver, CO 80225

1 L. D. Foust
Technical Project Officer for YMP
TRW Environmental Safety Systems
101 Convention Center Drive
Suite P-110
Las Vegas, NV 89109

1 A. L. Flint
U. S. Geological Survey
MS 721
P. O. Box 327
Mercury, NV 89023

1 Robert L. Strickler
Vice President & General Manager
TRW Environmental Safety Systems, Inc.
2650 Park Tower Dr.
Vienna, VA 22180

1 Jim Krulik, Geology Manager
US Bureau of Reclamation
Code D-8322
P.O. Box 25007
Denver, CO 80225-0007

1 B. T. Brady
Records Specialist
US Geological Survey
MS 421
P.O. Box 25046
Denver, CO 80225

1 M. D. Voegele
Deputy of Technical Operations
M&O/SAIC
101 Convention Center Drive
Suite P-110
Las Vegas, NV 89109

2 A. T. Tamura
Science and Technology Division
OSTI
US Department of Energy
P.O. Box 62
Oak Ridge, TN 37831

1 P. J. Weeden, Acting Director
Nuclear Radiation Assessment Div.
US EPA
Environmental Monitoring Sys. Lab
P.O. Box 93478
Las Vegas, NV 89193-3478

1 John Fordham, Deputy Director
Water Resources Center
Desert Research Institute
P.O. Box 60220
Reno, NV 89506

1 The Honorable Jim Regan
Chairman
Churchill County Board of
Commissioners
10 W. Williams Avenue
Fallon, NV 89406

1 R. R. Loux
Executive Director
Agency for Nuclear Projects
State of Nevada
Evergreen Center, Suite 252
1802 N. Carson Street
Carson City, NV 89710

1 Brad R. Mettam
County Yucca Mountain
Repository Assessment Office
P. O. Drawer L
Independence, CA 93526

1 Vernon E. Poe
Office of Nuclear Projects
Mineral County
P.O. Box 1600
Hawthorne, NV 89415

1 Les W. Bradshaw
Program Manager
Nye County Nuclear Waste Repository
Project Office
P.O. Box 1767
Tonopah, NV 89049

1 Florindo Mariani
White Pine County Coordinator
P. O. Box 135
Ely, NV 89301

1 Tammy Manzini
Lander County Yucca Mountain
Information Officer
P.O. Box 10
Austin, NV 89310

1 Jason Pitts
Lincoln County Nuclear Waste
Program Manager
P. O. Box 158
Pioche, NV 89043

1 Dennis Bechtel, Coordinator
Nuclear Waste Division
Clark County Dept. of Comprehensive
Planning
P.O. Box 55171
Las Vegas, NV 89155-1751

1 Juanita D. Hoffman
Nuclear Waste Repository
Oversight Program
Esmeralda County
P.O. Box 490
Goldfield, NV 89013

1 Sandy Green
Yucca Mountain Information Office
Eureka County
P.O. Box 714
Eureka, NV 89316

1 Economic Development Dept.
City of Las Vegas
400 E. Stewart Avenue
Las Vegas, NV 89101

1 Community Planning & Development
City of North Las Vegas
P.O. Box 4086
North Las Vegas, NV 89030

2 Librarian
YMP Research & Study Center
101 Convention Center Drive, Suite P-110
Las Vegas, NV 89109

1 Library Acquisitions
Argonne National Laboratory
Building 203, Room CE-111
9700 S. Cass Avenue
Argonne, IL 60439

1 Glenn Van Roekel
Manager, City of Caliente
P.O. Box 158
Caliente, NV 89008

1 Gudmundur S. Bodvarsson
Head, Nuclear Waste Department
Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 50E
Berkeley, CA 94720

1 Steve Hanauer (RW-2)
OCRWM
U. S. Department of Energy
1000 Independence Ave.
Washington, DC 20585

1 Marc V. Cromer
Newmont Gold Company
1700 Lincoln St.
Denver, CO 80203

1 Robert W. Clayton
M&O/WCFS
101 Convention Center Drive/MS423
Las Vegas, NV 89109

1 Richard C. Quitmeyer
M&O/WCFS
101 Convention Center Drive/MS423
Las Vegas, NV 89109

1 Mark C. Tynan
DOE/YMPSCO
101 Convention Center Drive/MS523/HL
Las Vegas, NV 89109

2 1330 B. Pierson, 6811
100/1232222/SAND95-2080/NA

20 1330 WMT Library, 6752
1 9018 Central Technical Files, 8523-2
5 0899 Technical Library, 4414
2 0619 Review and Approval Desk,
12630, For DOE/OSTI
5 1324 C. A. Rautman, 6115
2 1324 S. A. McKenna, 6115
1 1324 W. P. Zelinski, 6115
1 1326 S. J. Altman, 6851
1 1326 B. W. Arnold, 6851