RAW MATERIALS FOR PORTLAND CEMENT: APPLICATIONS OF CONDITIONAL SIMULATION OF COREGIONALIZATION

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF APPLIED EARTH SCIENCES

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

Gordon Ray Luster

September 1985

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

A. G. Journel

(Principal Advisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John W. Harbaugh

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

(Statistics)

Approved for the University Committee on Graduate Studies:

Dean of Graduate Studies & Research

PREFACE

Like a few others before me, I must admit that my dissertation did not turn out exactly as planned. My original idea was to write a "definitive" text on cement raw materials, discussing applications of geostatistics, mathematical programming, statistics and data analysis, process engineering, chemistry, geology, and mining engineering, roughly in decreasing order of emphasis. However, after I had written over two hundred pages just on the subject of "conditional simulation of coregionalization" -- a subset of the field of geostatistics -- I realized that to cover my whole list of topics in comparable detail, I would have to spend perhaps another three or four years at Stanford and write perhaps another two thousand pages. Lacking sufficient commitment to undertake such a task at graduate-student wages, I decided to restrict my dissertation research to applications of conditional simulation of coregionalization and simply devote my subsequent career to tackling the remaining issues. Fortunately, conditional simulation has a great deal of untapped potential for applications in the cement industry, as I hope the following pages will demonstrate.

Financial support for this dissertation was provided largely by a grant from Exxon Minerals Company. Additional support was furnished by the Stanford Geostatistics Program and by the Dean and Dorothea McGee Fund of the Stanford School of Earth Sciences. The following individuals also provided much direct assistance in the completion of

this project: my patient and hard-working adviser, Andre Journel, whose influence can be found on nearly every page; the members of my proposal, reading, and defense committees, and particularly Mike Davis, who pointed out a multitude of inadequacies in Chapter 3; and Ron Gebhardt and Charles Yost of Lehigh Portland Cement, who made numerous suggestions for the improvement of Chapter 2 and who arranged for release of the excellent data set used in Section 4.2.

This dissertation is dedicated in memory of my father and of Mrs. N. H. Davison, who gave me my first "science" book as soon as I was old enough to read the words.

ABSTRACT

Modern cement plants commonly rely on homogenization and proportioning facilities to reduce the natural variability of incoming raw materials and to blend different materials into plant feeds satisfying strict chemical specifications. To determine appropriate designs and capacities for such facilities, plant designers must know the statistical behavior of each raw material arriving at the plant. However, if the quarries have not been opened yet, or if historical data on raw-material quality have not been kept, or if the materials to be quarried in the future differ appreciably from those used in the past, then only data extracted from exploratory drilling of the materials in the ground will be available.

Fortunately, useful information can be obtained by combining conditional simulations of the deposits to be mined with simulations of likely mining procedures. A conditional simulation of a mineral deposit consists of a large set of simulated raw-material analyses distributed among the nodes of a fine two- or three-dimensional grid covering the region of the deposit that is to be mined. These simulated data should possess all statistical properties, including spatial properties, that are suggested by the available data from the real deposit; furthermore, they should equal the data obtained from the real deposit at actual data locations. The output of a mining simulation applied to a simulated deposit is a time-series simulation of material compositions that mimics

the statistical behavior of the real materials to be delivered to the plant in the future.

Improved methods have been developed for creating statistically and geologically realistic conditional simulations of multivariate, or coregionalized, data such as complete rock analyses. These methods have been used to simulate chemical analyses of two limestone deposits currently being mined for cement manufacture. Simulations of alternative mining and homogenization procedures applied to one of the simulated deposits illustrate how conditional simulations can be employed to select the best mining and homogenization procedures before mining of the deposit has actually begun.

CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter I

INTRODUCTION

This is a dissertation in Applied Earth Sciences. For any work to be "applied", it must be dedicated to the solution of a practical problem. The problem to be addressed here is how to characterize and deal with the variability of raw materials used for the manufacture of portland cement. This is not a new problem, but it is becoming more serious:

> The general character of the formation with which one is dealing must be determined in order to formulate proper plans for prospecting. It is positively a mistake, although not as serious, to drill 50 prospect holes in a property where marked uniformity prevails, if 10 holes would be sufficient, as to drill only 10 in a region of variable rocks where 50 are necessary to form an adequate picture of the situation. The writer has observed mistakes of both kinds. [B. L. Miller, 1934, p. 30]

> The complexities of both process and equipment will dictate an ever-increasing demand for highly qualified personnel and staff. . . . The geologist can no longer simply determine the quantity and grade of basic oxides and alkalies. The analysis must include the determination of variations . . . [H. M. Garrett, 1976, p. 16]

> Therefore, there is clearly a basic requirement for a two tier geological survey to include:
>
> 1. A preliminary investigation which establishes the particular deposit to be utilised and subsequently quarried.
> 2. . . . A more detailed study . . . which would provide the requisite process design data and information.
>
> Present practice within the cement industry tends to concentrate all the available effort and resources on the preliminary investigation . . . Point 2 above, although regarded as important, cannot be justified on a cost basis by the industry. . . . Such costs can be justified by considering, that as the industry is essentially moving more towards the utilisation of progressively lower grade reserves, there is a corresponding increase in the probability of a cement works being designed, built and onstream that is incapable of producing cement to the specification originally

> envisaged, due essentially to the inadequacies of the raw material . . . assessment. In the author's opinion (in the context of the totally inadequate methodology adopted within the industry for orebody quality characterization) such an event will take place, and although undesirable may well have to occur before the necessary shift of cost emphasis from quantity to quality considerations is forthcoming. [C. G. Schofield, 1980, p. 200]

B. L. Miller, a pioneer in the application of geological science to the study of cement raw materials, recognized fifty years ago that those of us who make a living assessing the suitability of raw materials for cement manufacture sometimes do not adequately take the variability of those materials into consideration in our assessments. Our shortcomings in this regard have become more serious as the technology of cement making has become more sophisticated and more dependent on uniform raw-material feeds. The second and third quotes above, by contemporary cement process-design engineers, make it clear that these shortcomings have been noticed. Nevertheless, cement managers and plant designers rarely ask geologists for more than an accounting of the total reserves and an impression of the <u>average</u> characteristics of their raw materials. Possibly no further information is sought simply because no further information is believed to be obtainable.

This dissertation demonstrates that useful information on raw-material variability <u>can</u> be obtained, using the same types of data (mainly drilling data) that cement geologists have traditionally used for determining only the average compositions of their raw materials. Such information can be particularly valuable if we wish to construct a new cement plant on a "greenfield" site, using new, previously unquarried raw materials. Plant designers must decide whether to construct a preblending system to smooth out the variability of those

raw materials as they are delivered to the new plant; then they may need to decide what kind of system it should be and how big it should be. Quantitative information for making these decisions can be obtained with the methods described in the following chapters.

Chapter 2 provides a review of cement-making processes and the types of information that are important in the characterization of cement raw materials. Chapter 3 contains a summary of basic "geostatistical" concepts, which are the basis for characterizing raw-material variability, and provides a detailed exposition of the methods and practices of "conditional simulation of coregionalization", a geostatistical technique particularly well suited to cement problems. Chapter 4 contains two cement-related case studies illustrating the methods of conditional simulation and some of its applications in mine planning and plant design. Chapter 5 offers some final comments about the costs of simulation studies, the types and amounts of data that are needed, and the criteria that should be used to decide when a simulation approach is warranted.

Chapter II

PORTLAND CEMENT AND CEMENT RAW MATERIALS

This chapter briefly summarizes some fundamentals of cement composition and manufacture, and describes the methods used for evaluation, selection, and quality control of cement raw materials. For readers wishing to learn more details about these subjects, the following sources are recommended. Bye (1983) provides a short, up-to-date introduction to all phases of cement technology, written from a materials-science viewpoint. More detailed reviews of individual topics, including extensive bibliographies, are provided in a large volume edited by Ghosh (1983), and by Duda (1977). Schofield (1980) discusses the state of the art in homogenization, proportioning, and process-control systems for cement raw materials, and includes a wealth of useful references. Lea (1971) provides an exhaustive account of cement and concrete chemistry. Two useful European journals, usually providing several articles of raw-materials interest each year, are World Cement (U.K.) and Zement-Kalk-Gips (Germany). Occasionally something of interest also appears in the less technical American trade journals Rock Products and Pit and Quarry, and in the proceedings of the IEEE Cement Industry Technical Conference.

2.1 PORTLAND CEMENT: DEFINITIONS AND SPECIFICATIONS

> Portland cement -- a hydraulic cement produced by pulverizing clinker consisting essentially of hydraulic calcium silicates, usually containing one or more of the forms of calcium sulfate as an interground addition. [ASTM, 1984, p. 156]
>
> Portland cement may be defined . . . as a product obtained by intimately mixing together calcareous and argillaceous, or other silica, alumina, and iron-bearing materials, burning them at a clinkering temperature, and grinding the resulting clinker. [Lea, 1971, p. 13]

The term "portland cement" is a generic name for the most common of several varieties of hydraulic cement used to manufacture concrete for construction purposes. It is usually produced by blending together calcareous and argillaceous raw materials, such as limestone and clay, then grinding the blended materials to a powder, then "burning" the ground material in a kiln at a temperature sufficient to recombine the nonvolatile constituents into a mixture of several new compounds, and finally grinding the resulting clinker together with a small amount of gypsum. (The gypsum is added primarily to control the rate at which the concrete hardens.) The product is a gray powder that reacts with water to form a hard crystalline cementing material. A typical portland cement might have the following chemical composition (Lea, 1971, p. 16):

SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO	SO_3	Other	TOTAL
22.0	5.5	3.0	64.1	1.4	2.1	1.9	100.0

Mineralogically, a good-quality portland cement consists mainly of the compounds[1] $(CaO)_3 \cdot SiO_2$ or C_3S, $(CaO)_2 \cdot SiO_2$ or C_2S, $(CaO)_3 \cdot Al_2O_3$ or

---

[1] The following abbreviated chemical notation, in common use in the cement industry, will be used in this chapter: $S=SiO_2$, $A=Al_2O_3$, $F=Fe_2O_3$, $C=CaO$, $\underline{S}=SO_3$, $M=MgO$, $K=K_2O$, $N=Na_2O$, $H=H_2O$. (An overbar, rather than an underscore, is normally used in the abbreviation for SO_3. Overbars are not available in the character set used to print this dissertation.)

C_3A, a ferrite similar to $(CaO)_4 \cdot Al_2O_3 \cdot Fe_2O_3$ or C_4AF, and either gypsum ($CaO \cdot SO_3 \cdot (H_2O)_2$, or $C\underline{S}H_2$) or a mixture of gypsum and anhydrite ($CaO \cdot SO_3$, or $C\underline{S}$). The major impurities usually are MgO, various compounds of Na_2O, K_2O, and SO_3, and some free CaO. The clinker compounds typically occur as poorly crystallized solid solutions that may deviate considerably from the ideal formulas above. They are generally interspersed with glassy material formed during the quenching of the hot clinker. For this and other technical reasons, it is difficult to obtain an accurate quantitative mineralogical (modal) analysis of a cement. Hence it is common practice to represent the composition of a cement in terms of a set of "potential" (normative) compound compositions calculated from a complete chemical analysis of the cement, in the same manner that normative mineral compositions are obtained for igneous rocks. The "Bogue equations" (Bogue, 1955, Chapter 10), derived below, are the most commonly used formulas for calculation of normative mineral compositions of cements and clinkers. These equations are listed in Table 1, in the slightly modified form given in ASTM Specification C 150 (ASTM, 1984, pp. 158-159).

Table 1 lists several <u>chemical</u> specifications that are commonly imposed on the two most common types of portland cement used in the United States. Additional <u>physical</u> specifications on the product will not be discussed here, as the physical properties of cements are determined primarily by the manufacturing process and are thus unimportant in a discussion of raw materials. Some of the specifications listed in Table 1 are promulgated by ASTM (in the United States), some are more commonly imposed by individual manufacturers, and

TABLE 1

Some common cement specifications and other functions used to describe the compositions of portland cements.

SPECIFICATION	FORMULA (weight % in cement)	Typical limits for: Type I	Type II
Bogue C_3S	4.071 C - 7.6 S - 6.718 A - 1.43 F - 2.852 $\underline{S}$	≃ 55	≃ 50
Bogue C_2S	2.867 S - 0.7544 C_3S	------	------
Bogue C_3A	2.65 A - 1.692 F	------	* < 8
Bogue C_4AF	3.043 F	------	(see F)
SiO_2	S	------	* >20.0
Al_2O_3	A	------	* < 6.0
Fe_2O_3	F	------	* < 6.0
MgO	M	* < 6.0	* < 6.0
SO_3	$\underline{S}$	* < 3.5	* < 3.0
Alkalies	N + 0.658 K	< 0.60	< 0.60
SM	S / (A + F)	2.3-2.8	2.6-3.1
AM	A / F	1.3-2.3	1.3-1.7
LSF	C / (2.8 S + 1.18 A + 0.65 F)	0.91-0.96	0.88-0.91

NOTES: Only specifications that may be important in judging the suitability of raw materials are included here.

Type I portland cement -- general-purpose cement
Type II portland cement -- cement with moderate sulfate resistance or moderate heat of hydration

Specifications marked by "*" are ASTM Standard Chemical Requirements, applicable to Type I or Type II cements produced in the USA, from ASTM Specification C 150. Other specifications are more commonly imposed by cement producers or customers, and may vary greatly among plants.

others may be required by government agencies or customers. The rationales underlying these specifications are briefly summarized in the following paragraphs.

The Bogue equations. To obtain a normative mineral composition of a cement from an analysis of SiO_2, Al_2O_3, Fe_2O_3, CaO, and SO_3, we simply assume that these five chemical components are allocated among the following five phases: C_3S, C_2S, C_3A, C_4AF, and $C\underline{S}$.[2] Thus the total Fe_2O_3 reported in the analysis consists entirely of the Fe_2O_3 that makes up 32.86% by weight of stoichiometric C_4AF, as this is the only phase containing Fe_2O_3; the Al_2O_3 in the analysis contributes 37.73% of the C_3A and 20.98% of the C_4AF; and so on for the other oxides. If we represent the normative composition of the clinker by the vector $N'=[C_3S, C_2S, C_3A, C_4AF, C\underline{S}]$ and the chemical composition by $X'=[S, A, F, C, \underline{S}]$, we can easily set up a matrix M to transform a normative analysis into a chemical analysis, i.e.:

$$X = M\,N$$

$$\begin{bmatrix} S \\ A \\ F \\ C \\ \underline{S} \end{bmatrix} = \begin{bmatrix} 0.2631 & 0.3488 & 0 & 0 & 0 \\ 0 & 0 & 0.3773 & 0.2098 & 0 \\ 0 & 0 & 0 & 0.3286 & 0 \\ 0.7369 & 0.6512 & 0.6227 & 0.4616 & 0.4119 \\ 0 & 0 & 0 & 0 & 0.5881 \end{bmatrix} \begin{bmatrix} C_3S \\ C_2S \\ C_3A \\ C_4AF \\ C\underline{S} \end{bmatrix}$$

Each column of M adds to 1.0, so that:

$$S + A + F + C + \underline{S} = C_3S + C_2S + C_3A + C_4AF + C\underline{S}$$

---

[2] Bogue (1955) adjusts his total CaO downward to account for $C\underline{S}$ before beginning his calculations, rather than incorporating $\underline{S}$ and $C\underline{S}$ directly into his equations.

Inverting M, we obtain a matrix M^{-1} of Bogue coefficients, which are incorporated (with more or less roundoff than provided below) into ASTM Standard C 150 (ASTM, 1984, p. 158):[3]

$$N = M^{-1} X$$

$$\begin{bmatrix} C_3S \\ C_2S \\ C_3A \\ C_4AF \\ C\underline{S} \end{bmatrix} = \begin{bmatrix} -7.60 & -6.72 & -1.43 & 4.07 & -2.85 \\ 8.60 & 5.07 & 1.08 & -3.07 & 2.15 \\ 0 & 2.65 & -1.69 & 0 & 0 \\ 0 & 0 & 3.04 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.70 \end{bmatrix} \begin{bmatrix} S \\ A \\ F \\ C \\ \underline{S} \end{bmatrix}$$

If we multiply the "typical" chemical composition on page 6 by M^{-1}, we obtain the following "typical" Bogue compositions for a portland cement: C_3S=46.5, C_2S=28.0, C_3A=9.5, C_4AF=9.1, $C\underline{S}$=3.6. This approximates a Type I cement according to Table 1. $C\underline{S}$ normally is not reported; sulfur actually occurs in several compounds ranging from $C\underline{S}$ to $C\underline{S}H_2$ in a typical cement.

C_3S is the essential ingredient needed to provide high early strength in a concrete. C_3A and C_4AF also appear to contribute to early strength, whereas C_2S appears to provide slower gains in strength. Although cements very high in silicates would provide high strength, they would be prohibitively expensive to manufacture because the mixes would have to be burned at very high temperatures to form the silicates. Nevertheless, the contents of C_3A and C_4AF must be limited in some applications. C_3A produces a high heat of hydration at early ages that is undesirable in massive concrete structures, and concretes made with high-C_3A cements have a low resistance to attack by sulfates in their

---

[3] A different set of equations, based on different assumptions about the phase composition, is used in Standard C 150 if the weight ratio A/F is less than 0.64.

environment. C_4AF and other iron-bearing phases are dark-colored and thus must be severely limited in white cements used for some architectural applications.

Ratios (moduli). Ratios between linear combinations of oxides have become popular for clinker quality control. Although a bewildering array of ratios (with conflicting formulas and names) has been proposed, the three most commonly used ratios appear to be:

$$\text{Silica modulus} = SM = S / (A + F)$$

$$\text{Alumina modulus} = AM = A / F$$

$$\text{Lime Saturation Factor} = LSF = C / (2.8\ S + 1.18\ A + 0.65\ F)$$

Mixes with low SM and LSF are more "burnable"; i.e., clinker compounds tend to form at lower temperatures, requiring less fuel consumption. A mix with higher SM will produce a clinker having a relatively higher content of silicates (C_3S and C_2S) versus aluminate and aluminoferrite (C_3A and C_4AF). Mixes with high AM will produce clinker with a larger proportion of C_3A to C_4AF, and mixes with high LSF will yield higher ratios of C_3S to C_2S. LSF was originally derived so that LSF=1.0 represented "the maximum lime content that can be present without free lime appearing at the clinkering temperature in equilibrium with the liquid present" in the quaternary system S-A-F-C (Lea, 1971, p. 164).

Impurities and Other Constraints. Maximum levels of MgO, SO_3, and in some cases alkalies (commonly expressed as soda equivalent, $Na_2O + 0.658\ K_2O$) are imposed on cements, and specific types of cement (particularly Type II -- see Table 1) are tightly constrained in their contents of other oxides and potential compounds. Undesirably high levels of MgO and alkalies are particularly common in limestones and

shales, respectively, and are among the most common reasons why specific limestone and shale deposits are rejected in the early stages of exploration for cement raw materials. Particles of periclase (MgO) occurring in a cement may undergo a slow hydration leading to long-term unsoundness of the concrete, and alkalies in the cement may react with some types of siliceous aggregates to produce long-term expansion and cracking of the concrete.

2.2 CEMENT MANUFACTURING

The cement manufacturing process (Figure 1) can be divided into five successive steps: (1) mining and crushing of the individual raw materials; (2) homogenization and proportioning of the raw materials; (3) grinding and homogenization of the raw mix; (4) dehydration, calcination, and sintering of the raw mix in a kiln system, followed by rapid cooling of the resulting clinker; and (5) grinding of the clinker, usually with addition of gypsum or a gypsum-anhydrite blend, and possibly with addition of other special-purpose additives. The means by which these steps are carried out may vary greatly among different plants, and in many plants there is some overlap among the steps. The underlying principles and major operations involved in each step are described briefly in the following sections. This description is rather nontraditional, strongly emphasizing the second step of the process, where the methods to be developed in this dissertation can be most profitably applied.

STEPS

(see page 12)

(1)

(2)

(3)

(4)

(5)

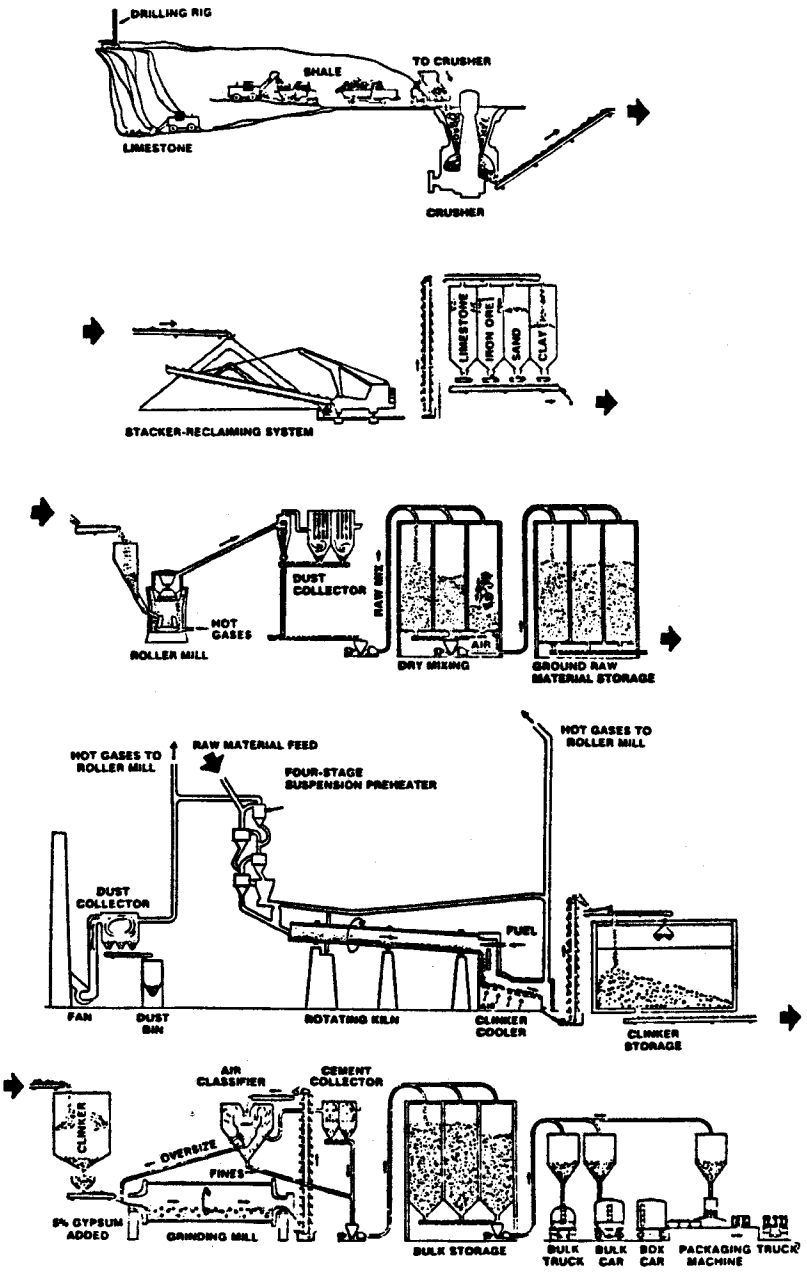


Figure 1: Process flow diagram depicting the organization of a preheater cement plant. [Source: Portland Cement Association]

2.2.1 Mining and Crushing

Cement plants are usually located in one of two areas: near the primary source of raw materials, or near the primary market for the cement. Total transportation costs for raw materials and product are usually lower if the plant is located nearer the raw materials. This alternative avoids the costs of transporting moisture, combined water, carbon dioxide, and components eventually discarded as waste dust (explained below) -- all of which are present in most of the raw materials, but not in the product.

Most cement plants are supplied by at least two nearby mining operations: a large quarry for limestone production, and a smaller pit for shale or clay. Because of the low unit values of the raw materials and product, expensive mining methods must be avoided if possible. Raw materials consumed in relatively small quantities (perhaps sand, iron ore, kaolin or bauxite, and gypsum or anhydrite) and the fuel used in the kiln (usually coal) are typically purchased from outside suppliers and transported to the plant.

All raw materials are crushed to a top size of a few centimeters prior to homogenization and proportioning. Newer plants usually have a primary crusher inside the limestone quarry, with a long conveyor belt leading to a secondary crusher near the plant.

2.2.2 Homogenization and Proportioning

"Homogenization" of raw materials is the reduction of variations in raw-material characteristics by mechanical means. Homogenization has become more important in recent years with the introduction of energy-

saving burning methods (discussed in Section 2.2.4), which tend to be sensitive to fluctuations in raw-material composition. "Proportioning"[4] of raw materials is the combining together of different, preferably homogenized, raw materials in proportions designed to satisfy a set of quality specifications. Proportioning has always been a fundamental concern in cement quality control, mine planning, and raw-materials exploration.

2.2.2.1 Homogenization

If we follow a particular mined or imported raw material "downstream" from its source to its final disposition as part of the finished product, we can find several locations along the way where homogenization (or its opposite -- segregation) can occur. Homogenization can begin in the quarry, where a given material exposed in several active quarry faces, or throughout a single large face, may be loaded together into the same truck (rarely), or loaded into different trucks but delivered to the same pre-crusher stockpile (commonly), thus smoothing out the effects of the local in-situ variability in the material. A simulated example of the effects of different loading schemes (but on a larger scale than truck loads) is provided in Section 4.2.5.

Once the material has entered the plant's process stream (beginning with the pre-crusher stockpile, if there is one), any location in the stream where large amounts of material are accumulated and mixed

---

[4] Some authorities (e.g., Schofield, 1980) use the term "blending" instead of "proportioning". Still others use "blending" to mean "homogenization", as in a kitchen "blendor".

together (i.e., "pools" in the stream) can, in principle, be used for homogenization. These are facilities in which the material can be said to have a distribution of "residence times". The capabilities of such facilities have been analyzed in detail by Schofield (1980). The facilities may be specifically designed for homogenization (e.g., stacker-reclaimer systems, stirred tanks, and fluidized silos), although storage stockpiles, surge bins, storage silos, crushers, mills, and even kiln systems may also achieve some homogenization. Of course, systems not specifically designed to homogenize may segregate instead (e.g., a stockpile in which coarse materials collect at the edges), or do nothing (e.g., a bin or silo through which material moves by "plug flow", as though in a pipeline).

Stacker-reclaimer systems (stockpile prehomogenizers) are nearly a necessity for cement plants with sensitive preheater or precalciner systems and chemically variable raw materials or fuel (i.e., most new plants). Several competing commercial designs are available (described conceptually by Schofield, 1980, Chapters 3 through 6; mechanically in Wohlbier, 1977; and concisely by Colijn, 1980). All but one of the systems operate on a "batch" principle.

A very simple type of batch system is illustrated in Figure 2(a). The basic idea is simply to add increments of material to the pile in series, then withdraw them in parallel. As crushed material is conveyed to the system it is dumped onto a long, narrow stockpile by a traveling stacker, which moves back and forth along the length of the pile many times (perhaps about 100 times each way) while the pile is being built, spreading a thin layer of material across the pile during each pass.

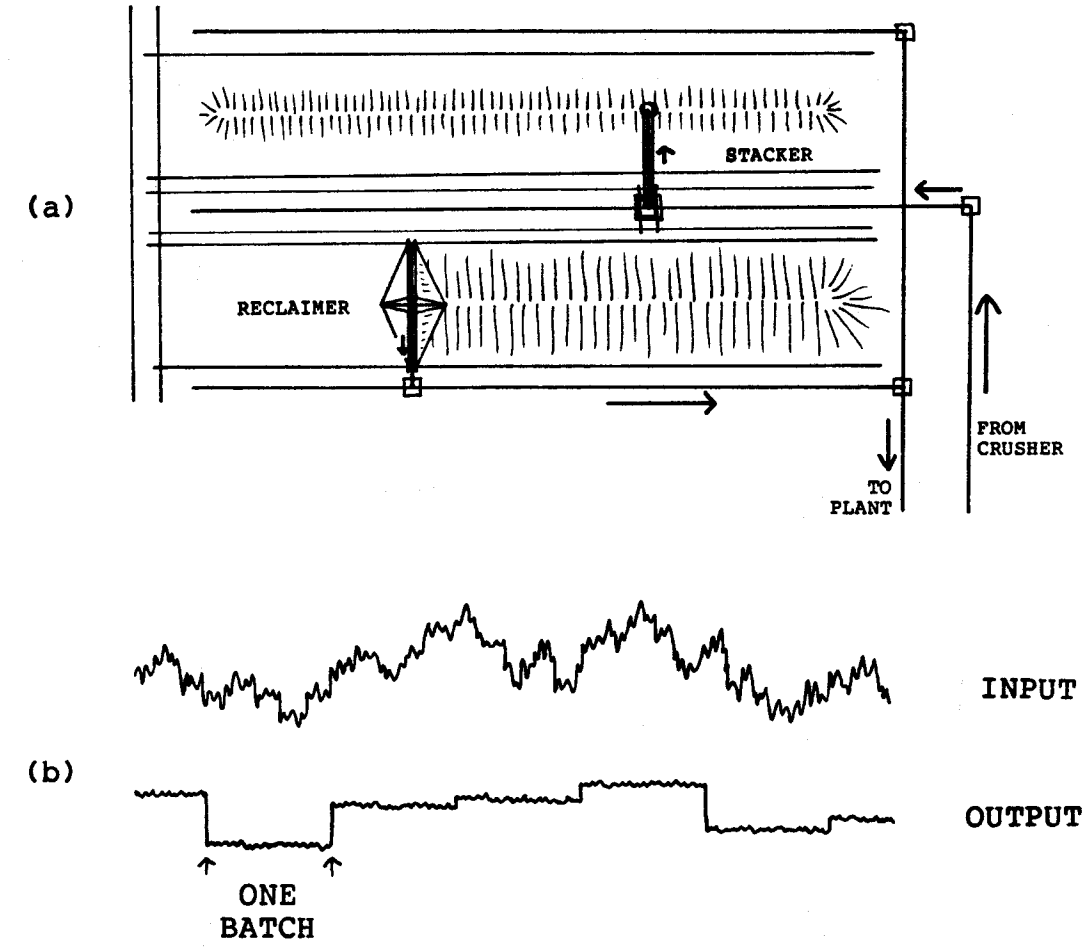


Figure 2: Homogenization of a raw-material input stream using a batchwise linear stockpile prehomogenizer.
(a) Plan view of a linear stacker-reclaimer system.
(b) Batchwise homogenization of an input time series.

When this pile is finished, the stacker moves to a second pile and builds it up in the same way, while a reclaimer goes to work on the original pile. The reclaimer begins at one end of the linear pile and works its way gradually to the other end, continuously reclaiming a complete cross section through the stacked material (about 200 layers of

material stacked onto the pile at 200 different times) from the triangular face of the pile. The reclaimer may simply be a bucket-wheel excavator, or preferably some device capable of shaving a slice of material off the whole face of the pile simultaneously. Any variability among the layers of material collected from the input stream during the time the pile was being built is thus averaged out by reclaiming all layers simultaneously. The composition of a slice reclaimed from each of 100 to 200 layers deposited at regular[5] intervals during the time in which the pile was built constitutes a very good approximation of the average composition of the whole pile.[6] The result is depicted in Figure 2(b), which compares the input and output time series of a batchwise stacking system. The large variability of the input stream deposited onto each successive pile is reduced to small random fluctuations about the mean composition of each pile. Major fluctuations in the output stream occur only when the reclaimer shifts from one pile to another pile with a different average composition. Larger piles are able to accomplish more averaging, thus reducing the variability among piles. Methods for choosing appropriate pile sizes are explained in Section 3.9.2.1 and illustrated in the case study of Section 4.2.

---

[5] Notice that effectively only 100 layers are deposited near the ends of the pile, where the stacker reverses direction, if 200 layers are deposited in the middle. Even fewer than 100 layers are present in the outermost edges of the end cones of the pile, where particle-size segregation also may occur. Between the ends and the middle of the pile, the amount of averaging that effectively takes place varies between an average of 100 evenly spaced samples and an average of 200 evenly spaced samples.

[6] In the language of Section 3.7, the "discretization error" involved in representing the composition of the pile by the average of 100 evenly spaced samples is very small.

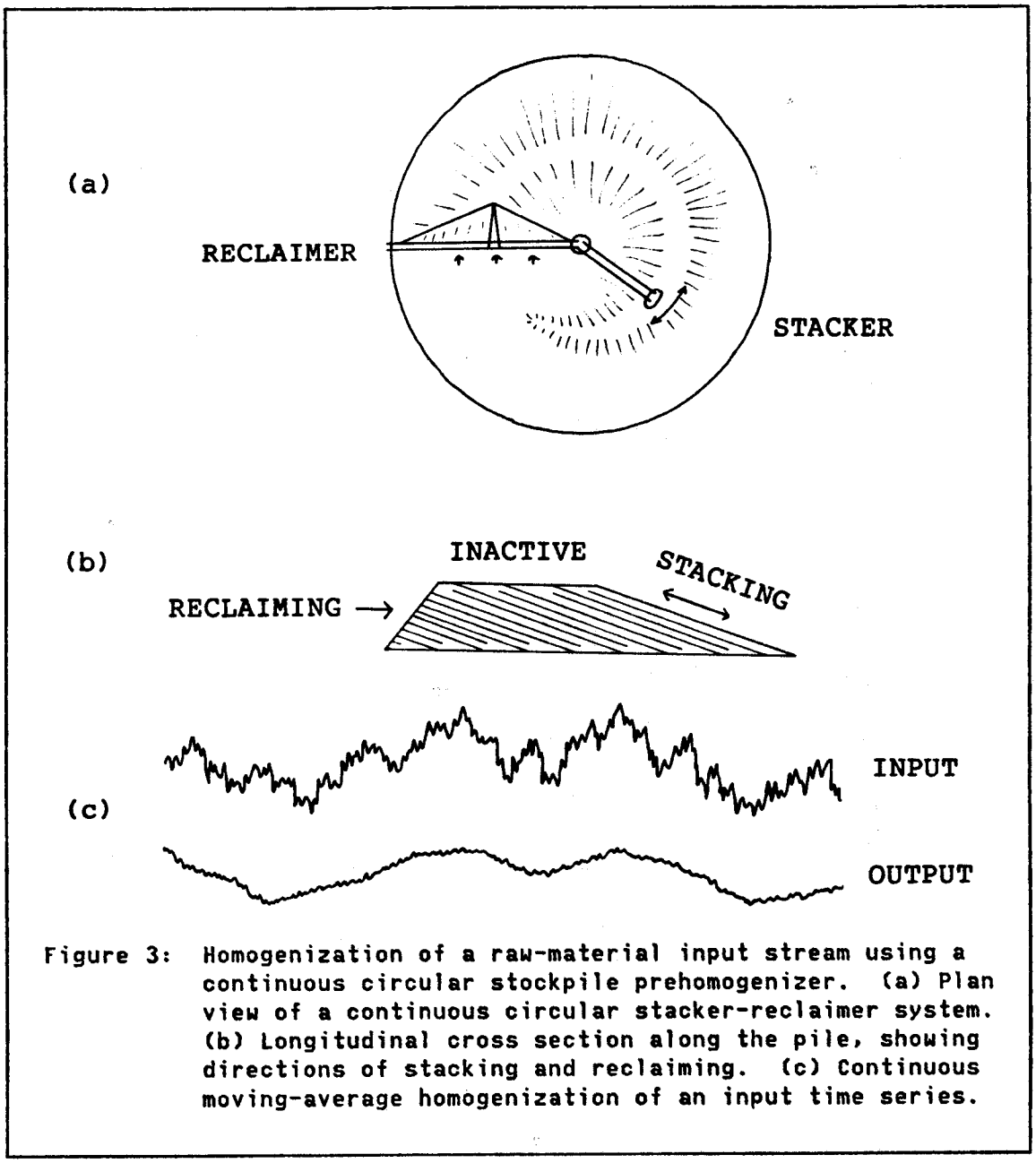


Figure 3: Homogenization of a raw-material input stream using a continuous circular stockpile prehomogenizer. (a) Plan view of a continuous circular stacker-reclaimer system. (b) Longitudinal cross section along the pile, showing directions of stacking and reclaiming. (c) Continuous moving-average homogenization of an input time series.

The one system that does not operate on a batch principle is the continuous circular pile illustrated in Figure 3(a,b). This system features a wedge-shaped pile whose tapered end gradually migrates around a circular track as the pile is built. The stacker moves back and forth

along the crest of the pile as before, but on each forward pass (clockwise in Figure 3(a), away from the reclaimer) it travels a bit farther forward before turning back, and on each backward pass (counterclockwise, toward the reclaimer) it correspondingly turns forward again a bit short of its previous turning point. Thus the tapered forward end of the pile gradually advances in the forward (clockwise) direction. Simultaneously, the reclaimer recovers the stacked material from the completed back end of the pile, moving forward (clockwise) at the same average rate at which the front of the pile moves forward. The time-series output of this system is essentially a simple moving average of the input series, illustrated in Figure 3(c). The composition of the output stream varies slowly and continuously, avoiding the sudden inter-batch jumps in composition that appear in Figure 2(b). One disadvantage of this system is that its capacity cannot be easily expanded. Another is that the stacker should be able to stack continuously over the whole length of the pile, eliminating the inactive completed part of the pile, in order for maximum homogenization (maximum residence time in the active part of the pile) to be achieved.

2.2.2.2 Proportioning

The calculation of raw-material mix proportions is deterministic, in that we must implicitly assume that all raw-material analyses are perfectly known and unchanging. In practice this is never the case, so proportioning always involves errors. Ways of dealing with the errors that occur because of short-term changes in raw-material compositions are discussed in Section 2.2.2.3. Some longer-term problems are illustrated in Section 2.3.3.

The distinction between long-term (design) and short-term (control) proportioning calculations is fundamental. When a plant is first designed, a long-term estimate of the quantity of each raw material to be used in the plant must be made available for proper sizing of the storage, weighing, and mixing facilities, and for a check on the adequacy of raw-material reserves. After the plant is built, long-term estimates are still needed occasionally for quarry planning and for the negotiation of contracts with suppliers of purchased raw materials. Linear programming is usually employed to obtain these estimates. (Details on this procedure and some alternatives are provided in Section 2.3.2.) The problem in these cases is to determine what proportions of raw materials will be needed over a long period of time to produce an acceptable clinker at minimum cost. It is recognized that the short-term mix proportions actually required during production will deviate from the long-term proportions because of variations in the compositions and costs of materials, and because of changes in the plant's product mix (e.g., mix proportions for Type I usually differ from those for Type II -- see Table 1). Plant designers incorporate some flexibility into a plant's proportioning facilities to accommodate these variations. (Just <u>how much</u> flexibility should be incorporated is addressed in Section 3.9.2.2.)

Several methods can be used for short-term control proportioning. Most plants proportion their materials to achieve a fixed quality "aim", or ideal mix composition.[7] The objective of control proportioning is

---

[7] Usually an ideal cement or clinker composition is determined first, and then adjustments are made for dust wastage, stack losses, fuel contributions, and perhaps gypsum addition, to obtain the corresponding raw-mix aim. These adjustments are explained more fully in Sections 2.3.1 and 2.3.2.

either to achieve this aim exactly, regardless of the resulting costs or quantities of materials used, or to minimize deviations from the aim, subject to constraints on cost, mix composition, or raw-material tonnages. Some plants may still try to minimize costs in a control situation (e.g., Laney, 1977), but in these cases the chemical constraints might have to be set very tightly so that large fluctuations in mix composition cannot occur in response to short-term changes in raw-material costs or compositions.

The most straightforward short-term proportioning calculation is by solution of a set of linear equations representing equality constraints on the composition or raw-material proportions of the mix.[8] The number of equations must equal the number of raw materials available. For example, suppose we are given analyses of SiO_2, Al_2O_3, Fe_2O_3, and CaO for four raw materials, and we want to produce 100 tons of raw mix satisfying the equality constraints LSF=0.95, SM=2.5, and AM=1.5 (Table 1). First we linearize the three ratio constraints. Instead of using LSF = C/(2.8 S +1.18 A +0.65 F) = 0.95, we use:

$$l_{lsf} = C - 0.95 (2.8 S + 1.18 A + 0.65 F) = 0$$

Similarly, SM = S/(A+F) = 2.5 becomes

$$l_{sm} = S - 2.5 (A + F) = 0$$

and AM = A/F = 1.5 becomes:

$$l_{am} = A - 1.5 F = 0$$

---

[8] Complicated "cookbook" procedures, evidently designed for those who do not have a computer and do not wish to perform matrix operations by hand, are still in widespread use. Actually they amount to special cases of the method described here. Cookbook algorithms for two-, three-, and four-component mixes are described by Peray (1979, Chapter 2) and Witt (1966, Chapter 4).

Then for each raw material, numbered i = 1 to 4, we can calculate $l_{lsf}(i)$, $l_{sm}(i)$, and $l_{am}(i)$ and set up the following system of equations:

$$\sum_{i=1}^{4} l_{lsf}(i)\, p(i) = 0$$

$$\sum_{i=1}^{4} l_{sm}(i)\, p(i) = 0$$

$$\sum_{i=1}^{4} l_{am}(i)\, p(i) = 0$$

$$\sum_{i=1}^{4} p(i) = 100 \text{ tons}$$

where p(i) is the proportion of raw material "i", expressed in tons.

If L represents the 4x4 matrix of coefficients, P the 4x1 vector of proportions, and K the 4x1 vector of equality constraints (the "aim"), this system is simply

$$L\ P = K$$

and the mix proportions are found from:

$$P = L^{-1}\ K$$

The solution P is feasible if none of the elements p(i) is negative, and if no p(i) lies outside the proportions that can be handled by the plant's weighing and conveying facilities. If in fact an infeasible P is found, then either the vector K must be altered slightly (e.g., to drive a negative p(i) to zero[9]), or if this is not possible, at least one additional material with a more suitable composition must be made available (along with an additional constraint). Some cement plants

---

[9] A procedure for doing this, involving inspection of the elements of L^{-1}, is explained by Niederjohn (1969).

keep small amounts of high-grade "additives" on hand for this type of situation. If a plant operates with the three specifications used in this example, then some (probably expensive) additives that could be used are silica sand, bauxite or kaolin, iron ore, or high-calcium limestone, which would provide very high-grade sources of S, A, F, or C, respectively. Of these, iron ore seems to be the most commonly used.

In control situations, we may not be required to satisfy an "aim" composition exactly but merely to minimize deviations from the aim. (The aim could actually be the composition of the immediately preceding mix, in which case we would be minimizing mix-to-mix variability.) This is a mathematical-programming problem of the general form:

MINIMIZE objective function

SUBJECT TO constraints on mix composition and mix proportions

Schofield (1980, Section 8.4) proposes an objective function of the form

$$\alpha(LSF_k-LSF_m)^2 + \beta(SM_k-SM_m)^2 + \gamma(AM_k-AM_m)^2 + (T_k-T_m)^2$$

where the subscripts k and m denote the aim and mix characteristics, respectively, T denotes the total tonnage of material being mixed, and α, β, and γ are weights chosen subjectively to reflect the relative importance of the three components of the aim. The optimal solution is found by a hill-climbing method that allows upper and lower constraints to be placed on any number of additional variables.

Niederjohn (1969) has proposed a Lagrangian approach starting with a similar objective function (stated more generally here),

$$\sum_n \mu_n (Q_{nk}-Q_{nm})^2$$

where the subscripts n denote different aim characteristics Q_n, the μ_n's denote the subjective weights, and the subscripts k and m denote the aim

and mix characteristics, respectively. Equality constraints to be met exactly are added to the objective function as additional terms, i.e.,

$$\sum_{n} \mu_n \, (Q_{nk}-Q_{nm})^2 - \sum_{j} \lambda_j \, (R_{jk}-R_{jm})$$

where the subscripts j denote the exact constraints R_j, and the λ_j's denote the Lagrange multipliers. The minimum is found by setting the partial derivatives of this expression with respect to each Q_{nm} and each λ_j to zero, and then solving the resulting set of simultaneous equations (much as in the "kriging system" of Section 3.1.4).

Another method that might be used for short-term proportioning is linear goal programming (Hillier and Lieberman, 1980, Chapter 5), which has been applied by Lonergan (1984) to the proportioning of coal supplies for power-plant feed. This is a "multi-objective" optimization technique, in which each target value of the aim composition is reformulated as an objective function. For example, if we want LSF to be as close as possible to 0.95 (as expressed in the example on page 23), then the linearized objective function for LSF is expressed as

$$\sum_{i=1}^{4} l_{lsf}(i) \, p(i) + a_{lsf} - b_{lsf} = 0$$

where a_{lsf} is a negative deviation variable and b_{lsf} is a positive deviation variable, representing the extent to which the actual mix composition (represented by the summation) deviates from the target value of zero. These deviation variables, plus those associated with other target values, are weighted according to their relative importance and combined into an "achievement function", which is then minimized by the goal-programming algorithm. Goal programming is proposed by Lonergan (1984) as an attractive alternative to linear programming

(Section 2.3.2) for both long- and short-term proportioning calculations in which some constraints may have to be violated to obtain a feasible solution.

Proportioning and homogenization can be performed simultaneously in a batch stacker-reclaimer system by stacking different raw materials onto the same pile and reclaiming them together. The raw materials might be stacked onto the pile one at a time, although better control can be achieved by continuously proportioning all materials onto the same conveyor belt ahead of the stacker. If the raw materials have constant known compositions (or have already been homogenized), the only control that is needed to achieve a correctly proportioned mix is an accurate weighing of the amount of each material added to the feed stream (Colijn, 1983). More commonly the raw materials have variable compositions, so the individual input stream from each raw material source should be repeatedly sampled and analyzed so that the total composition of the pile can be continually updated as the pile is built. (Methods for sampling streams of crushed material accurately and precisely are described and analyzed in great detail by Gy, 1982.) Any of the mathematical proportioning methods discussed in this section can then be applied periodically to determine the proportions of each raw material that should be added to the feed stream to guide the pile toward its desired final composition. Each time the proportioning calculation is performed, the composition of the existing partly built pile, which should be well established from previous sampling of the feed streams, is entered into the proportioning algorithm as a single raw material that must be used in a fixed amount. The compositions of

other raw materials to be added to the stream can be represented by their long-term expected values, or preferably forecasted by one of the methods described in Section 2.2.2.3. The proportioning algorithm then calculates the amount of each raw material that should be added to the pile to achieve the correct feed composition, and the feed rates are adjusted accordingly. These operations are performed periodically using newly updated pile compositions and newly estimated feed compositions. When the pile has achieved a large enough size and an acceptable total composition, a new pile can be started. A procedure of this type, using linear programming for proportioning, is described by Laney (1977). If the raw mix can be consistently proportioned and homogenized correctly by such a system, there may be no need for further homogenization of the raw mix before it enters the kiln. However, segregation commonly occurs during subsequent handling of the output stream, so that a powder homogenization system (Section 2.2.3) is usually required.

2.2.2.3 Sampling and Control of Mix Composition

> In some instances, one or more sources of disturbance may be measured and these measurements used to compensate potential deviations in the output. Such action is called feedforward control. In other situations, the only evidence we have of the existence of the disturbance is the deviation from target it produces in the output. When this deviation itself is used as a basis for adjustment, this action is feedback control. In some instances, a combination of the two modes of control is desirable, and this is referred to as feedforward-feedback control. [Box and Jenkins, 1976, p. 423]

The proportioning calculations described above require that the compositions of all raw-material inputs be known or reliably estimated. This would seem to be an obvious necessity in any plant producing a chemical product from variable raw materials. But at the present time,

a surprising number of cement plants (seemingly a majority in the United States) do not routinely sample individual raw-material feed streams, do not homogenize their raw materials prior to proportioning, and do not routinely use core-drill or blast-hole sample data to estimate the local compositions of individual mined raw materials. Quality control of mix composition in such a plant must be performed entirely by feedback, using only blended raw-mix compositions and perhaps the "average" compositions of the individual raw materials to calculate new mix proportions (Figure 4). The raw-mix composition is obtained either from samples of blended material leaving the raw mill or from samples extracted somehow from the post-mill homogenizing system. With only this information, only the most general kinds of control decisions can be made (e.g., "if LSF is too low in the mix, add more limestone"). This approach can work reasonably well for a plant with relatively uniform raw-material inputs or with a batch blending system that will not allow the consequences of occasional anomalies in raw-mix composition to cascade immediately into the kiln. In the case of extremely uniform raw materials (prehomogenized materials, or naturally uniform materials such as the "layer-cake" sedimentary deposits of the central United States), a plant may operate with practically constant mix proportions and obtain a practically constant mix composition.

This traditional approach cannot be expected to perform well in a plant with highly variable raw materials and a low tolerance for mix variability. (Then if LSF is too low in the mix, we would want to know precisely <u>how much</u> more limestone to add.) In such a case noticeable improvement may be obtained just by sampling the individual feed

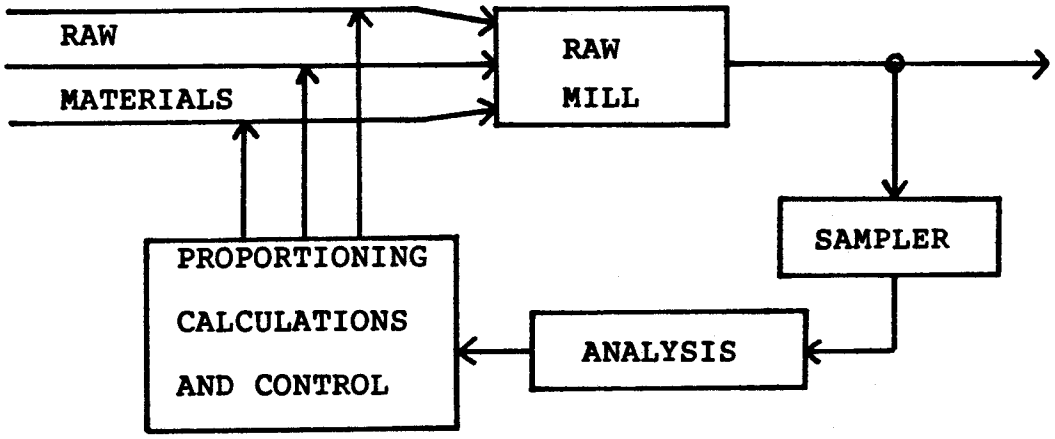


Figure 4: Feedback control of raw-mix quality using only raw-mix analyses.

streams, so that the source of any anomalous mix compositions that appear can be readily identified and corrected (Figure 5). By the definitions quoted above, the use of these individual analyses to control mix compositions constitutes feedforward control of the mix. However, each feed stream is being controlled only by feedback if only the most recent analyses are used in the proportioning calculations. If raw-material fluctuations are rapid and the time lag between sampling and control is long, direct corrections to mix proportions based solely on these samples may be no longer appropriate (possibly resulting in "oversteering" or "tail-chasing", in plant parlance). This is a common weakness of feedback control.

Forecasting of incoming raw-material qualities can be performed by time-series methods (Box and Jenkins, 1976; Chatfield, 1980) or by

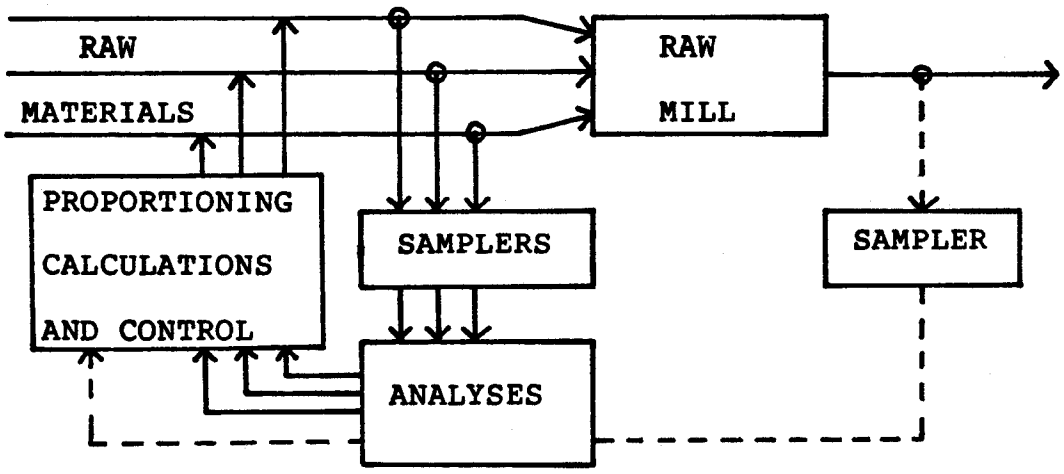


Figure 5: Control of raw-mix quality using individual raw-material analyses. The addition of a feedback loop (dashed) using raw-mix analyses may be advisable if some individual raw materials are poorly sampled.

geostatistics, provided that the individual feed streams have been well sampled and their statistical properties (particularly their autocorrelation functions or their variograms, Section 3.1.2) are well known (Schofield, 1980, Chapter 11). The general approach is to use present and past analyses of the individual feed streams to forecast their future compositions by regression, and to adjust mix proportions based upon these compositions.

Additional information on future compositions can be obtained from the quarry (Figure 6). Some plants with variable raw materials routinely analyze blast-hole cuttings to determine the compositions of blocks of of stone before they are mined. (The data set used for the first case study in Chapter 4 consists entirely of block compositions

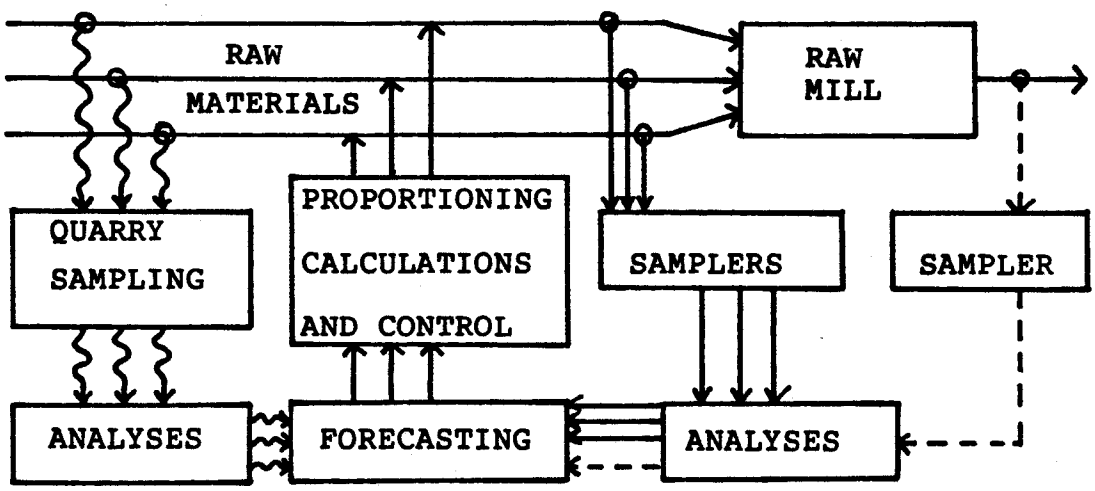


Figure 6: The full capabilities of feedforward control, using forecasts of individual feed-stream compositions. Both present and past feed-stream samples and samples collected far upstream in the process (wavy lines), usually in the quarry, are employed to predict future raw-material compositions.

estimated in this way.) Even before blast holes are drilled, forecasts of block compositions can be obtained by kriging (Section 3.1.4.), using data from surrounding core-drill holes and previously sampled blocks. Predictions of future feed-stream compositions can then be made by regression, using estimates obtained by time-series methods in combination with the estimated compositions of the mining blocks to be fed into the raw-material stream at that time. The exact forecasting formula to be adopted would vary widely with the situation.

The most effective, though initially most expensive, means of controlling raw-mix compositions in plants with extremely variable raw materials is to employ stacker-reclaimer systems to smooth out the variability of the individual streams, so that proportioning can be performed on homogenized materials. The most fruitful cement-related applications of the methods to be introduced in Chapter 3 (and of geostatistics in general) probably are in the design of these systems.

2.2.2.4 Sensitivity of Control Variables

Control difficulties result not only from variable raw materials, but also from the fact that some quality indices are exceedingly sensitive to changes in raw-material composition. A particularly bad offender is the Bogue formula for C_3S (Table 1), which is commonly used as a control variable. Consider the variance of C_3S:

$$\begin{aligned}\mathrm{Var}(C_3S) = {} & 4.071^2\mathrm{Var}(C)+7.6^2\mathrm{Var}(S)+6.718^2\mathrm{Var}(A)+1.43^2\mathrm{Var}(F)+2.852^2\mathrm{Var}(\underline{S}) \\ & - 2(4.071)\,[7.6\mathrm{Cov}(C,S) + 6.718\mathrm{Cov}(C,A) + 1.43\mathrm{Cov}(C,F) + 2.852\mathrm{Cov}(C,\underline{S})] \\ & + 2(7.6)\,[6.718\mathrm{Cov}(S,A) + 1.43\mathrm{Cov}(S,F) + 2.852\mathrm{Cov}(S,\underline{S})] \\ & + 2(6.718)\,[1.43\mathrm{Cov}(A,F) + 2.852\mathrm{Cov}(A,\underline{S})] \\ & + 2(1.43)\,[2.852\mathrm{Cov}(F,\underline{S})]\end{aligned}$$

Now consider that, in a mixture of limestone and clay (the usual cement raw mix), when CaO goes up, everything else involved in C_3S tends to go down. Furthermore, because SiO_2, Al_2O_3, Fe_2O_3, and sometimes SO_3 are major constituents of the clay or shale, when one goes up in the mix, usually they all do. In other words, all of the covariances of the form Cov(C, not C) listed in the second row above are large and negative, whereas all of the others are large and positive. Finally, notice that

the negative covariances are multiplied by large negative coefficients, whereas the positive covariances, and the variances in the first row, are multiplied by large positive coefficients. Clearly Bogue C_3S can have a huge variance, and yet many plants try to control it within a few percentage points. In this situation, even a slight sampling or analytical error can cause alarm. For example, in the "typical" cement analysis on page 6, if SiO_2 gained 1% and CaO lost 1% (an easily imaginable occurrence), the Bogue equations and moduli would be altered as below:

Index	Old Analysis	New Analysis
C_3S	46.52	34.85
C_2S	27.98	39.65
C_3A	9.50	9.50
C_4AF	9.13	9.13
SM	2.59	2.71
AM	1.83	1.83
LSF	0.92	0.87

Either C_3S or LSF (a similarly volatile index) would likely trigger an "add-more-limestone" order. This example illustrates the desirability of homogenizing each raw material before proportioning and the need for accurate and precise sampling and analysis for quality control. Although analytical capabilities in American plants are usually fairly good, and homogenization of the raw materials is being practiced in some newer plants, the sampling and sample-preparation techniques (Gy, 1982) are commonly crude and need improvement.

2.2.3 Grinding and Homogenization of the Raw Mix

The proportioned raw materials are fed into a mill that reduces the raw mix to a fine powder, typically with a top size of about 100 microns. This degree of fineness is necessary so that reactions among the chemical constituents of the different raw materials in the kiln can proceed quickly and completely. After grinding, a final stage of homogenization, commonly in an air-agitated silo containing at least several hours of raw-mix production, is usually included before the raw mix is released to the kiln. Homogenization at this stage is intended primarily to filter out short-scale fluctuations in feed composition caused by segregation in the reclaiming and grinding systems. Capacities usually are too small to filter out fluctuations on a day-to-day scale. Homogenization may be performed in one or more "cascade" silos, into which raw materials are constantly being added and withdrawn, or in a pair of "batch" silos, one of which is filled while the other is emptied. Batch systems may be preferred if some final fine tuning of the mix composition is likely to be needed. The characteristics of various silos are discussed by Schofield (1980, Section 2.6 and Chapter 9).

In many older cement plants, water is added to the raw mix at the grinding stage, or perhaps earlier in the process if some of the raw materials are already very moist. The ground raw-material slurry is then pumped into a stirred homogenization tank prior to entering the kiln. Because of the large increases in fuel prices that occurred in the 1970's, this "wet process" is now out of favor, as the water contained in the mix (commonly around 30-35%, unless reduced by

filtration) must be evaporated in the kiln. New plants usually employ a "dry process", wherein the raw mix is thoroughly dried before or during grinding, commonly with the use of hot exit gases from the kiln.

2.2.4 Burning

The following reactions occur during "burning" of the raw mix (Lea, 1971, p. 123):

Temperature (Celsius)	Reaction	Heat change
100	Evaporation of free water	Endothermic
≥500	Evolution of combined water from clay	Endothermic
≥900	Crystallization of amorphous dehydration products of clay	Exothermic
≥900	Evolution of carbon dioxide from $CaCO_3$	Endothermic
900-1200	Reaction between lime and clay	Exothermic
1250-1280	Commencement of liquid formation	Endothermic
≥1280	Further formation of liquid and completion of formation of cement compounds	Mainly endothermic

In a wet-process plant, all of these reactions occur within a long rotary kiln that is inclined about three degrees from the horizontal and rotates about once per minute. The raw material tumbles slowly down the length of the kiln toward a flame at the lower end. A fan draws hot air from the clinker cooler and combustion products from the flame up through the kiln, opposite the direction of raw-material flow. In the lower part of the kiln, nearest the flame, clinker temperatures may reach 1500 degrees and gas temperatures 2000 degrees Celsius. Shortly after reaching its maximum temperature, the clinker falls from the lower end of the kiln into a cooler, where fans blow outside air through the clinker as it passes over a grate.

Dry-process plants may also operate with a single long kiln, but most new plants realize some additional energy savings by combining preheater

or preheater-precalciner systems with a shorter, more rapidly rotating kiln. In a preheater system, the raw mix is first heated to about 800 degrees as it falls through turbulent exit gases from the rotary kiln. (Exit gases may also be used to dry the raw mix as it is being ground in the mill.) More efficient heat transfer from gas to raw material, with resulting energy savings, can be accomplished in this way. Some calcination (decarbonation of $CaCO_3$) occurs in the lower part of the system. Further calcination and clinker formation then occur in the short rotary kiln.

A precalciner, which may be inserted between the preheater and the rotary kiln, uses hot gases from the clinker cooler and a secondary flame to elevate mix temperatures well into the calcining range before material enters the rotary kiln. As the calcination reaction is highly endothermic, the application of additional heat to the mix at the location where most calcination occurs, instead of farther down the kiln, can result in further energy savings and permit greater production from a given size of kiln.

Some K_2O, Na_2O, and SO_3 evaporate from the raw mix in the clinkering zone of the rotary kiln and enter the gas stream. Additional SO_3 and some K_2O and Na_2O are contributed by the fuel (particularly by ashy coal). After the gas enters the cooler parts of the kiln system, these constituents condense onto dust particles entrained in the gas. Finer dust particles collect more of the condensate because of their greater surface-to-volume ratio. If the raw materials and fuel contain excessive levels of these constituents and an old-fashioned kiln system is being used, the heavily contaminated fine dust can be captured in

dust collectors and discarded, while the coarser dust is returned to the mix, thus lowering the content of K_2O, Na_2O, and SO_3 in the clinker. This is not possible in preheater systems, because the outgoing dust becomes thoroughly mixed with the incoming raw mix inside the preheater. To achieve the same effect in a preheater system, some energy efficiency may have to be sacrificed by allowing some of the kiln gas to bypass the preheater.

Additional problems may be created in the preheater if the molar ratio of alkalies (K_2O and Na_2O) to SO_3 in the gas departs greatly from 1.0 (Garrett, 1976). In this case, sticky coatings of low-melting alkali or sulfate compounds may accumulate inside narrow passages within the preheater and plug up the system. Even if the long-term ratio is in balance, short-scale fluctuations between excess SO_3 and excess alkalies, as well as fluctuations in other components of the kiln gas, may still cause plug-ups. Thus strong fluctuations in the compositions of the raw materials and fuel cannot be tolerated in a preheater system. This is one reason why stacker-reclaimer systems are desirable in preheater-equipped installations. Another reason is that the residence time of raw materials in a preheater-equipped kiln system is much shorter than in a long kiln system (about half an hour versus two to three hours), so preheater systems are inherently more sensitive to all kinds of short-scale changes in feed quantities or compositions. This is particularly true in the preheater itself, where residence times are about thirty seconds (Garrett, 1976, p. 6).

2.2.5 Cement Grinding

After the clinker leaves the cooler, it is generally sampled and stockpiled for a time while a chemical analysis is obtained. If the analysis of a clinker departs significantly from the aim composition, it may be blended with other clinkers to achieve the proper overall quality. The final stage in the manufacturing process involves grinding the clinker together with a small amount of gypsum or a gypsum-anhydrite blend -- usually enough to raise the SO_3 content of the cement by about one to two percent. These materials retard the setting times of concrete mixes and may somewhat enhance the strength of the concrete. Fineness of the ground cement is measured indirectly, and specifications are expressed in terms of surface area per unit weight.

2.3 CEMENT RAW MATERIALS

2.3.1 Specifications

Specifications on the compositions of cement raw materials naturally depend to a large extent on the chemical specifications applied to clinker or cement (Table 1). Usually only the chemical compositions of the raw materials are actively considered, as their mineral compositions and physical characteristics are destroyed in the kiln. However, materials that are known to be exceptionally tough, abrasive, or unreactive (coarse quartz grains, for example) may be avoided even if chemically suitable.

To understand the characteristics that cement raw materials should have, we must understand what happens to these materials in the cement-making process. The various chemical constituents of the raw materials

and fuel tend to segregate inside the kiln system and end up in different products. The flow of materials consumed and generated by the cement-making process is depicted in Figure 7. Notice that the raw mix is not the only input to the system. The fuel may introduce a large amount of sulfur and ash (essentially sintered shale and pyrite) into the kiln system, most of which ends up in the clinker. Some oxygen from the outside air introduced to the kiln also is picked up by the clinker and by other products of the kiln system. Raw-material and ash analyses (but not fuel sulfur) are usually quoted as though the materials were fully oxidized[10], so they do not have to be adjusted for oxygen added from the air.

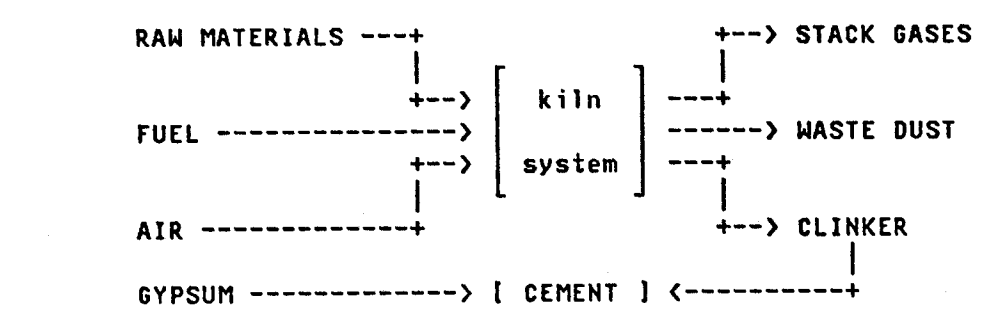


Figure 7: Materials consumed and produced during the manufacture of portland cement.

---

[10] This may result in analyses totaling well over 100% for some highly reduced materials, such as black shales or any material containing sulfides.

The kiln system produces three products: clinker, waste dust, and stack gases. Waste dust (if any) typically has a composition similar to the clinker, except that the proportions of K_2O, Na_2O, SO_3, CO_2 (reported as "loss on ignition"), and perhaps other volatile components (e.g., chloride) are elevated. Stack gases evolved from the raw materials and fuel consist mostly of CO_2, H_2O, and minor SO_2. The clinker that remains after the gases and waste dust are removed from the system is then interground with gypsum to produce the final product.

Because factors in addition to raw-material compositions determine the extent to which various raw materials can be used in a cement raw mix, and because "undesirable" characteristics in one raw material may be innocuous when the material is diluted among other raw materials, strict specifications (e.g., the "cutoff grades" commonly used in metals mining) are rarely placed on the characteristics of individual cement raw materials.[11] The suitability of raw materials is instead determined from the calculation of "mix designs", which take the influence of all components of Figure 7 into account in determining proper raw-material proportions.

2.3.2 Mix Designs

Out of the dust, . . . Now you are dust,
Out of the slime, Limestone and rust.
A little rust, I mold and I stir
And a little lime. . . . And make you again.

-- attributed to Beelzebub by E. L. Masters, Spoon River Anthology

---

[11] Of course, chemical restrictions do have to be written into most contracts for purchased raw materials and fuel, so that both parties understand what is being purchased.

The term "mix design" is industry jargon for a raw-material recipe that is calculated in advance to assure a good-quality cement. Short-term mix-design calculations for control proportioning have already been explained in Section 2.2.2.2. If one is satisfied with matching raw-material proportions to a fixed "aim" composition, these methods can be used for long-term planning as well. But more commonly, long-term mix designs are obtained by linear programming.[12] The general problem is stated as follows:

MINIMIZE: Total cost of raw materials

SUBJECT TO: Constraints on product composition

Constraints on raw-material availability

The calculations require the following data: unit production costs (usually excluding fixed costs) for each raw material, a "typical" or "average" analysis for each raw material (plus moisture if the costs are quoted for "as-received" materials), constraints (if any) on the availability of each raw material, specifications on the chemical composition of the clinker (in the form of linear constraints), estimates of the chemical contributions of fuel to the product (e.g., an estimate of tons of coal required to produce each ton of clinker, the ash content and sulfur content of the coal, and an analysis of the ash), an estimate of the composition of the waste dust, and an estimate of the proportion of total sulfur in the system that escapes as stack gas. The fuel contribution is treated as a raw material, and the waste products

---

[12] There are many good textbooks on linear programming. A best-selling general introduction to operations research, including linear programming and other optimization methods discussed here and in Sections 2.2.2.2 and 2.3.3, is Hillier and Lieberman (1980).

(dust and stack gas) are usually treated as raw materials with negative compositions.[13]

Clinker specifications are normally used as constraints. All other material compositions, costs, and proportions therefore must be converted to a "clinker", or "ignited", basis before the proportioning calculations begin, and then reconverted to their original forms to report the results.[14] Raw-material costs, tonnages, and analyses must all be adjusted for "loss on ignition" (essentially CO_2 and chemically bound water), which is reported in the analyses but is entirely lost to the stack gases during calcination of the raw materials.[15] Costs and tonnages may also have to be adjusted for loss of moisture. The tonnage and analysis of the waste dust must be adjusted for loss on ignition. Several waste dusts with different compositions may be used in the calculation, to reflect the fact that fine, high-alkali dusts are always thrown away first and coarse, relatively low-alkali dusts are thrown away last. The relative amounts and compositions of these dusts depend on the type of dust-collection equipment installed in the plant. Normally sulfur (represented by SO_3 in the calculations) is the only

---

[13] Data on gas and dust compositions are lacking if the plant has not yet been built. Furthermore their compositions depend heavily on burning conditions in the kiln and on the amount of dust wasted, and both of these factors are under the control of plant personnel. Hence for plant design purposes these compositions must be subjectively chosen from a wide range of possible values.

[14] However, if only ratio constraints such as LSF, SM, and AM (Table 1) are being used, it is not necessary to convert to ignited basis, as the numerators and denominators of the ratios are inflated equally by the calculation.

[15] For instance, to convert each oxide percentage, x_r, in a raw-material analysis to an ignited percentage, x_i, use the relation

$$x_i = 100\, x_r / (100 - y_r)$$

where y_r is the loss on ignition of the raw material.

component of the stack gas whose flow must be measured or estimated, as the other components of the gas disappear from the calculation as loss on ignition. The contribution of outside air does not need to be measured either, as all analyses are reported as though the materials were fully oxidized.

Once all materials have been adjusted to ignited basis, the linear-programming problem can be straightforwardly written:

$$\text{MINIMIZE: } \sum_{i=1}^{I_r} p(i)\, c(i)$$

SUBJECT TO:

(Chemical constraints) $$\sum_{i=1}^{I_t} p(i)\, l_j(i) \langle = \rangle k_j,\ j = 1 \text{ to } J$$

(Material availabilities or fixed tonnages) $$\sum_{i=1}^{I_t} p(i)\, l_m(i) \langle = \rangle q_m,\ m = 1 \text{ to } M$$

(Material balance) $$\sum_{i=1}^{I_r} p(i) + p(\text{fuel}) - p(\text{dust}) - p(\text{gas}) = \text{clinker tonnage}$$

(Nonnegativity) $$p(i) \geq 0,\ i = 1 \text{ to } I_t$$

where $p(i)$ denotes the ignited tonnage of material i

$c(i)$ denotes the ignited cost per ton of material i

I_r denotes the number of raw materials

I_t denotes the total number of materials aside from clinker; i.e., raw materials, fuel contribution, dust, and gas

$l_j(i)$ denotes the value of linear "chemical function" j, for material i

k_j denotes the value of chemical constraint j (a clinker specification)

J denotes the number of clinker chemical constraints

$l_m(i)$ denotes the coefficient of linear "availability function" m, for material i

q_m denotes the value of "availability constraint" m

M denotes the total number of availability constraints

<=> denotes ≥, ≤, or =.

In this formulation, waste dust and stack SO_3 are assigned positive tonnages p(i) to satisfy the nonnegativity constraints, but negative chemical compositions are used to calculate the coefficients $l_j(i)$ for dust and stack SO_3.

Some embellishments to this type of formulation can be made to reflect more accurately the total cost of the mix. For instance, costs for dust disposal can be added to the objective function, although dust wastage will be minimized by the above formulation anyway. (More dust wastage necessitates more raw materials, because of the material-balance constraint. If several dust analyses are entered, the most contaminated dusts will normally be discarded first, to minimize wastage.) It is also possible to develop objective functions that account for the relative energy costs of different mixes. For example, Xirokostas and Zoppas (1977) have proposed a nonlinear objective function incorporating an extra term in which the energy cost is regarded as a function of the silica modulus SM of the mix; i.e.,

$$\sum_{i=1}^{I_r} p(i)\, c(i) + A\,(\,SM - SM_{min}\,)$$

where SM = S/(A+F) is the mix silica modulus, SM_{min} is the minimum constraint on SM for their problem, and the factor A represents the approximate increase in energy costs per unit of SM above its minimum.

For long-range-planning purposes, one may wish to incorporate one or more mix designs into a much larger optimization model, wherein an optimal combination of raw-material mixes, product mixes (proportions of different types of cement to be produced), and shipping strategies, possibly formulated as a multi-period cash-flow problem, can be found simultaneously (e.g., Gershon, 1982). Integer programming can be incorporated into the model to include fixed costs incurred with the use (versus non-use) of different raw materials or products, to devise optimal mining plans requiring either-or or conditional decisions, and to decide on alternative locations for facilities (e.g., Barbaro and Ramani, 1983; Hillier and Lieberman, 1980, Section 18.5).

2.3.3 An Illustration of a Mix Design and Its Weaknesses

Simple linear-programming problems involving two or three raw materials can be represented graphically. For example, suppose that over the life of a certain cement plant, we expect to produce 100 million tons of clinker from three raw materials. We will waste no dust, use a low-sulfur natural gas as fuel, and use low-sulfur raw materials. Therefore the problem simply involves proportioning the three raw materials to match a set of specifications. Ignited-basis analyses and costs of the three raw materials are:

Raw Material		S	A	F	C	$/ton
High-Calcium Limestone	(X)	4	1	0	95	3
Siliceous Limestone	(Y)	20	4	1	75	2
Clay	(Z)	60	35	5	0	1

We wish to keep C_3S between 45% and 55%, SM between 2.5 and 3.0, and C_3A less than 10%. Furthermore, the siliceous limestone lies on top of the high-calcium limestone, such that to avoid wasting some siliceous limestone, we must use at least two (ignited) tons of siliceous stone for every ton of high-calcium stone. Finally, we have reserves of only 40 million tons of clay. Thus the problem is (in millions of tons):

$$\text{MINIMIZE:} \quad 3X + 2Y + Z$$

$$\text{SUBJECT TO:}$$

$$(C_3S)_X X + (C_3S)_Y Y + (C_3S)_Z Z \geq (45)(100)$$

$$(C_3S)_X X + (C_3S)_Y Y + (C_3S)_Z Z \leq (55)(100)$$

$$\frac{(\ 4X + 20Y + 60Z\ )}{(\ X + 4Y + 35Z\ ) + (\ Y + 5Z\ )} \geq 2.5$$

$$\frac{(\ 4X + 20Y + 60Z\ )}{(\ X + 4Y + 35Z\ ) + (\ Y + 5Z\ )} \leq 3.0$$

$$(C_3A)_X X + (C_3A)_Y Y + (C_3A)_Z Z \leq (10)(100)$$

$$Y / X \geq 2$$

$$Z \leq 40$$

$$X + Y + Z = 100$$

$$X, Y, Z \geq 0$$

Calculating the C_3S and C_3A coefficients and linearizing the ratio constraints, the above formulation is simplified to:

$$\text{MINIMIZE:} \quad 3X + 2Y + Z$$

$$\text{SUBJECT TO:}$$

$$349.6\ X + 125.0\ Y - 698.3\ Z \geq 4500$$

$$349.6\ X + 125.0\ Y - 698.3\ Z \leq 5500$$

$$1.5\ X + 7.5\ Y - 40.0\ Z \geq 0$$

$$1.0\ X + 5.0\ Y - 60.0\ Z \leq 0$$

$$2.7\,X + 8.9\,Y + 84.3\,Z \leq 1000$$

$$-2.0\,X + 1.0\,Y \geq 0$$

$$Z \leq 40$$

$$X + Y + Z = 100$$

$$X, Y, Z \geq 0$$

The last two constraints above (material-balance and nonnegativity) define an equilateral triangle within the three-dimensional space of (X,Y,Z). The problem can thus be displayed on a triangular composition diagram, as in Figure 8, by plotting lines representing the intersections of all other constraint planes with the plane X+Y+Z=100.

Figure 8 displays several interesting features that can help us understand the typical behavior of mix-design problems. First notice that the two constraints on C_3S plot as two parallel lines, exceedingly close together. These constraints alone remove the vast majority of the triangle from feasibility. The closeness of the two lines reflects the extreme sensitivity of the Bogue C_3S formula to small changes in raw-mix composition. Even a tiny shift in mix proportions (especially, a change of more than 1% in the amount of clay) could result in an unsatisfactory mix. Obviously then, C_3S is a highly restrictive specification to use in a mix design, particularly if it is given both upper and lower limits.

The two constraints on SM plot as not-quite-parallel lines that are somewhat less restrictive than C_3S. The C_3S constraints, in conjunction with the nonnegativity constraint on high-calcium limestone (i.e., $X \geq 0$), render the constraint $SM \leq 3.0$ redundant, as the entire feasible region already lies within the area where $SM \leq 3.0$. Notice that the two tonnage

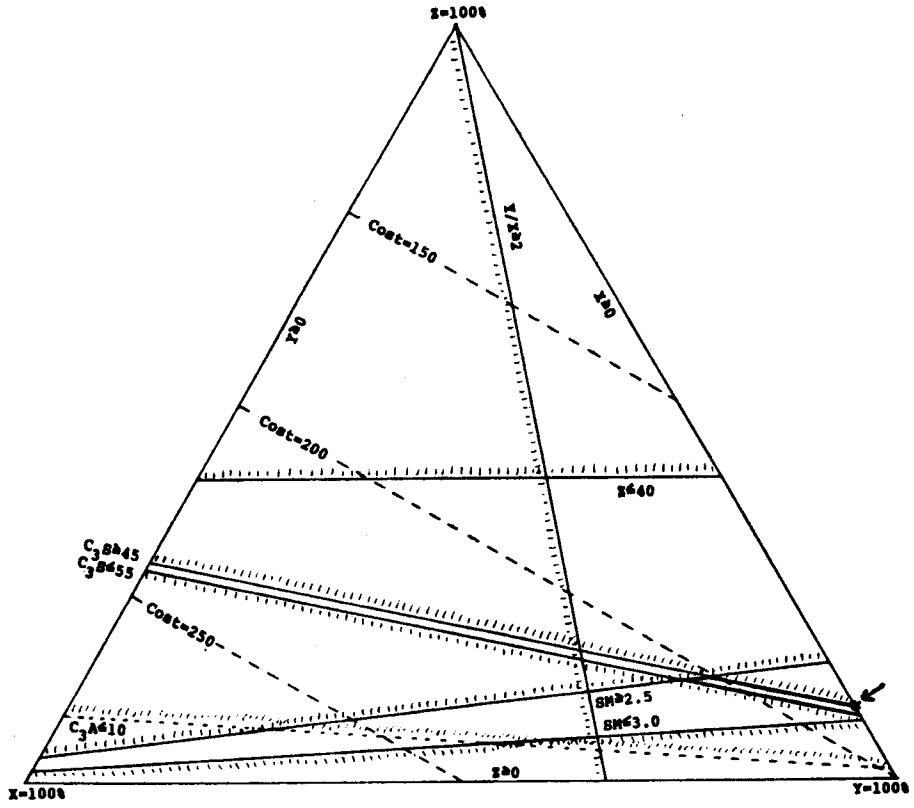


Figure 8: Graphical solution of a simple linear-programming mix-design problem. The hachures mark the infeasible sides of the inequality constraints. The constraint on C_3A (dotted) had to be removed to obtain a feasible solution. The orientation of the contours of the cost function (dashed lines) make it plain that the minimum-cost feasible mix is located at (X,Y,Z)=(0, 90.3, 9.7), indicated by "*" on the graph.

constraints, $Y/X \geq 2$ and $Z \leq 40$, are also redundant; we do not need much clay or high-calcium limestone to make this mix.

Now a dire situation erupts when we apply the single constraint on C_3A. This constraint excludes the entire region bounded by $X \geq 0$, $SM \geq 2.5$, and $45 \leq C_3S \leq 55$, which had previously been considered feasible. Thus the

problem has no feasible solution, regardless of cost. We simply cannot make a low-C_3A clinker with only these three raw materials. Even removing the restrictive constraints on C_3S would not help, as then the constraints SM≤3.0 and Y/X≥2 would still conflict with C_3A≤10.

In practice, this difficulty would probably be solved by purchasing a fourth material very low in C_3A, such as iron ore. But then the problem could not be visualized on a single two-dimensional graph, so for purposes of illustration let's assume that we can live with a very high-C_3A clinker. Then the only remaining issue is the choice of a single set of mix proportions (X,Y,Z) from the infinite number of choices within the feasible region bounded by X≥0, SM≥2.5, and 45≤C_3S≤55 (and of course by X+Y+Z=100 million tons). In the statement of this problem, the criterion chosen to select such a proportion was minimization of the linear cost function 3X+2Y+Z. The dashed lines in Figure 8 are the contours of this function where it intersects the plane X+Y+Z=100 million tons. The minimum raw-material cost of $1.90 per ton of mix is achieved at (X,Y,Z) = (0, 90.3, 9.7), which is fixed at C_3S=45 (the minimum), X=0 (the minimum -- no high-calcium stone is used), and X+Y+Z=100 million tons.

This solution, like all optimal linear-programming solutions, lies exactly on a set of constraints. In this case three constraints are needed to fix the three variables X, Y, and Z. It is always wise, then, to set the chemical constraints somewhat narrower than their official limits (e.g., narrower than the ASTM limits in Table 1), so that very slight changes in mix proportions or raw-material compositions cannot easily cause the product to fall out of specifications.

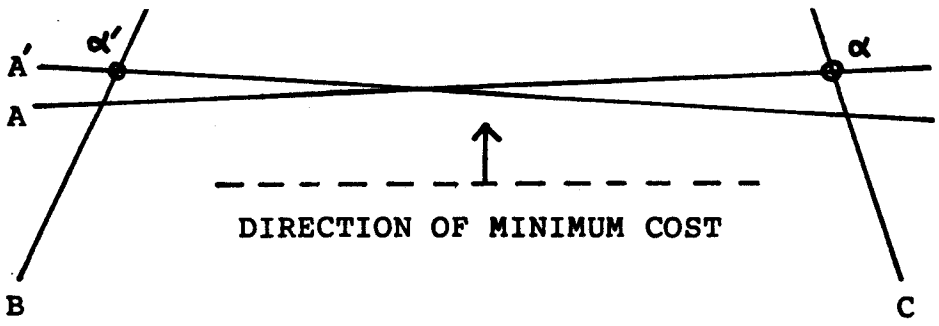


Figure 9: A possible effect of a small change in raw-material compositions on optimal mix proportions. If compositional changes shift constraint A to constraint A′, the minimum-cost mix will shift from point α to point α′.

Now what if the raw materials vary greatly in composition, instead of having constant known compositions, as implicitly assumed here? In Figure 8, variable compositions would cause the slopes of the chemical-constraint lines to vary. If the contours of the cost function were nearly parallel to one or more constraints, and if the constraints changed orientations slightly because of changing raw-material compositions, the location of the optimal solution could change a great deal (Figure 9). The convex feasible region might even disappear or become defined by a different set of constraints. For example, with a small shift in either the SM or the C_3S of one of the raw materials in the example, the constraint SM≤3.0 might become a boundary of the feasible region, possibly even replacing X≥0 and thus forcing the addition of high-calcium limestone to the mix.

Clearly, when raw materials are highly variable in composition, a deterministic linear-programming solution of this type is of limited

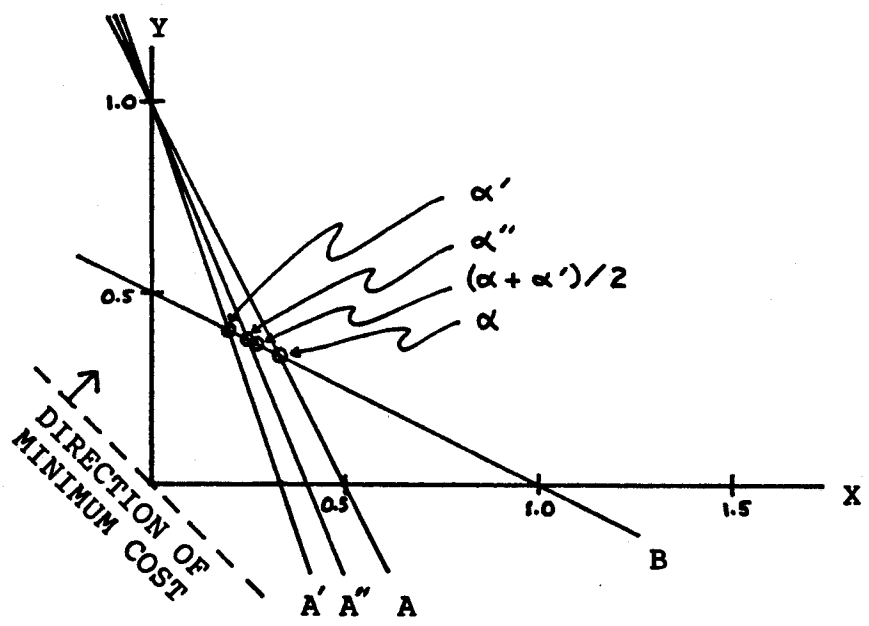


Figure 10: A simple example of the difference between a minimum-cost mix using average analyses and the average of minimum-cost mixes. Two constraints A and A′, resulting from different compositions of raw-material X, are used to determine two minimum-cost mixes, α and α′. The average compositions of raw-material X, resulting in constraint A″, yields another minimum-cost mix, α″, which is higher in cost than the average of mixes α and α′. In this example the constraints are:

A	2X + Y = 1
A′	3X + Y = 1
A″	2.5X + Y = 1
B	X + 2Y = 1

value, and might be severely misleading. Furthermore, the average minimum raw-material cost of the mix is not generally the minimum cost obtained by using average raw-material compositions (Figure 10), so the solution obtained by using the average analyses of variable materials may not be truly optimal, although it is probably close in most cases.

Nevertheless, some method must be used to obtain an impression of the mix proportions needed by the plant. In Section 3.9.2.2, some methods are suggested for determining whether the answer yielded by linear programming is likely to be reasonable, and how much flexibility in mix proportions should be anticipated when designing a new proportioning system to accommodate previously unmined materials.

2.4 INVESTIGATIONS OF CEMENT RAW MATERIALS

> After all is said, the services of a geologist in the cement industry are needed solely for the purpose of enabling the company concerned to obtain the raw materials necessary with the least expenditure of money and effort. [B. L. Miller, 1934, p. 40]

2.4.1 Exploration

Cement is a low-unit-value product, and most cement plants operate with a low profit margin. This has three important implications for raw-materials exploration. First, the major raw materials (and the plant site) should be located close to the major cement market on good transportation routes, so that transportation costs do not greatly increase the cost of cement delivered to the market. Second, the production costs of raw materials must be kept low. Expensive mining methods -- whether necessitated by thick overburden, complicated geologic structure, an abundance of caves or clay seams, excessive groundwater inflows, or zones of waste or off-grade materials that would have to be mined selectively -- are rarely affordable. And third, raw materials that would greatly increase handling or processing costs in the plant are preferably avoided. Such materials might include: sticky

wet clays that create handling problems; cherty or sandy limestones that cause excessive wear of crushing and grinding equipment and drive up fuel costs; or any materials with undesirable or highly variable characteristics that necessitate extra homogenization and proportioning efforts, require importation of exotic materials to counteract their effects, require dust wastage, or increase fuel consumption.

Of course, in many parts of the world all of these preferences cannot be satisfied simultaneously. This is particularly true in regions with a shortage of high-quality raw materials. In these regions, the need to minimize transportation costs drives many cement producers to put up with raw-material problems that producers in other areas would find intolerable. Even in some formerly favorable areas, the best plant sites have now been occupied or withdrawn from consideration for environmental or political reasons, and the best raw materials have been depleted. Thus production costs have increased as producers have resorted to less favorable raw materials and sites.

Increased fuel prices have also made some formerly desirable materials less so (particularly high-moisture and low-burnability materials), and have spurred the installation of more energy-efficient processes that are less tolerant of variations in unhomogenized raw materials. Environmental regulations have also necessitated stricter monitoring of the composition and variability of plant emissions and waste products. Thus external developments are bringing stricter controls on raw-material quality, even though marginal-quality raw materials are being forced into greater use each year. These trends must be kept in mind when selecting potential plant and mine sites.

Cement is made primarily from relatively widespread sedimentary raw materials. This is fortunate, as it helps to keep transportation costs down in most regions. The geographic distribution of materials likely to be usable for cement manufacture can usually be readily determined from geologic maps, government geological survey reports, and quick field reconnaissance. Except in less developed areas where there is little published information, or in areas where some of the needed materials happen to be scarce (e.g., limestones on the west coast of the United States), exploration focuses less on finding the materials than on finding the best sites for mining and processing the materials. Ideally, a mine site should have at least the following characteristics: abundant[16] good-quality raw materials, low mining costs and safe mining conditions, absence of serious environmental constraints (including close neighbors), and proximity to a good plant site. The plant site should be located on stable ground (cement plants are very heavy), adjacent to major highway and rail facilities, preferably adjacent to a navigable waterway, near electric-power supplies, near a source of water, and near sources of any necessary purchased raw materials and fuel. The site should also be near enough to populated areas that the plant can be staffed, routine services can be provided, and some of the product can be sold locally, but not so near that environmental and zoning restrictions would impede operations.

---

[16] Typically a new plant should have reserves adequate for at least 50 to 100 years of production. This is desirable because payback times for plant investments can be long, because creeping urbanization or other unforeseen factors may ultimately rule out some reserves, and because future plant expansions may result in much more rapid depletion of reserves.

The first stage of an exploration project usually is to define a search area (typically an area surrounding a good potential market) on a map and immediately locate all sites within that area where seemingly acceptable raw materials coexist with adequate infrastructure (transportation, power, water, and services). This simple step usually narrows the search to only a few sites (and in many cases will reveal that all of the best sites are already occupied). Then zoning, environmental restrictions, property ownership and likely availability of property, and the general geology and mining situation are checked out for those few sites. A "most favorable" site can usually be selected at this point. Then the detailed geology of this site is mapped in the field, and surface sampling of the raw materials is performed wherever outcrops are available, to determine whether the compositions of the materials appear to be suitable. If the detailed geology and field analyses appear favorable, the property may be optioned or claimed so that more detailed field studies can be performed to obtain information adequate for a thorough feasibility study.

2.4.2 General Site Evaluation

Widely spaced core drilling, groundwater pumping tests, and perhaps geophysical methods are used to obtain general subsurface geologic and hydrologic information to assess geologic structure and mining conditions, and to obtain reliable chemical and mineralogical analyses of the raw materials available at the mine site. A geotechnical study of the plant site may also be performed at this time. At this point, it is usually not necessary to spend large sums of money for a dense

program of core drilling, sampling, and analysis. One needs only enough chemical information to identify the distinctive rock types that are present on the property, to obtain a reliable estimate of the total quantity of each rock type, and to obtain a reliable estimate of the average composition of each rock type and a rough appraisal of its variability. A preliminary mining plan is needed to determine the limits within which reserve calculations should be performed.

The average compositions and total tonnages of the various materials identified on the property are used in mix-design calculations to determine whether a suitable cement mix can be made, and to estimate whether the property will be essentially self-sufficient in raw materials, or whether additional property or purchased materials will be needed. All of this information, combined with independently gathered marketing, technical, and economic data, should be adequate to perform a careful feasibility study of the total project, and in most cases will be adequate to reach a decision on whether to commit over $100 million to construction of a new plant.

2.4.3 Detailed Site Evaluation

Traditionally, the evaluation of cement raw materials has stopped at the "general" stage described above. The drilling of more than fifty or so holes on a single property at this stage has been widely viewed as a profligate expenditure of company funds, and the drilling of two holes closer together than a few hundred feet has been considered a sign of geological ineptitude, unless the rocks were clearly in a structural mishmash. These views developed for good reasons, based on many

people's accumulated experience in evaluating cement raw materials. First of all, sparse drilling is perfectly adequate in many cases because of the virtually constant compositions of many sedimentary raw materials. Furthermore, the looser restrictions on feed variability that were permissible in many older plants made it unnecessary to evaluate local raw-material variability very carefully. And finally, until the advent of geostatistics in the past few years, no method existed for translating information on the variability of drill-core analyses into information on the variability of raw-feed compositions.

Nevertheless, the remarks quoted at the beginning of Chapter 1 indicate that the traditional methods of cement raw-material assessment are no longer adequate for the design of most modern plants. In particular, if the chemical compositions of the raw materials appear to be extremely variable, then the more sophisticated mining methods, extra homogenization capacity, increased flexibility in proportioning and storage, and more sophisticated sampling and quality-control methods that will be needed may add several million dollars to the capital cost of a plant and millions more to the total discounted operating cost of the plant throughout its life. Therefore it is important to obtain information on raw-material variability, so that these additional costs can be properly evaluated.

Reliable information on <u>variability</u>, as opposed to average compositions, requires considerably more drilling data in most cases. Furthermore, information on <u>small-scale</u> variability -- the hour-to-hour and day-to-day variability in the compositions of delivered raw materials that is of greatest importance in quality control and process

stability -- can be determined only from closely spaced data on raw-material compositions. Still, the additional expense of gathering more data (treated further in Chapter 5) can be justified only if methods are available to recast the drilling data into a form usable by plant designers. Chapter 3 of this dissertation describes these methods, Chapter 4 provides some examples of their execution and practical application, and Chapter 5 discusses the costs involved and the factors that should be considered in making a decision about whether more data and more sophisticated methods of raw-material assessment are needed for a specific project.

Chapter III

SIMULATION OF COREGIONALIZATION: BASIC METHODS, EXTENSIONS, AND APPLICATIONS

This chapter introduces a few basic concepts that are needed for an understanding of the methods of conditional simulation of coregionalization (spatial cross-association among variables), and then describes those methods and some of their applications. Section 3.1 provides a brief introduction to linear geostatistics. Sections 3.2 through 3.4 discuss the methods used for classical conditional simulations of coregionalization for stationary gaussian spatial processes (the simplest case). The remaining sections present some extensions to these techniques for the simulation of more complicated phenomena and provide some practical advice on real-world applications. Two case studies, involving simulations of limestone deposits, are presented in Chapter 4.

3.1 GEOSTATISTICAL CONCEPTS: A BRIEF OVERVIEW

3.1.1 Scope: Geostatistics and Its Applications

Let's begin by reviewing a very unfortunate bit of terminology. The term "geostatistics" naturally should mean "statistics applied to the earth sciences", and originally that is exactly what it did mean. Unfortunately, during the past several years the definition of geostatistics has gradually been transformed into "applications of the theory of regionalized variables", regardless of whether those

applications happen to be in the earth sciences. Thus we now find the prefix "geo-" being used in the name of a discipline that is not necessarily concerned with earth-science problems at all. Moreover, the general field of statistical applications in the earth sciences has now been rendered nameless. (The common terms "mathematical geology" and "geomathematics" certainly are not appropriate replacements.)

Geostatistics (as the term is now applied) is used to describe phenomena that appear to vary continuously in space or time according to a spatial probability law. Examples of such variables from geology include ore grades in mineral deposits, elevations or thicknesses of sedimentary beds, chemical analyses of rocks, porosities of aquifers or oil reservoirs, and strengths of earth materials. Examples outside geology include rainfall data, concentrations of environmental pollutants, crop yields, densities of forest species, and most data commonly analyzed by time-series methods. In mining applications, useful results that can be obtained by geostatistical methods include optimal estimates of total reserves, local ore grades, and recoverable reserves in mineral deposits, estimates of the variability of different-sized blocks of ore, selection of the best drilling locations for the improvement of the quality of ore estimation, and simulations of the spatial distribution of ore grades for use in mine and process planning (covered in the case study of Section 4.2).

A clear and concise introduction to basic geostatistical concepts can be found in Chapter I of Journel and Huijbregts (1978). The terminology and most of the notation used in this dissertation conform to their

usage.[17] Their volume and that of David (1977) are the most complete texts covering all of linear geostatistical theory and its mining applications. This dissertation provides only a very brief introduction, stressing those concepts that are most important in applied simulations of coregionalization.

There is an important conceptual hurdle that one must leap before one can apply, with a clear conscience, the methods of geostatistics in a study of real-world phenomena: one must be willing to accept the idea that a geostatistical point of view is valid for the phenomenon under study. Some geologists, and other natural scientists who are closely acquainted with complex natural phenomena in a field setting, are uncomfortable with the idea that such complicated variables as ore grades can be realistically represented by any kind of mathematical model. But several points can be made in response:

(1) It is true that some situations cannot be rendered into a neat probabilistic form very easily. A situation where several different rock types have been mixed together by folding, faulting, and igneous intrusion is a fine example. But a thinking geostatistician working in such a setting must be, among other things, a thinking earth scientist, and as an earth scientist he will recognize situations where a straightforward probabilistic model does not apply and alter his model or the scale of his study accordingly.

---

[17] A few exceptions to the notation of Journel and Huijbregts (1978) have been necessary, primarily because of limitations in the character set used to print this dissertation.

(2) For many applications, the geostatistical model employed does not need to fit all observable aspects of a natural phenomenon exactly. The important question to ask when fitting a model to a set of data is not "How perfect is this model as a representation of the complete phenomenon?", but rather "How adequate is this model as a representation of those aspects of the phenomenon that are of practical importance?". It is especially important to consider the scale at which observations have been made in comparison with the scale of the geostatistical study. For example, if we are interested only in the ore grades of large mining blocks, then the microscopic features of the ore mineralization are irrelevant to the problem, even if the model we are using would be clearly inappropriate at very small scales.

(3) Ultimately, geologic processes may be regarded as deterministic, yet this does not mean that we cannot use a probabilistic approach in analyzing the results of those processes. So many different influences converge to produce a given result that, given our ignorance of their individual effects, we might as well view the outcome as probabilistic. Moreover, Journel (1985) has demonstrated that deterministic and probabilistic formulations of spatial estimation problems ultimately boil down to the same estimation methods, merely expressed in different mathematical languages.

(4) Geostatistics, in common with other probabilistic approaches, is a way of recognizing and accounting for our ignorance of the phenomenon under study. A situation in which the geologist has reason to doubt the accuracy of a carefully constructed geostatistical estimate is usually one in which an estimate by any other means would be equally bad or

worse; but at least the geostatistical estimate is accompanied by an "estimation variance", which can tell the geologist how bad the estimate is likely to be.

(5) Finally, although the introduction to geostatistical methods presented in this chapter covers only very simple models, this simplicity is only in the interest of brevity. If enough data are available (including qualitative geological data), geostatistical techniques can be "custom-built" to fit virtually any situation. The case study of Section 4.2 provides one example of how several simple techniques can be combined and adapted to a nonideal situation, through careful data analysis, consideration of qualitative geological information, and "conditioning" of the model to the data.

3.1.2 Random Functions and Regionalized Variables

The type of phenomenon that we can examine by means of geostatistics is one that can be viewed as a realization of a spatial random function (random process, or random field) having a value at every point in space.[18] The space can have any number of dimensions -- usually one, two, or three. A two-dimensional example of such a realization is presented in Figure 11, which is a contour map of simulated porosities in a thin (effectively two-dimensional) sandstone aquifer.[19] Porosity is considered to be a random quantity distributed over this two-dimensional

---

[18] As emphasized in the previous section, the phenomenon does not have to be a realization of a random function, but it must resemble one enough to allow a random-function model to be constructed from the data.

[19] This simulation was a preliminary step in the generation of a more complicated simulation, called "Field 1" by Helwick and Luster (1984).

space, but "randomness" does not necessarily imply a complete lack of structure. The porosities of any two neighboring points are not independent but tend to be positively correlated with each other -- in general, the closer together the points, the more similar their porosities. Furthermore, neighboring points a given distance apart tend to be most similar if they are located along a northwest-southeast line, and least similar if they are located along a northeast-southwest line. Thus, the degree of correlation depends on direction as well as on distance.

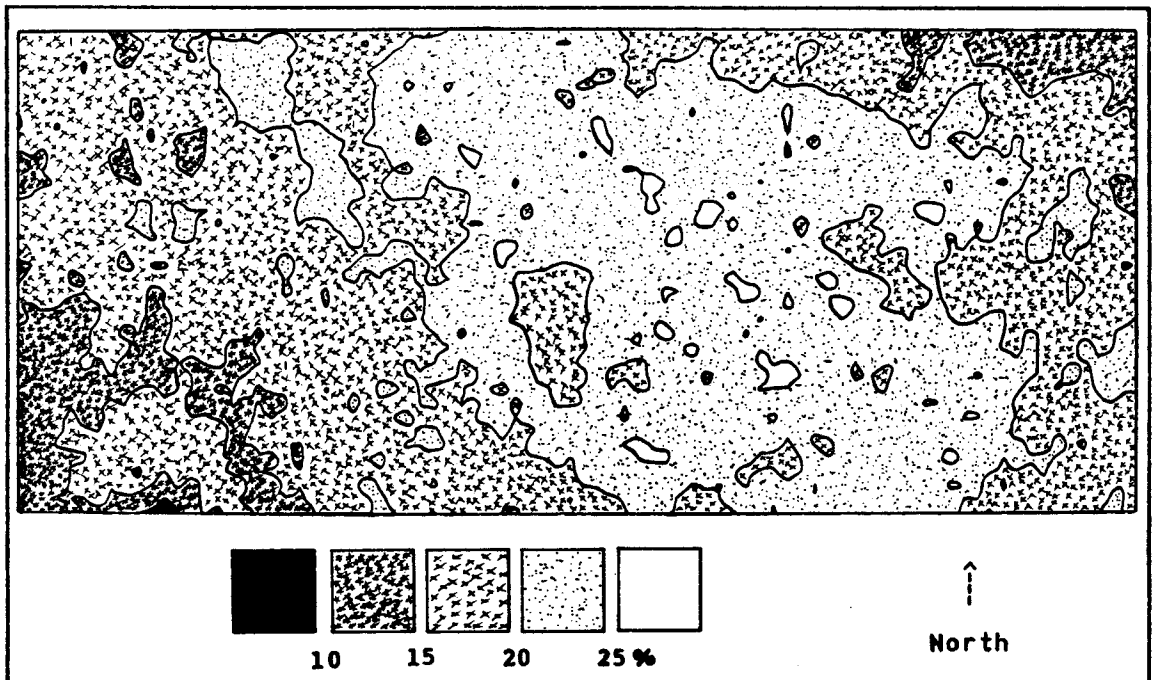


Figure 11: Contour map of porosity in a conditionally simulated sandstone aquifer. Each simulated value represents the average porosity for the total thickness of the aquifer at that point. Contour interval: 5% porosity. Sample semivariograms of this simulation are provided in Figure 12.

In geostatistics, we regard the set of porosities at all points x (actually a vector of coordinates, $\underline{x}$) as a "regionalized variable" z(x), which in turn is regarded as one particular realization of a random function Z(x), which obeys a spatial probability law that must be inferred from the data. In a simple case, we can consider this random function to be "second-order-stationary".[20] Second-order stationarity entails first that the mean (expected value) of the random function Z(x) is a constant for all locations x, i.e.:

$$EZ(x) = m(x) = m \quad \text{for all } x$$

This is a way of saying that there are no large-scale "trends" or "drifts" in porosity across the area of interest. Second, it entails that the covariance (or "autocovariance") function C(h), representing the covariance between Z(x) and Z(x+h) at two points separated by vector h, is also a constant for any given vector h, independent of the location x; i.e., for points x_1 and x_2 separated by vector h:

$$C(x_1,x_2) = E\{[Z(x_1)-m(x_1)][Z(x_2)-m(x_2)]\}$$

$$= E\{Z(x+h)Z(x)\}-m^2 = C(h) \quad \text{for all } x$$

The variogram[21] function $2\gamma(h)$ is defined as:

$$2\gamma(h) = E\{[Z(x+h)-Z(x)]^2\} = 2[C(0)-C(h)] \quad \text{for all } x$$

---

[20] Second-order stationarity differs from "strict" stationarity, which requires that the multivariate distribution of the set of random variables $\{Z(x_i)\}$ for any configuration of points $\{x_i\}$ be invariant under any translation of the set of points. Different authors unfortunately use somewhat different terminology in reference to stationarity. The terminology in common use in geostatistics is explained by Journel and Huijbregts (1978, Section II.A.2).

[21] In informal usage the semivariogram function $\gamma(h)$ is commonly called a "variogram"; however, an attempt has been made to maintain the more formal usage in this dissertation.

Under second-order stationarity, this is just the variance of the increments $Z(x+h)-Z(x)$, which have mean zero. If we have two spatially cross-correlated regionalized variables of interest, $z_1(x)$ and $z_2(x)$, there is an analogous cross-variogram function, representing the covariance of the increments $Z_1(x+h)-Z_1(x)$ and $Z_2(x+h)-Z_2(x)$:

$$2\gamma_{12}(h) = E\{[Z_1(x+h)-Z_1(x)][Z_2(x+h)-Z_2(x)]\} \quad \text{for all } x$$

The variogram is fundamental to all geostatistical calculations, so it is treated in detail in the next section.

3.1.3 Structural Analysis of Regionalized Data

In a real situation we do not know the random function $Z(x)$, so we must be content with estimates of its properties. Fortunately, for most geostatistical calculations (those required in the application of "linear" geostatistics, which is used in this dissertation), we need only a good estimate of the semivariogram function $\gamma(h)$ for all directed distances h that are of interest to us. For a set of N data within an area of interest, the usual semivariogram estimator is[22]

$$\gamma^*(h) = [1/(2N(h))] \sum_{i=1}^{N(h)} [z(x_i+h)-z(x_i)]^2$$

where $N(h)$ is the number of pairs of data separated by the vector h. The corresponding cross-semivariogram estimator is

$$\gamma^*_{12}(h) = [1/(2N_{12}(h))] \sum_{i=1}^{N_{12}(h)} [z_1(x_i+h)-z_1(x_i)][z_2(x_i+h)-z_2(x_i)]$$

---

[22] Some other estimators are discussed in Part I of Verly et al. (1984).

where $N_{12}(h)$ is the number of pairs with observed values of both z_1 and z_2. In practice, distances and directions are grouped into classes for irregularly spaced data. "Sample semivariograms" $\gamma^*(h)$ of the contoured data in Figure 11 are shown in Figure 12. These semivariograms have been calculated in four directions from a set of 3024 simulated porosity data, which are located on a regular square grid covering the entire area. Because there is such a large number of ideally located points from which to calculate these semivariograms, the resulting plots look much smoother than most sample semivariograms calculated from real data sets.

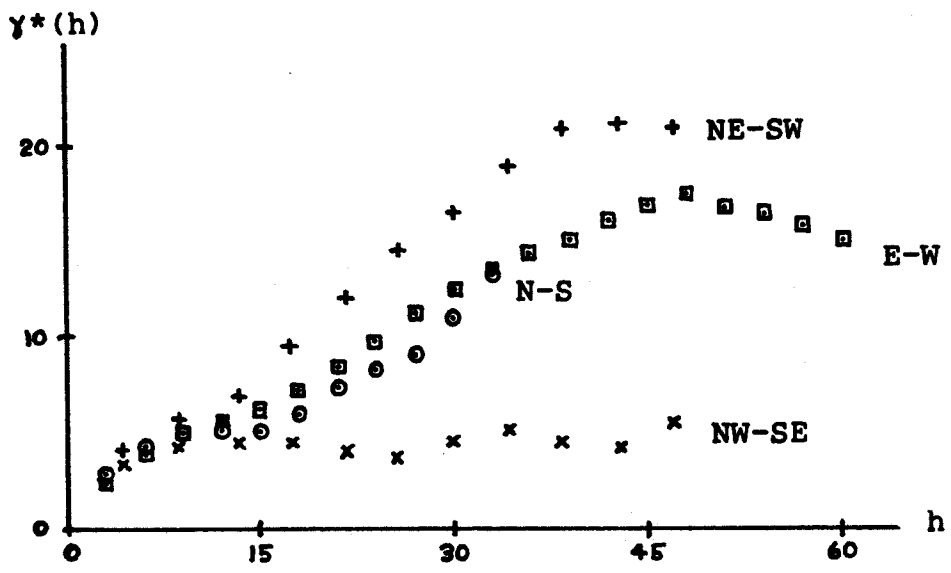


Figure 12: Exhaustive sample semivariograms for the porosity example in Figure 11.

Once we have obtained a sample semivariogram $\gamma^*(h)$, we need to fit a model semivariogram function $\gamma(h)$ to the sample points. In practice, this fitting is usually done by eye, although automatic procedures exist. We cannot fit just any function to such a plot: we must select only "conditionally positive-definite" (Journel and Huijbregts, 1978, p. 35) semivariogram models that will assure nonnegative variances for all finite linear combinations of random variables $Z(x_i)$ drawn from the random function $Z(x)$. Standard geostatistical texts list many such models, which can be used separately or in linear combinations ("nested structures") to fit any sample plot encountered in practice. Several of these models are illustrated in Figure 13. Two models that are especially useful for simulation purposes are the spherical model,

$$\gamma(r) = \tfrac{1}{2}C(3r/a - r^3/a^3) \quad \text{for } r \leq a$$
$$= C \quad \text{for } r \geq a$$

which is positive-definite in one, two, or three dimensions, and the less commonly known circular model,

$$\gamma(r) = [2C/(\pi a^2)][r\sqrt{(a^2-r^2)} + a^2\text{Arcsin}(r/a)] \quad \text{for } r \leq a$$
$$= C \quad \text{for } r \geq a$$

which is positive-definite in one or two dimensions. In the above expressions, "r" is the modulus of vector h: $r=|h|$. The parameter "a" is a distance parameter, the "range" of the semivariogram in a particular direction. Beyond this distance, values of the random function are uncorrelated. (This parameter can be regarded as a "range of influence" around a data point.) The maximum value "C" of the semivariogram is called its "sill" value; this value is achieved at distances beyond the range. If the random field has infinite spatial

extent, the scale parameter C corresponds to the variance of the random function.

Although the two models above reach their sill values at finite distances, others reach their sills only asymptotically. Models with sills are collectively called "transition" models. Still other models, corresponding to processes without finite variances, have no sill: the semivariogram continues to increase with increasing distance. For example, the Wiener-Levy (continuous random-walk) process has a linear semivariogram $\gamma(r)=Cr$, where C is just the semivariogram's slope, not a variance. "Hole-effect" semivariogram models do not increase monotonically with distance and may correspond to periodic or pseudoperiodic random functions. Finally, purely random processes with no correlation structure at any observable scale have "pure-nugget-effect" models; these models have flat semivariograms with constant values equal to the sill value at all distances greater than zero.

All semivariogram functions are necessarily equal to zero at their origins, i.e. at distance r=0. The shape of a semivariogram model near its origin is related to the short-scale continuity of the random function. A parabolic (concave-upward) behavior, in which $\gamma(r)\cong Ar^2$ as $r\to 0$, corresponds to very continuous (almost differentiable) spatial variability; for example, a one-dimensional deterministic function consisting of just a sloping line would have a semivariogram exactly equal to Ar^2, where A is the square of the line's slope. Semivariograms with linear behavior near the origin, $\gamma(r)\cong Ar$ as $r\to 0$, correspond to random functions that are mean-square continuous (i.e., $\lim \gamma(r)=0$ when $r\to 0$) but not differentiable (i.e., not "smooth"); good examples are

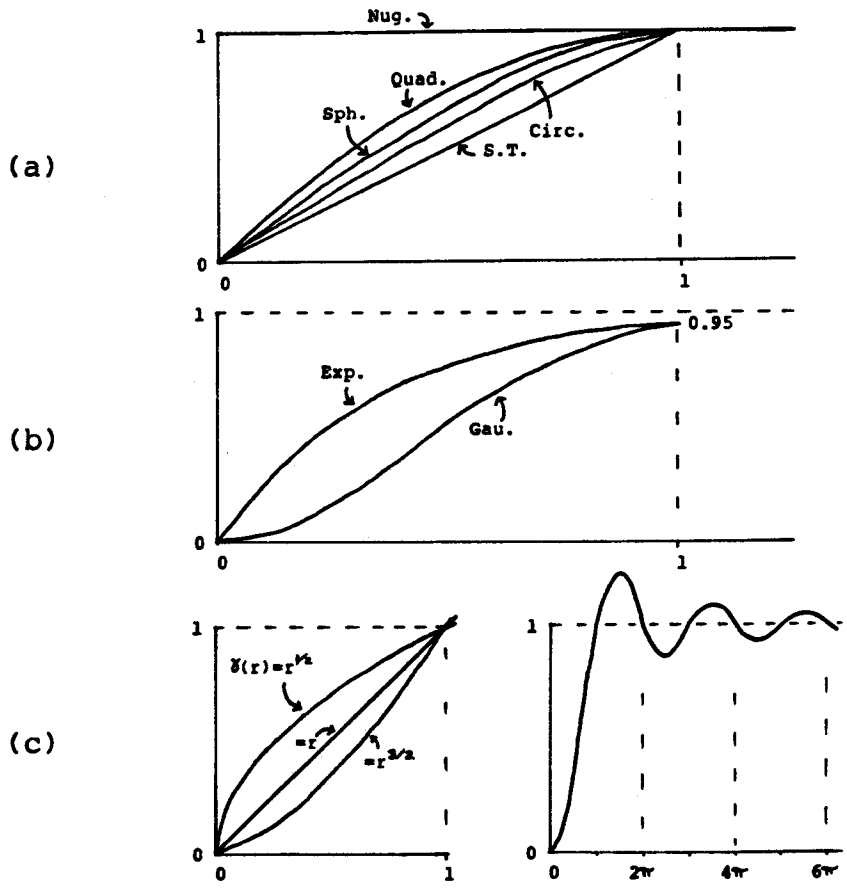


Figure 13: Graphs of several commonly used model semivariogram functions. Models with a sill (spherical, circular, simple-transitive, quadratic, exponential, gaussian, pure-nugget-effect) are scaled to sill=1. (a) Models with a definite range (spherical, circular, simple-transitive, quadratic, pure-nugget-effect) are scaled to range=1, except for pure-nugget-effect. (b) The "practical range" of the exponential and gaussian models is taken as the distance where the model reaches 95% of its sill value; this distance is scaled to 1 in the diagram. (c) Power models (including the linear model) and the 3-D hole-effect model can be useful if the data exhibit local drifts or pseudoperiodicities. Details on the quadratic model are provided by Alfaro (1984); the simple-transitive, circular, and spherical models are derived in Section 3.3.1.2 of this dissertation; the spherical model and the other models are described by Journel and Huijbregts (1978, Chapter III). Note that the spherical, circular, and simple-transitive models are positive-definite in ≤ 3, ≤ 2, and 1 dimension, respectively. The other models are positive-definite in ≤ 3 dimensions.

random functions with spherical and circular semivariograms, and the Wiener-Levy process. Finally, pure-nugget-effect models and nested models incorporating a discontinuity at the origin ($\gamma(0)=0$ but $\gamma(r)\neq 0$ for $r>0$) correspond to phenomena having a purely random ("white-noise") component with no spatial correlation at any distance beyond zero. Figure 14 compares time-series plots and sample semivariograms of realizations of white-noise, random-walk, and integrated random-walk processes to illustrate the differences among these types of short-distance semivariogram behavior.[23]

As illustrated by Figure 12, semivariograms of multi-dimensional phenomena need not be "isotropic": semivariograms may differ for different directions of vector h. The most obvious example of anisotropy among geologic phenomena is probably a bedded sedimentary deposit, in which one usually finds a shorter range and probably a larger variance (higher sill) in directions that cut across the bedding. Even within beds there may be a direction in which the rock is least variable, e.g. the direction parallel to a paleo-shoreline.

There are two types of anisotropy. If the anisotropy can be removed simply by a linear transformation of the coordinate system, the anisotropy is said to be "geometric". In this case the sill values in all directions must be equal, and the distances at which the semivariogram reaches a particular value (e.g. the range in the case of the sill value) must trace out an ellipse (in two dimensions) or

---

[23] In Section 3.3.2, we will see that models with parabolic behavior correspond to random functions in which most of the variability of the process is concentrated at low frequencies, whereas models with linear behavior correspond to random functions with a lot of high-frequency (localized) variability.

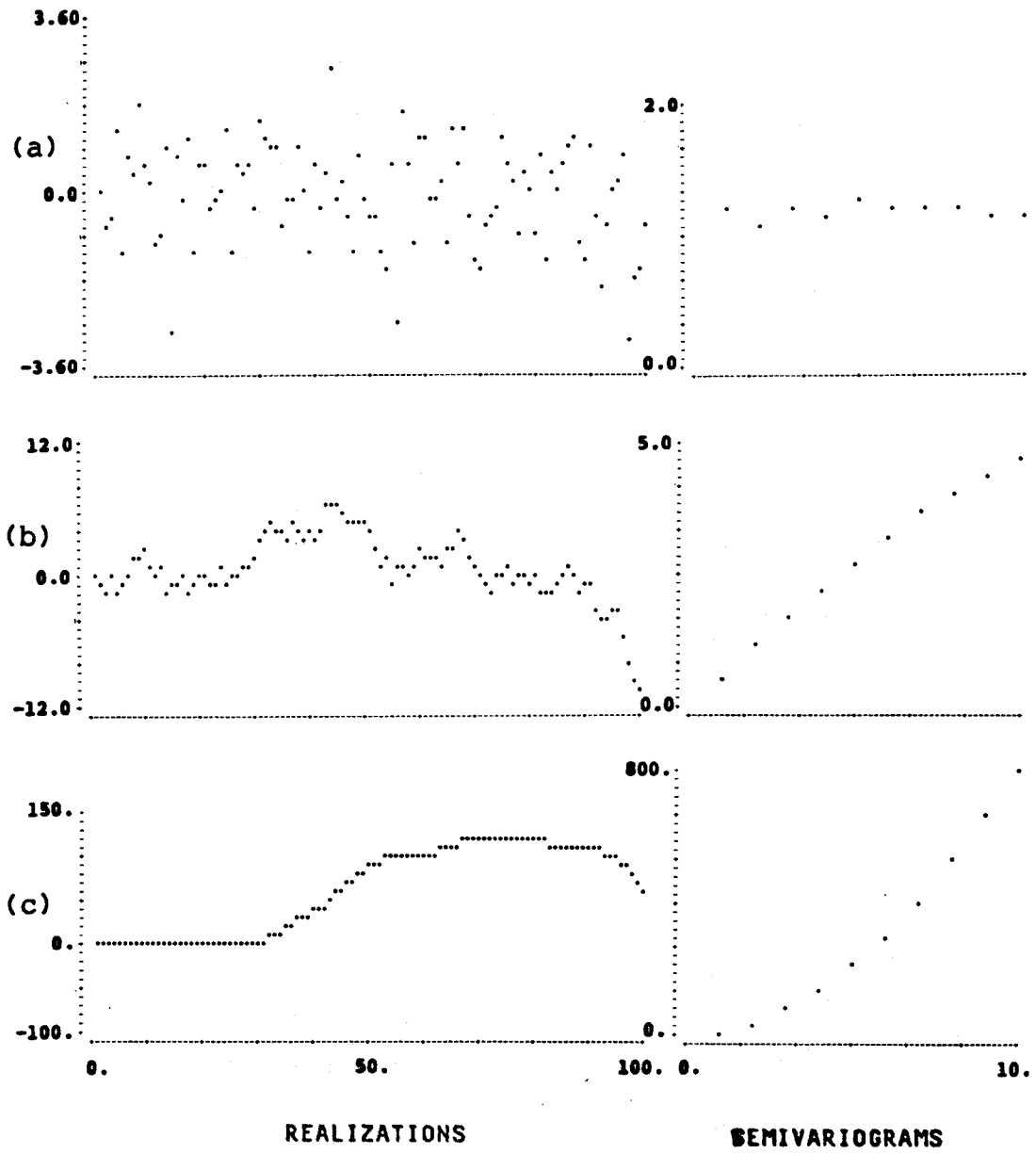


Figure 14: The behavior of some sample semivariogram functions at short distances. (a) Normally distributed white noise, illustrating a semivariogram with a discontinuity at the origin. (b) Integrated white noise (random walk), consisting of the cumulative sums of the process in (a), illustrating a semivariogram with linear behavior at the origin. (c) Integrated random walk, consisting of the cumulative sums of the process in (b), illustrating a semivariogram with parabolic behavior at the origin.

ellipsoid when plotted against all possible directions of vector h. The linear transformation of coordinates that would restore isotropy is that which would deform the ellipsoid back into a sphere. A three-dimensional semivariogram function showing geometric anisotropy can be represented in the form

$$\gamma(h) = \gamma'\{\sqrt{[(h'_u)^2+(h'_v)^2+(h'_w)^2]}\}$$

where h represents a vector with components (h_u,h_v,h_w) in the coordinate directions (u,v,w), and (h'_u,h'_v,h'_w) are linearly transformed components. Figure 15(a) illustrates a geometric anisotropy for a two-dimensional model.

Anisotropies that do not have this simple structure are said to be "zonal". Zonal anisotropies are modeled by systems of "nested structures" (linear combinations of several semivariogram models). Zonal anisotropies are very common in three-dimensional models, especially of bedded deposits. An example is provided in Figure 15(b).

3.1.4 Estimation Variance and Kriging

Returning to the porosity example of Figures 11 and 12, suppose we know the porosities z(x) at three points x_1, x_2, and x_3 within the area of interest, and we want to estimate (interpolate) the porosity at another point x_4 from these three data (see Figure 16). The estimate $z^*(x_4)$ will have to be some function of the three data. In terms of the underlying random function Z(x), we can define an "estimation variance" $\sigma^2_e=E\{[Z(x)-Z^*(x)]^2\}$ as a measure of the quality of this estimate. In linear geostatistics, we restrict ourselves to estimators that are linear combinations of the available data, i.e.:

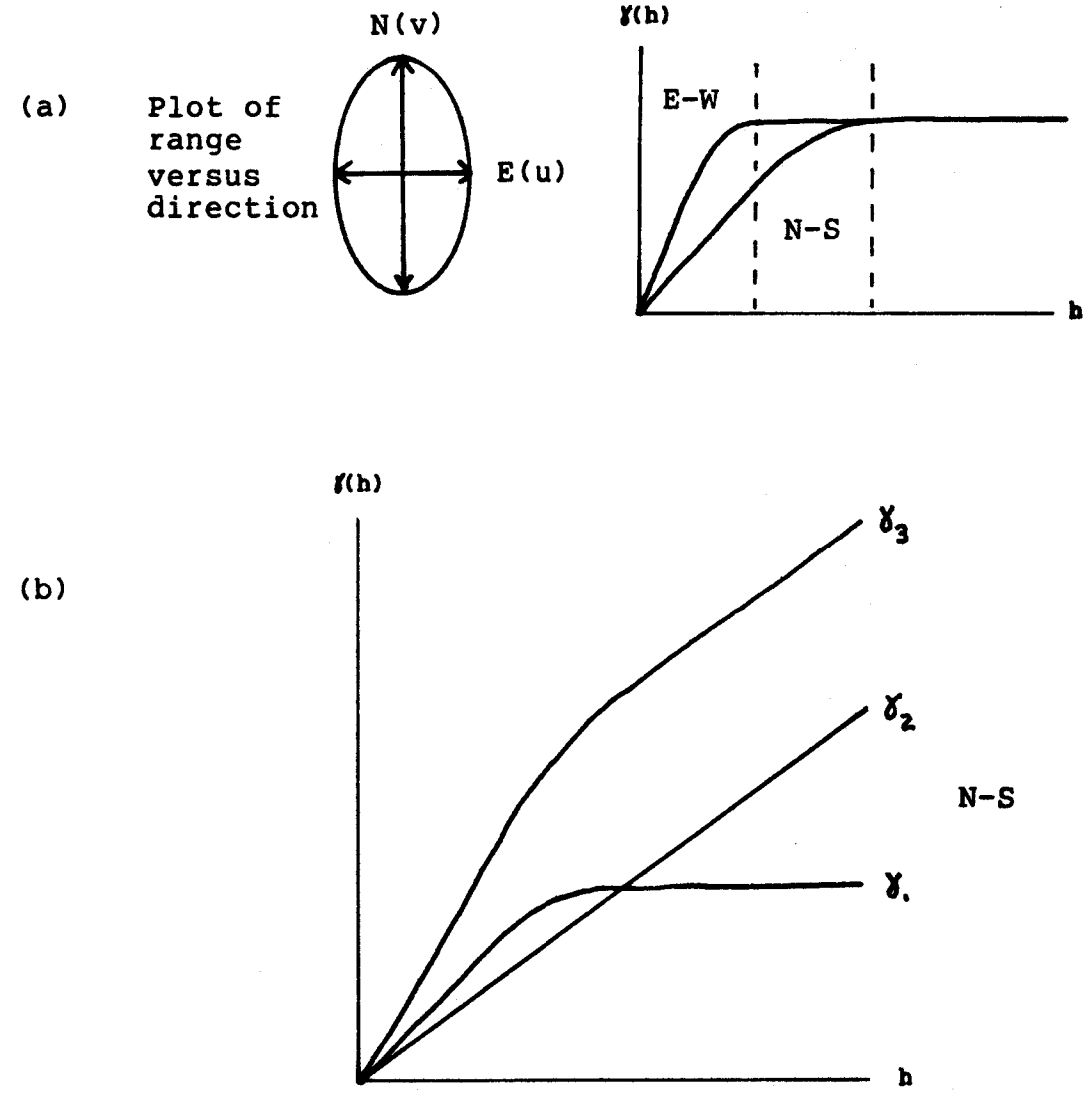


Figure 15: Models of anisotropy illustrated by a typical sedimentary deposit. (a) Geometric anisotropy within a bedding plane. The range is 6.0 in the north-south (v) direction, 3.0 in the east-west (u) direction. Isotropy can be restored by multiplying the north-south distance by $\frac{1}{2}$. Thus the model is:

$$\gamma(h_u, h_v) = \gamma'(\sqrt{(h^2_u + h^2_v/4)})$$

(b) Zonal anisotropy in three dimensions. The overall model is composed of a geometric-anisotropic structure for the horizontal direction (γ_1) nested with another structure (γ_2) representing the difference between the horizontal semivariogram (γ_1) and the more variable vertical semivariogram (γ_3). The final model is thus:

$$\gamma(h_u, h_v, h_w) = \gamma_1(\sqrt{(h^2_u + h^2_v/4 + h^2_w)}) + \gamma_2(h_w)$$

$$z^*(x) = \sum_{i=1}^{n} \lambda_i z(x_i) \quad \text{where } n = \text{the number of data}$$

If Z(x) is a second-order-stationary random function, this estimate will be unbiased if $\sum \lambda_i = 1$. But how should we allocate the weights λ_i among the data points x_i? To get a feel for the answer, let's start by comparing the data points two at a time.

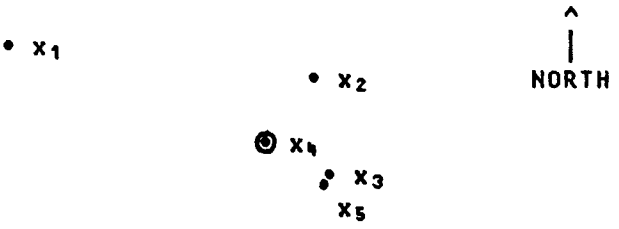


Figure 16: Data configuration for a simple linear estimation problem. We wish to estimate an unknown value of the regionalized variable z(x) at point x_4. Two cases are considered: (a) the available data are at x_1, x_2, and x_3, and (b) there is an additional datum at x_5.

If we were given only x_1 and x_3, it would seem to make sense that the value at x_3, which is closer to the unknown point x_4 than is x_1, should receive more weight than the value at x_1, <u>provided</u> that the distance $|x_3 - x_4|$ is within the range of the variogram in the northwest-southeast direction -- otherwise both data would be uncorrelated with the value at x_4, so both would be equally good estimators. Given only x_3 and x_2, it would also make sense that x_3 should receive more weight than x_2, even though both are equidistant from x_4; this is because the regionalized

variable z(x) is less variable in the northwest-southeast (x_4-x_3) direction than in the northeast-southwest (x_2-x_4) direction, so we naturally expect x_3 to be a better estimator than x_2. Given only x_1 and x_2, it is not so easy to decide which point should receive greater weight. Recall that the variogram $2\gamma(h)$ is the expected value of the squared difference between the values at two points x and x+h; it is thus equivalent to the estimation variance when estimating z(x) by the single value z(x+h). As the regionalized variable z(x) is least variable in the northwest-southeast direction, the variogram for a given distance will be lowest in that direction. Thus it is reasonable that, in order to minimize estimation variance, we should decide on the relative weights of x_1 and x_2 by checking to see which distance, $|x_1-x_4|$ or $|x_2-x_4|$, corresponds to a lower variogram value in the corresponding direction.

Now in comparing the points two at a time, we have overlooked an important consideration: the relative weights of any two points may not be independent of the position of the third point. To see this more clearly, suppose that point x_3 has a companion point, x_5, a very short distance away (Figure 16). Now it is not so obvious that x_3 should receive more weight than, for instance, x_1. Because x_3 and x_5 contribute roughly the same information, x_3's weight should be only about half of what it was before the advent of x_5, and might thus be less than x_1's weight.

Obviously the task of assigning an intelligent set of weights to a linear estimator z*(x) may not be trivial, especially in a practical case where the number of data is large and their spatial configuration

is very complicated. In such a situation, we need a way to calculate an "optimal" set of weights, i.e. a set of weights that will result in a minimum estimation variance.

To approach this problem, let's consider the general case where we want to determine the average value $z_v(x)$ of a regionalized variable $z(x)$ over some volume v, centered at x, from a known average value $z_u(x')$ in some other volume u, centered at x'. Each volume, v and u, can be composed of a single point, a compact volume, or a set of several volumes and/or points. In terms of Z(x), the estimation variance is

$$\sigma^2_e = E\{[Z_v(x)-Z_u(x')]^2\}$$

$$= E\{[(1/v)\int_{v(x)} Z(y)dy - (1/u)\int_{u(x')} Z(y)dy]^2\}$$

$$= (1/v^2)\int_{v(x)}\int_{v(x)} Z(y)Z(y')dy'dy$$

$$- (2/(vu))\int_{v(x)}\int_{u(x')} Z(y)Z(y')dy'dy$$

$$+ (1/u^2)\int_{u(x')}\int_{u(x')} Z(y)Z(y')dy'dy$$

$$= (1/v^2)\int_{v(x)}\int_{v(x)} C(y-y')dy'dy + m^2$$

$$- (2/(vu))\int_{v(x)}\int_{u(x')} C(y-y')dy'dy - 2m^2$$

$$+ (1/u^2)\int_{u(x')}\int_{u(x')} C(y-y')dy'dy + m^2$$

where C(y-y') is the covariance function at vector (y-y'). This can be expressed more compactly in the following way:

$$\sigma^2_e = \underline{C}(v,v) - 2\underline{C}(v,u) + \underline{C}(u,u)$$

where $\underline{C}(A,B)$ is the average value of the covariance function C(h) when the opposite extremities of vector h independently vary over the volumes

A and B.[24] Recalling that the semivariogram function $\gamma(h)$ is equal to $C(0)-C(h)$, we can conveniently rewrite the above expression simply by reversing the signs (and rearranging terms):

$$\sigma^2_e = 2\underline{\gamma}(v,u) - \underline{\gamma}(v,v) - \underline{\gamma}(u,u)$$

This is the fundamental catch-all formula for estimation variance. From the first term, we see that as the volumes v and u become more distant, the estimation variance becomes larger -- as we might expect. As the domain v to be estimated becomes larger, the variance decreases: it is easier to estimate an average over a large volume than over a small one. As the domain u of the data becomes larger, the variance also decreases: it is easier to estimate using a large, spatially dispersed set of data than using a small, compact set. Notice that the variogram function $2\gamma(h)$ may be regarded as an estimation variance for the trivial case in which the volumes v and u are points separated by vector h:

$$\sigma^2_e = 2\underline{\gamma}(v,u) - \underline{\gamma}(v,v) - \underline{\gamma}(u,u) = 2\gamma(v,u) - 0 - 0 = 2\gamma(h)$$

In practice the $\underline{\gamma}$ values can be calculated with the use of auxiliary functions (the approach used when calculating an estimation variance by hand), or by discretization of the volumes v and u into fine grids of points (for a numerical integration over the volumes -- the computerized approach). These alternatives are discussed by Journel and Huijbregts (1978, pp. 95-147).

Now back to the problem of minimizing the estimation variance by obtaining an optimal set of weights. Suppose we are estimating our unknown $z_v(x)$ by not one but several data sets $z_{ui}(x_i)$, i = 1 to n, with

---

[24] Journel and Huijbregts (1978) and most other references use an overbar instead of an underscore to indicate an average covariance or semivariogram. Overbars are not available in the character set used to print this dissertation.

associated weights λ_i, i = 1 to n. The estimation variance then becomes:

$$\sigma^2_e = E\{[Z_v - \sum_{i=1}^{n} \lambda_i Z_{u_i}]^2\}$$

$$= 2\sum_{i=1}^{n} \lambda_i \gamma(v,u_i) - \gamma(v,v) - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \gamma(u_i,u_j)$$

We want to minimize σ^2_e subject to the nonbias condition $\sum\lambda_i=1$. Notice that once we have calculated the relevant γ values from the semivariogram model, the expression for σ^2_e is just a quadratic equation with n unknowns, the λ_i values. To minimize this expression subject to the nonbias condition, we can use the Lagrange procedure found in many basic calculus texts, setting up n partial differential equations of the form:

$$\partial/\partial\lambda_i \quad \{\sigma^2_e - 2\mu[\sum\lambda_i - 1]\} = 0 \quad \text{for all } i = 1 \text{ to } n$$

The expression $[\sum\lambda_i - 1]$ above is equal to zero when the nonbias condition is satisfied, although after differentiation a nonzero term, -2μ, will appear at this location in each equation. The parameter μ then appears as an (n+1)th unknown, so we add the nonbias condition above as an (n+1)th equation, thus constraining the weights to obey this condition. Substituting the expression for σ^2_e and simplifying, we arrive at a system of n+1 linear equations:

$$\sum_{j=1}^{n} \lambda_j \gamma(u_i,u_j) + \mu = \gamma(u_i,v) \quad \text{for all } i = 1 \text{ to } n$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

Solving this "ordinary kriging system" provides us with the set of "kriging weights" $\{\lambda_j, j = 1 \text{ to } n\}$ that results in a minimum estimation variance, σ^2_k, which is called the "kriging variance".

There are other types of linear kriging systems. "Cokriging" systems, taking advantage of cross correlations among different regionalized variables, can be obtained for estimating values of one regionalized variable from data on the same variable and other "coregionalized" variables (Journel and Huijbregts, 1978, p. 324). Cokriging is especially helpful if the variable being estimated has been "undersampled" in comparison to other variables. "Drift" functions can also be incorporated, resulting in "universal kriging", applicable to some nonstationary cases (Journel and Huijbregts, 1978, p. 313). Nonlinear geostatistics incorporates transformations of the data so that nonlinear estimates can be made (described in several papers in Verly et al., 1984). Finally, some research on optimization criteria other than minimum estimation variance has been initiated recently (Journel, 1984b).

3.1.5 Dispersion Variance

Expressions of the variability of regionalized phenomena must always make reference to the scale of measurement. The variability of thousand-ton blocks of ore should be less than the variability of core samples, and the variability of monthly composite chemical analyses of raw materials delivered to a processing plant should be less than that of hourly quality-control samples. The variogram provides the information we need to calculate the theoretical values of these

variances for any sample size (called "support" in geostatistical jargon) larger than the size used in calculating the variogram itself. In most mining situations, the variogram is calculated from exploratory drilling data, defined on "core support", whereas the variability that is of interest for mine-planning purposes is usually the variability on "block support".

Consider a large block of material of volume (support) "v" (Figure 17) within a much larger ore deposit of volume "w". This block may be divided up into N smaller blocks of volume "u". If we measure the average value of a regionalized variable z(x) within each of the small blocks $u(x_i)$ centered at x_i,

$$z_u(x_i) = (1/u)\int_{u(x_i)} z(y)dy$$

and also the average within the larger block v,

$$z_v(x) = (1/v)\int_{v(x)} z(y)dy = (1/N)\sum_{i=1}^{N} z_u(x_i)$$

then the "dispersion variance" $D^2(u/v)$ of small blocks of size u within large blocks of size v is defined as

$$D^2(u/v) = E\{(1/N)\sum_{i=1}^{N}[Z_v(x)-Z_u(x_i)]^2\}$$

which, under second-order stationarity, is a constant for fixed configurations of u and v, independent of the location of v. Notice that actually more than just the <u>volumes</u> of u and v must be specified here: the <u>shapes</u> and <u>orientations</u> of the blocks are also important. Volumes with elongated shapes usually encompass more different types of material (and thus accomplish more smoothing of the variability of the material) than blocks with compact shapes. Furthermore, the

regionalized phenomenon may be anisotropic and thus more variable (and more susceptible to averaging) in one direction than in another.

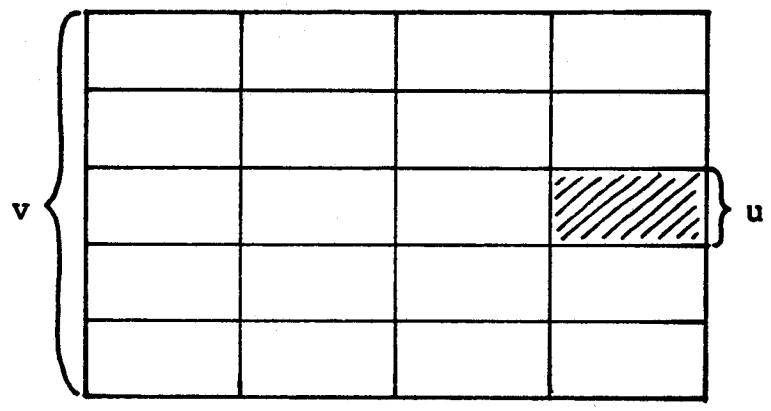


Figure 17: A large block of ore of volume v divided into smaller blocks of volume u.

We need not consider only volumes v divided into an integral number N of units u_i. A more general expression for the dispersion variance is

$$D^2(u/v) = (1/v)\int_{v(x)} E\{[Z_v(x)-Z_u(y)]^2\}dy$$

which is the mean value over v of the estimation variance of $Z_v(x)$ by $Z_u(y)$. Substituting the expression for estimation variance into this formula, we have:

$$D^2(u/v) = (1/v)\int_{v(x)} [\underline{C}(v(x),v(x))-2\underline{C}(v(x),u(y))+\underline{C}(u(y),u(y))]dy$$

Now recall that the average covariances $\underline{C}(v(x),v(x))$ and $\underline{C}(u(y),u(y))$ depend on the internal geometries of the supports v and u but not on their respective locations x and y in the above expression. Thus the average values of the average covariances over volume v(x) are just $\underline{C}(v,v)$ and $\underline{C}(u,u)$. The remaining term, $\underline{C}(v(x),u(y))$, when averaged over the volume v(x), just becomes $\underline{C}(v(x),v(x))$, or simply $\underline{C}(v,v)$. Thus the expression for dispersion variance becomes

$$D^2(u/v) = \underline{C}(v,v)-2\underline{C}(v,v)+\underline{C}(u,u)$$

$$= \underline{C}(u,u)-\underline{C}(v,v)$$

or, in terms of the semivariogram:

$$D^2(u/v) = \underline{\gamma}(v,v) - \underline{\gamma}(u,u)$$

An important implication of this simple relationship is that dispersion variances calculated for progressively larger supports are additive, e.g.:

$$D^2(u/w) = D^2(u/v) + D^2(v/w) \quad \text{for } u < v < w$$

Now consider a random function Z(x) measured on point support and having finite variance $\mathrm{Var}\{Z(x)\}=E\{[Z(x)-m]^2\}=C(0)$. This variance may be regarded as the dispersion variance $D^2(0/\infty)$ of point-support volumes within an infinite volume. Thus:

$$D^2(0/\infty) = \underline{C}(0,0) - \underline{C}(\infty,\infty) = C(0) - 0 = C(0)$$

$$= \underline{\gamma}(\infty,\infty) - \underline{\gamma}(0,0) = \gamma(\infty) - 0 = \gamma(\infty)$$

So the semivariogram sill has the equivalent representations:

$$D^2(0/\infty) = C(0) = \gamma(\infty) = \mathrm{Var}\{Z(x)\}$$

The concept of dispersion variance can be extended to describe covariances between the grades of different coregionalized variables measured on block support. Matheron (1965, pp. 146-148) demonstrates that the "dispersion covariance",

$$D^2_{\alpha\beta}(u/v) = E\{(1/N)\sum_{i=1}^{N}[Z_{v\alpha}(x)-Z_{u\alpha}(x_i)][Z_{v\beta}(x)-Z_{u\beta}(x_i)]\}$$

for random functions $Z_\alpha(x)$ and $Z_\beta(x)$ measured on support u inside volume v can be simply calculated by

$$D^2_{\alpha\beta}(u/v) = \gamma_{\alpha\beta}(v,v) - \gamma_{\alpha\beta}(u,u)$$

where $\gamma_{\alpha\beta}(h)$ is the cross semivariogram between $Z_\alpha(x)$ and $Z_\beta(x)$. The important additivity property still holds:

$$D^2_{\alpha\beta}(u/w) = D^2_{\alpha\beta}(u/v) + D^2_{\alpha\beta}(v/w) \quad \text{for } u < v < w$$

3.1.6 Regularization

Variograms are not necessarily expressed on point support. Point-support data exist only in theory anyway. In practice, core-support data can usually be considered as "practically" point-support data, because (except perhaps in the down-hole direction) the dimensions of the core samples are very small in comparison to the dimensions of mining blocks, so that $D^2(\text{core/block}) \simeq D^2(0/\text{block})$. But sometimes we have data on larger supports, such as block averages estimated from a large volume of combined blast-hole samples, or from crushed ore sampled in a processing plant. We can then define a "regularized" variogram,

$$2\gamma_u(h) = E\{[Z_u(x+h)-Z_u(x)]^2\}$$

where x and x+h are block centers separated by vector h, and u denotes the nonpoint support. If we formulate this variogram as an estimation variance, we have:

$$2\gamma_u(h) = 2\bar{\gamma}(u(x),u(x+h))-\bar{\gamma}(u(x),u(x))-\bar{\gamma}(u(x+h),u(x+h))$$

Under stationarity, the last two terms are equal, thus:

$$\gamma_u(h) = \bar{\gamma}(u(x),u(x+h))-\bar{\gamma}(u,u)$$

The first term, $\gamma(u(x),u(x+h))$, can be approximated by $\gamma(h)$ if the distance $|h|$ is very large in comparison to the block size. This is certainly true as $|h|$ approaches infinity; thus we see that the sill $\gamma_u(\infty)$ of the regularized semivariogram is always lower than the sill of the point semivariogram by an amount equal to $\gamma(u,u)$, so that:

$$D^2(u/\infty) = C_u(0) = \gamma_u(\infty) = \mathrm{Var}\{Z_u(x)\} < \mathrm{Var}\{Z(x)\}$$

The regularized range (for semivariogram models having ranges), which is the distance at which the values of $Z_u(x)$ in two volumes u(x) and u(x+h) are uncorrelated, must be increased by the width of the volume u in the direction of vector h.

Figure 18 compares the characteristics of sample semivariograms $\gamma^*(h)$ and $\gamma^*_u(h)$ for a typical transition phenomenon with finite range. Notice that, in addition to a decrease in the sill and an increase in the range, there is a tendency toward parabolic behavior near the origin of the regularized curve. This short-scale smoothness of the regularized phenomenon is the result of overlap between volumes u(x) and u(x+h) at distances $|h|$ less than the dimension of u.[25] As the size of volume u increases, all three of these effects become more pronounced. If the volume u is not isotropic (a circle in two dimensions, a sphere in three), the regularized variogram will not be isotropic even if the point-support process is isotropic.

In the absence of measurement errors, a sample semivariogram of a regularized phenomenon should show no discontinuity at the origin (no nugget effect). In practice, such a discontinuity might still be

---

[25] Parabolic behavior arising from large support is rarely observed clearly in sample semivariograms. Samples do not usually overlap in practice, so sample semivariogram points are usually farther apart than the dimension of the samples.

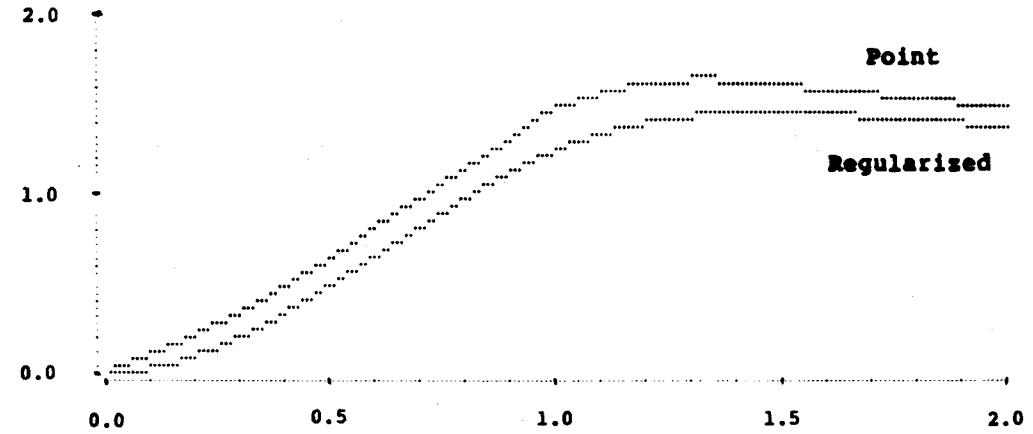


Figure 18: Comparison of sample semivariogram functions for point and regularized phenomena. The example is a point-support time series having a simple-transitive semivariogram with range=1 and sill=1. The time series is regularized by averaging the data within a moving line segment of length u=½.

observed if sampling, preparation, or analytical errors are responsible for an appreciable part of the observed variability of the phenomenon.

3.1.7 Simulation and Conditioning

The two variances discussed above -- estimation variance and dispersion variance -- are linked to the two basic types of problems that can be solved by linear geostatistical methods. If we are interested in estimating the value of a regionalized variable at some point where we have no data, or its average over some volume where we have missing or incomplete data, then we are interested in estimation variance and its minimization by kriging. If we are more interested in the variability of a phenomenon, and the relationship of that variability to the support (volume, shape, and orientation) of the measurement units, then we are interested in dispersion variance.

There are some practical applications in which dispersion variance may not give us all of the information we need about variability. For example, if we are designing a mineral-processing plant at a new mine site, we would like to know in advance the variability of the ore to be processed. But we really need more than just a few numbers representing the variances of small blocks of ore within large ones, large ones within still larger ones, and so forth. The raw-material input to this processing plant will be the output of a complex mining system operating on a regionalized variable (the ore grades in place in the mine). In effect, the mining system is a very complicated transformation, having no neat mathematical form, which converts a three-dimensional realization of ore grades in place into a one-dimensional realization: a time series of ore grades delivered to the plant. The more we know about this time series, the better job we can do in designing the plant. Dispersion variance alone can give us only, even under ideal conditions, the long-term variance (i.e., the semivariogram sill) of this time series. To characterize the time series completely, we need both a model of the mining system and a model of the ore grades in place. The ore-grades model should reflect not just the variability of the ore grades at different scales, but their whole spatial distribution.

The ideal model for ore grades in place would be a complete realization of the true regionalized variable $z_0(x)$ at all points x within the region of interest. But in practice we have only a finite data set $\{z_0(x_i), i = 1 \text{ to } N\}$ within this region. Infinitely many different realizations of the random function Z(x), all passing through the N available data points, and all having the same histogram and

variogram as the real $z_0(x)$, are possible. Only one of these realizations is the real $z_0(x)$.

From the available data we can generate kriged estimates $z^*_{0k}(x)$ at all points x within the region of interest. Although these estimates are optimal in the sense that the estimation variance at each point is minimized, the complete set of estimated values $\{z^*_{0k}(x)\}$ over the whole region of interest will have a variance smaller than the observed variance of the data. It is easy to see why this "smoothing effect" (Journel and Huijbregts, 1978, p. 450) occurs: the kriged estimates are just weighted averages of the data, and the averaging process filters out much of the variance observed in the original values.

Now suppose that we apply a computerized model of the mining process to the kriged estimates of the ore grades. Because the kriged estimates are so smooth, the output of this mining model would be a time-series input of ore grades to the processing plant that also would be very smooth -- creating an overly optimistic impression of the quality-control problems that we would actually experience in the real plant. To obtain a more realistic impression of these problems, we need a realization of the random function Z(x) of ore grades in place that (1) equals the observed data values at points where the real regionalized variable $z_0(x)$ has been sampled, and (2) has the same spatial distribution as the real $z_0(x)$, or at least the same mean and variogram. In geostatistical jargon, such a realization is called a "conditional simulation" of $z_0(x)$ -- "conditional" because it is constrained to pass through the observed data values, and "simulation" because, as in other types of simulations, this realization is drawn

randomly from an infinite population of possible realizations having the same statistical characteristics as the data.[26] This is an important point: a conditional simulation, unlike a kriged estimate, is not unique. Given a single finite set of conditioning data and the associated statistics (at least a mean and semivariogram model), an infinite number of conditional simulations can be produced.

There are many applications for conditional simulation in addition to mine and process planning (discussed more fully in Section 3.9). Several of these applications are described in Journel (1980), in Journel and Huijbregts (1978, Section VII.B, particularly VII.B.3), and in Verly et al. (1984, particularly Part XII).

To get an idea of how a conditional simulation is created in practice, let's look at a simple one-dimensional case -- a published time series of $CaSO_4$ percentages of an anhydrite-shale mixture delivered to a stockpile (Schofield, 1980, pp. 95 and 259). Figure 19(a) shows a plot of the normal scores[27] of part of the original data, point-sampled at fifteen-minute intervals. Figure 20(a) shows the semivariogram of the normal scores. Now suppose, for some mineral-processing study, we would like to fill in some additional data at five-minute intervals.

We begin by obtaining kriged values at the intervening five-minute positions, shown also in Figure 19(a).[28] Next we generate an

---

[26] In conformance to popular (but sloppy) usage, the term "simulation" is used throughout this dissertation to denote both the simulated realization and the act and methods of creating it. Some authors call the realization a "numerical model".

[27] The reason for using normal scores rather than raw data will become clear in Section 3.5.

[28] Some risk is being taken in this example. There are no data at five-minute intervals with which to estimate the semivariogram at five-

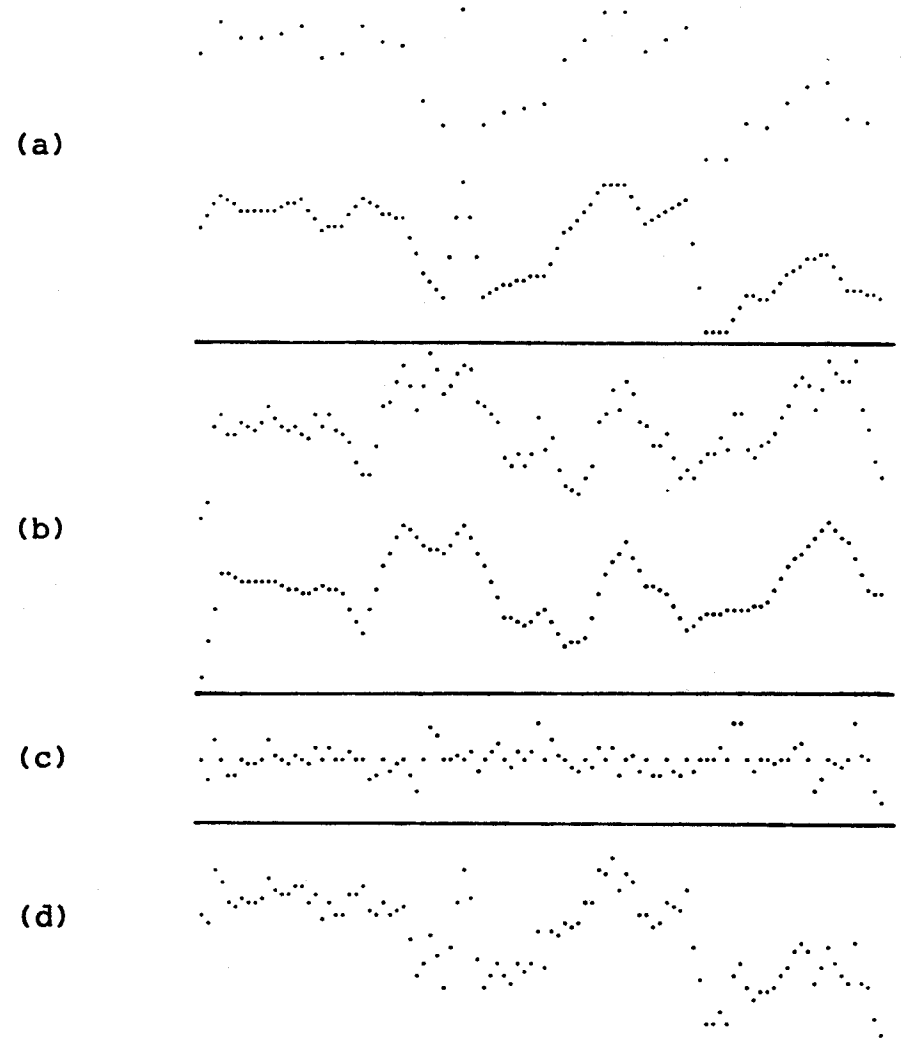


Figure 19: Steps in the conditional simulation of a time series of $CaSO_4$ analyses. (a) Normal scores of original data at fifteen-minute intervals, from Schofield (1980, pp. 95 and 259), with kriged estimates (lower plot) at five-minute intervals. (b) Unconditional simulation at five-minute intervals, with kriged values (lower plot) at five-minute intervals using simulated data at fifteen-minute intervals. (c) Kriging errors obtained by subtracting the kriged estimates in (b) from the simulated data. (d) Conditional simulation obtained by adding the kriging errors in (c) to the kriged estimates from the original data in (a).

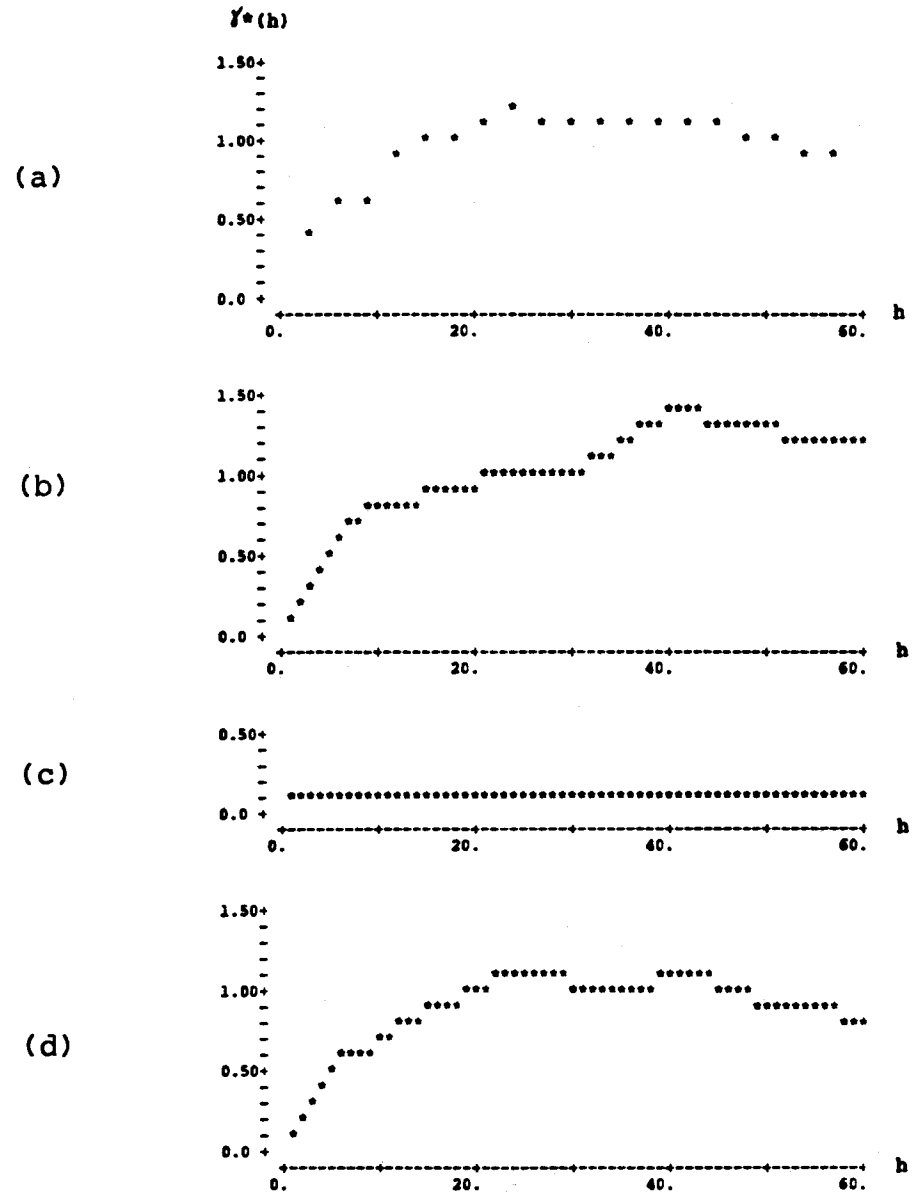


Figure 20: Sample semivariograms of the time series depicted in Figure 19. The lags on the horizontal axes are scaled to five-minute units. (a) Original data; (b) unconditional simulation; (c) kriging errors; (d) conditional simulation. The kriging errors do not generally display a pure-nugget-effect structure as in this example; the apparent lack of fluctuations in (c) is due to the high density of conditioning data and the coarse vertical scale of the plot.

"unconditional simulation" of the process at five-minute intervals, shown in Figure 19(b), using the moving-average technique explained in Section 3.3.1.2.[29] These simulated values have essentially the same semivariogram as the model semivariogram fitted to the original sample plot (compare Figures 20(a) and (b)), but the simulated values are not constrained to equal the data values at the fifteen-minute data locations -- hence the simulation is "unconditional". Next we select from the simulated series the subset of data located at points corresponding to the real fifteen-minute data locations. We use this smaller simulated data set to produce kriged estimates at the intervening five-minute locations, shown also in Figure 19(b). Next we subtract the kriged values of Figure 19(b) from the unconditionally simulated values, to produce a set of simulated residuals, or kriging errors, shown in Figure 19(c). Finally, we add these residuals to the kriged values we obtained with the original data (Figure 19(a)), producing a "conditional simulation", depicted in Figure 19(d).

Two important claims can be made about this conditional simulation: (1) its variogram is equal to the variogram of the original data (compare Figures 20(a) and (d)), and (2) the simulated realization passes through all of the original data points (compare the fifteen-minute values in Figures 19(a) and (d)). These two claims further imply that the mean and variance of the conditionally simulated data are equal

---

minute and ten-minute lags, so the short-distance structure of the semivariogram has been extrapolated from sample points at longer distances. A shortage of data at close spacings is a common problem in semivariogram modeling.

[29] The sample semivariogram was fitted by two nested simple-transitive models, which are easy to simulate in one dimension.

to those of the original data. Unless the random function is gaussian (as explained in the introduction to Section 3.3), these claims do not imply that the whole spatial distribution of simulated values is the same. More complicated methods for better reproducing the distribution are provided in Section 3.5.

Now let's see why this method works.[30] It is easy to see qualitatively what is being done: we have created two "smooth" realizations (the two krigings), subtracted the unconditional smooth realization from its associated "rough" unconditional simulation, and then simply added these residuals back to the other smooth realization, in order to roughen it up a bit. To demonstrate that the two claims made above are correct, we first need to formalize the series of steps that were performed above. We started with a data set selected from a true realization $z_0(x)$, which is unknown at all but a few data points. At these same data points and at several additional points, we created a conditionally simulated realization $z_{sc}(x)$ by the following operation:

$$z_{sc}(x) = z^*_{0k}(x) + [z_s(x) - z^*_{sk}(x)]$$

where: $z_{sc}(x)$ is the conditionally simulated value at point x,

$z^*_{0k}(x)$ is the value kriged from the real data,

$z_s(x)$ is the unconditionally simulated value, and

$z^*_{sk}(x)$ is the value kriged from the selected unconditionally simulated data.

A nice property of kriged estimates is that they are "exact interpolators"; i.e., the kriged values at data points with known real values are the real values themselves. This is to be expected, because

---

[30] More extensive discussions can be found in Journel (1974a, 1974b) and Journel and Huijbregts (1978, Chapter VII).

kriged estimates are calculated to minimize squared estimation errors, and the estimation error of a known value obviously should be zero. So at a particular point x' where we have observed the real value of $z_0(x')$ and the corresponding value of $z_s(x')$, the relation above can be rewritten

$$z_{sc}(x') = z_0(x') + [z_s(x') - z_s(x')]$$

$$\text{or} \quad z_{sc}(x') = z_0(x')$$

demonstrating the second claim above.

To demonstrate the first claim, let's look at the semivariogram of the conditionally simulated points in terms of some underlying random functions:

$$\gamma_{sc}(h) = E\{[Z_{sc}(x)-Z_{sc}(x+h)]^2\}$$

$$= E\{[Z^*_{0k}(x)+(Z_s(x)-Z^*_{sk}(x)) - Z^*_{0k}(x+h)-(Z_s(x+h)-Z^*_{sk}(x+h))]^2\}$$

$$= E\{[Z^*_{0k}(x)-Z^*_{0k}(x+h)]^2\}$$

$$+ E\{[(Z_s(x)-Z^*_{sk}(x))-(Z_s(x+h)-Z^*_{sk}(x+h))]^2\}$$

$$+ 2E\{[Z^*_{0k}(x)-Z^*_{0k}(x+h)][(Z_s(x)-Z^*_{sk}(x))-(Z_s(x+h)-Z^*_{sk}(x+h))]\}$$

$$= \gamma_{0k}(h) + \gamma_{s-sk}(h) + 0$$

The third term in the above expression is zero, as the realizations $z_0(x)$ and $z_s(x)$, along with their associated krigings, are independent.

To see that this is equivalent to the semivariogram $\gamma_0(h)$ of the original realization $z_0(x)$, we rewrite the random function $Z_0(x)$ as:

$$Z_0(x) = Z^*_{0k}(x) + [Z_0(x) - Z^*_{0k}(x)]$$

Then the semivariogram of $Z_0(x)$ can be re-expressed:

$$\gamma_0(h) = E\{[Z_0(x)-Z_0(x+h)]^2\}$$

$$= E\{[Z^*_{0k}(x)+(Z_0(x)-Z^*_{0k}(x)) - Z^*_{0k}(x+h)-(Z_0(x+h)-Z^*_{0k}(x+h))]^2\}$$

$$= E\{[Z^*_{0k}(x)-Z^*_{0k}(x+h)]^2\}$$

$$+ E\{[(Z_0(x)-Z^*_{0k}(x))-(Z_0(x+h)-Z^*_{0k}(x+h))]^2\}$$

$$+ 2E\{[Z^*_{0k}(x)-Z^*_{0k}(x+h)][(Z_0(x)-Z^*_{0k}(x))-(Z_0(x+h)-Z^*_{0k}(x+h))]\}$$

$$= \gamma_{0k}(h) + \gamma_{0-0k}(h) + 0$$

In this case, to show that the third term is zero we use a proof in Journel and Huijbregts (1978, pp. 496-498), which shows that in ordinary kriging the kriging error $[Z_0(x)-Z^*_{0k}(x)]$ is orthogonal to differences between the kriged values $[Z^*_{0k}(y)-Z^*_{0k}(y')]$:

$$E\{[Z^*_{0k}(y)-Z^*_{0k}(y')]\cdot[Z_0(x)-Z^*_{0k}(x)]\} = 0 \quad \text{for all } x, y, y'$$

Thus both cross products involved in the third term above are zero. Finally, the semivariograms of the two kriging errors, $\gamma_{s-sk}(h)$ and $\gamma_{0-0k}(h)$ must be equivalent, as the spatial configuration of kriging data is the same in both cases, and the two realizations $z_s(x)$ and $z_0(x)$ (and their semivariograms $\gamma_s(h)$ and $\gamma_0(h)$) are of the same random function $Z_0(x)$,

$$\gamma_{sc}(h) = \gamma_{0k}(h) + \gamma_{s-sk}(h)$$

$$= \gamma_{0k}(h) + \gamma_{0-0k}(h) = \gamma_0(h)$$

demonstrating the first claim.

To recapitulate: a conditional simulation $z_{sc}(x)$ of a single regionalized variable $z_0(x)$ is nothing more than an ordinary kriging $z^*_{0k}(x)$ of that $z_0(x)$, perturbed by some artfully constructed artificial kriging errors, $[z_{sc}(x)-z^*_{sk}(x)]$. These kriging errors are designed to have the same structure as the true kriging errors, $[z_0(x)-z^*_{0k}(x)]$. In particular, the variance of the simulated error at any given point x' is the same as that of the actual unknown kriging error: it is the kriging variance σ^2_k described in Section 3.1.4. However, the simulated error

is statistically independent of the true error, and the errors of independently simulated values at x′ (using independent sets of random numbers) are independent of one another. Hence, the estimation variance of a true value $z_0(x')$ at location x′ by a conditionally simulated value $z_{sc}(x')$ is:

$$\mathrm{Var}\{[z_0(x')-z_{sc}(x')]\} = \mathrm{Var}\{[z_0(x')-z^*_{0k}(x')]+[z^*_{0k}(x')-z_{sc}(x')]\} = 2\sigma^2_k$$

If we were to generate a large number N of independent conditional simulations and average them, the estimation variance of their average would be:

$$\mathrm{Var}\{[z_0(x')-\sum_{n=1}^{N} z_{scn}(x')/N]\}$$

$$= \mathrm{Var}\{[z_0(x')-z^*_{0k}(x')] + \sum_{n=1}^{N}[z^*_{0k}(x')-z_{scn}(x')]/N\}$$

$$= \sigma^2_k + \sigma^2_k/N = \sigma^2_k(N+1)/N$$

As $N\to\infty$, this variance converges to the kriging variance σ^2_k. The second term above, σ^2_k/N, can be viewed as the estimation variance of the kriged value by the average of the simulated values. As the number N of simulations increases, this average converges to the kriged value. The overall properties of a realization composed of averaged conditional simulations correspondingly converges toward the smooth properties of the kriging. So here are two final remarks on what not to do with a conditional simulation:

(1) There is little point in treating a conditional simulation as a local estimator. A conditional simulation is a bad local estimator compared to a kriging.

(2) There is little point in averaging several conditional simulations together. The average is both a bad local estimator and a bad representation of the variability of the simulated phenomenon.

3.2 THE LINEAR MODEL OF COREGIONALIZATION AND ITS USE IN SIMULATION

In this dissertation, the principal technique being used is conditional simulation of coregionalization, which differs in practice from the procedure outlined in the previous section only in that the unconditional-simulation step must provide simulated realizations of more than one random variable, and those realizations must exhibit the same variograms and cross variograms as the original data. The conditioning can be done by cokriging to be rigorous, although in practice ordinary kriging usually will suffice. (More will be said about this in Section 3.7.5.)

Suppose we have a set of k second-order-stationary random functions $\{Z_i(x), i = 1 \text{ to } k\}$, with a positive-definite matrix of positive-definite direct and cross covariance functions $C_{ij}(h)$, of the form

$$C_{ij}(h) = E\{Z_i(x)Z_j(x+h)\} - m_i m_j$$

where m_i and m_j are the stationary expected values of any two of the k random functions, $Z_i(x)$ and $Z_j(x)$ (including the cases i=j). The requirement of positive definiteness of the matrix and covariance functions ensures that the variances of all finite linear combinations of the Z's will be positive.[31] To confirm the positive definiteness of a matrix of nested covariance functions, we split the matrix into additive

---

[31] We assume in this model that there is no "lag effect" such that $C_{ij}(h) \neq C_{ij}(-h)$. Journel and Huijbregts (1978, p. 173) explain how to deal with such a geologically remarkable situation.

matrices of component covariance structures with the same mathematical form and the same range. For any vector h, the matrix of $\gamma(h)$ values will be positive-definite if the associated matrix of sills (or other vertical scale parameters, if there are no sills) is positive-definite. Then the total matrix of covariances summed from the nested structures must also be positive-definite. A simple example is provided below.

The problem is to simulate, first unconditionally, then conditionally, a set of k coregionalized (spatially cross-correlated) variables with a covariance matrix equal to the matrix of covariance functions observed in the sample data. Clearly we cannot accomplish this by merely performing k independent conditional simulations, as described in the preceding section: although the conditioning would impart some cross structure to the simulated data, the simulated values lying far from any conditioning points probably would not be properly cross-correlated. We must instead find some trick allowing us to simulate these k coregionalized variables dependently. This trick is the linear model of coregionalization.

For simplicity, let's consider an example with only two coregionalized variables, $z_1(x)$ and $z_2(x)$. These have means m_1 and m_2, direct semivariograms $\gamma_1(h)$ and $\gamma_2(h)$, and cross semivariogram $\gamma_{12}(h)$. For further simplicity, assume that all of these semivariograms have the same simple nested form: a nugget constant plus a single isotropic spherical structure with a common range "a" for all three semivariograms, i.e.,

$$\gamma_1(h) = CO_1 + C1_1\ Sph_a(h)$$

$$\gamma_2(h) = CO_2 + C1_2\ Sph_a(h)$$

$$\gamma_{12}(h) = CO_{12} + C1_{12} Sph_a(h)$$

where C0 denotes nugget constants, C1 denotes the sills of the spherical models, $Sph_a(h)$ denotes a spherical semivariogram model with sill=1 and range=a, and the subscripts denote the direct and cross structures.

We can deal with the two sets of nested structures separately. For the matrix of covariance functions defined above to be positive-definite, the matrices of nugget constants C0 and spherical sills C1 must both be positive-definite. This implies that the matrix of total sills, C=C0+C1, also is positive-definite. One of several checks on positive definiteness is to check that the determinants of all of the upper-left submatrices are positive. In the case of the total sill values, this amounts to verifying that:

$$C_1 > 0 \quad \text{and} \quad \begin{vmatrix} C_1 & C_{12} \\ C_{12} & C_2 \end{vmatrix} > 0$$

This restriction on the semivariogram models (and their component nested structures) must be satisfied when fitting models to sample semivariogram plots. If the sample plots do not exhibit clear sill values, we must either fit nontransition models with a positive-definite covariance matrix or at least fit a positive-definite set of transition models that will still fit the sample plots for small values of h. Only the quality of the variogram fit near the origin is of great importance in a well conditioned simulation, as the larger-scale behavior of the simulation is imposed by conditioning (Section 3.7.5).

Now how do we obtain an unconditional simulation of coregionalization that will fit our positive-definite model for $z_1(x)$ and $z_2(x)$? Consider two new random functions, $Y_1(x)$ and $Y_2(x)$, which have mean zero and direct covariance functions $K_1(h)$ and $K_2(h)$, but are independent, so

$K_{12}(h)=0$ for all h. Suppose we form two linear combinations of these functions:

$$Z_1(x) = a_{11}Y_1(x) + a_{12}Y_2(x)$$

$$Z_2(x) = a_{21}Y_1(x) + a_{22}Y_2(x)$$

As the Y's are independent, the covariance functions C(h) of the Z's are as follows:

$$C_1(h) = a^2_{11}K_1(h) + a^2_{12}K_2(h)$$

$$C_2(h) = a^2_{21}K_1(h) + a^2_{22}K_2(h)$$

$$C_{12}(h) = a_{11}a_{21}K_1(h) + a_{12}a_{22}K_2(h)$$

Now we see a way by which we can unconditionally simulate the two coregionalized variables $z_1(x)$ and $z_2(x)$. We begin by simulating two independent regionalized variables $y_1(x)$ and $y_2(x)$; then we pick suitable coefficients a_{11}, a_{12}, a_{21}, and a_{22} such that the correct positive-definite covariance matrix,

$$\begin{bmatrix} C_1(h) & C_{12}(h) \\ C_{12}(h) & C_2(h) \end{bmatrix}$$

will be produced. If the variograms of the z's have nested structures, we simulate each structure independently,[32] verifying that the sill matrix of each set of like structures is positive-definite. The coregionalized z's for the different structures are then added up to get the final set of unconditionally simulated coregionalized variables. The choice of coefficients, a_{ij}, is not unique. For example, given the

---

[32] The construction of a nested semivariogram model implies that the hypothetical underlying random functions corresponding to the different nested structures are uncorrelated, and thus independent if the random function is gaussian (Section 3.3). Otherwise, the variances (sills) of the nested structures could not simply be added up to get the total variance.

three equations for the known C_1, C_2, and C_{12} above, we cannot solve uniquely for the four unknown coefficients. Therefore, we usually just set one of the coefficients to zero, thereby setting one of the Z's equal to one of the Y's, e.g.:

$$Z_1(x) = Y_1(x)$$

$$Z_2(x) = a_{21}Y_1(x) + a_{22}Y_2(x)$$

Then:

$$C_1(h) = K_1(h)$$

$$C_2(h) = a^2_{21}K_1(h) + a^2_{22}K_2(h)$$

$$C_{12}(h) = a_{21}K_1(h)$$

The above expressions of $Z_1(x)$ and $Z_2(x)$ in terms of the independent functions $Y_1(x)$ and $Y_2(x)$ are one possible "linear model" for the coregionalization of $Z_1(x)$ and $Z_2(x)$. Linear models of this type can also be invented for each set of independent structures involved in a matrix of nested semivariogram models. The restriction to positive definiteness of the sill matrices of all independent nested structures implies that the sill values of any pair of like structures must satisfy the relationship $C_{12} < \sqrt{(C_1C_2)}$ (the Schwarz inequality). This means that if a particular nested structure with sill C_{12} appears on the sample cross semivariogram $\gamma_{12}(h)$ of variables $z_1(x)$ and $z_2(x)$, then that structure must also appear on the direct semivariograms $\gamma_1(h)$ and $\gamma_2(h)$. However, a structure appearing on one or both of the direct semivariograms need not be present on the cross semivariogram.

If there are many coregionalized variables with complicated nested structures, our problem can become rather tedious. First we have to form a sill or covariance matrix for each set of like structures and verify that this matrix is positive-definite. It is necessary to

isolate structures with the same form (e.g., spherical) and the same range in a separate matrix. Then, given these sill matrices, we must find a set of coefficients which, applied to a set of independently simulated Y's with the same structure, will give us the required sills. The general form of the linear model, expressed in necessarily rather cumbersome notation, is outlined in pages 171-175 of Journel and Huijbregts (1978), and its application in conditional simulation is formulated on pages 516-517. In this dissertation, the formalism is avoided in favor of worked-out examples, which are provided in the cement-related case study in Section 4.1, and in the paragraphs below. Another example can be found in Journel and Isaaks (1985).

For a brief and very simple numerical example of the procedure, consider the coregionalization of porosity and horizontal (assumed isotropic) permeability[33] in the thin sandstone aquifer whose porosity is mapped in Figure 11. The sample semivariograms of the original data set suggested the following simple nugget+circular models:

For $z_1(x)$ = porosity, units $(\%)^2$:

$$\gamma_1(h) = 1 + \text{Circ}[\text{range}=60, \text{sill}=6]$$

For $z_2(x)$ = permeability, units $(\text{md})^2$:

$$\gamma_2(h) = 10000 + \text{Circ}[\text{range}=60, \text{sill}=4000]$$

For the cross semivariogram, units $(\%)(\text{md})$:

$$\gamma_{12}(h) = 70 + \text{Circ}[\text{range}=60, \text{sill}=80]$$

---

[33] Permeability usually is approximately lognormally distributed, so in most real applications the structural analysis, linear model, and simulations would actually be performed using logarithms of the permeability data instead of the untransformed permeabilities used here. In the present example the simulated permeabilities would look normal, as explained in the introduction to section 3.3. No transformation was performed for this example because the original (quite unusual) data appeared more normal than their logarithms.

The sill matrices are:

$$\begin{array}{c} \\ 1 \\ 2 \end{array}\begin{array}{c} \text{Nuggets} \\ \begin{array}{cc} 1 & 2 \end{array} \end{array}\begin{bmatrix} 1 & 70 \\ 70 & 10000 \end{bmatrix} + \begin{array}{c} \text{Circulars} \\ \begin{bmatrix} 6 & 80 \\ 80 & 4000 \end{bmatrix} \end{array} = \begin{array}{c} \text{Total} \\ \begin{bmatrix} 7 & 150 \\ 150 & 14000 \end{bmatrix} \end{array}$$

These matrices are positive-definite, as the Schwarz inequality is satisfied for each:

$$\text{Nuggets:} \quad 70 < \sqrt{(1 \cdot 10000)} = 100$$

$$\text{Circulars:} \quad 80 < \sqrt{(6 \cdot 4000)} = 154.92$$

$$\text{Total:} \quad 150 < \sqrt{(7 \cdot 14000)} = 313.05$$

The nested nugget and circular structures can be modeled independently. First the nugget variables, $ZO_1(x)$ and $ZO_2(x)$:

$$ZO_1(x) = aO_{11}YO_1(x) + aO_{12}YO_2(x)$$

$$ZO_2(x) = aO_{21}YO_1(x) + aO_{22}YO_2(x)$$

$$CO_1(h) = aO^2_{11}KO_1(h) + aO^2_{12}KO_2(h) = 1$$

$$CO_2(h) = aO^2_{21}KO_1(h) + aO^2_{22}KO_2(h) = 10000$$

$$CO_{12}(h) = aO_{11}aO_{21}KO_1(h) + aO_{12}aO_{22}KO_2(h) = 70$$

If we generate independent normal nugget random variables YO_1 and YO_2 with mean=0 and variance=KO_1=KO_2=1, the solution is straightforward. First, arbitrarily set one coefficient equal to zero: aO_{12}=0. Therefore aO_{11}=1. Now $aO_{11}aO_{21}$=70, so aO_{21}=70. As $aO^2_{21}+aO^2_{22}$=10000, then aO^2_{22}=10000-4900, or aO_{22}=71.41. Thus the linear model for the nugget structures is:

$$ZO_1(x) = (1)\ YO_1(x)$$

$$ZO_2(x) = (70)YO_1(x) + (71.41)YO_2(x)$$

Next, the circular variables, $Z1_1(x)$ and $Z1_2(x)$:

$$Z1_1(x) = a1_{11}Y1_1(x) + a1_{12}Y1_2(x)$$

$$Z1_2(x) = a1_{21}Y1_1(x) + a1_{22}Y1_2(x)$$

$$C1_1(h) = a1^2_{11}K1_1(h) + a1^2_{12}K1_2(h) = 6$$

$$C1_2(h) = a1^2_{21}K1_1(h) + a1^2_{22}K1_2(h) = 4000$$

$$C1_{12}(h) = a1_{11}a1_{21}K1_1(h) + a1_{12}a1_{22}K1_2(h) = 80$$

Now we generate independent circular realizations $Y1_1$ and $Y1_2$ with mean=0, variance=$K1_1$=$K1_2$=1, and range=60. We set $a1_{12}$=0, so ai_{11}=$\sqrt{6}$=2.45. Then $a1_{21}$=80/$\sqrt{6}$=32.65. As $a1^2_{21}+a1^2_{22}$=4000, then $a1^2_{22}$=4000-6400/6, or $a1_{22}$=54.16. So the linear model for the circular structures is:

$$Z1_1(x) = (2.45)\ Y1_1(x)$$

$$Z1_2(x) = (32.65)Y1_1(x) + (54.16)Y1_2(x)$$

Finally, to get the unconditional simulations of $Z_1(x)$ and $Z_2(x)$, we combine the independent nugget and circular structures. Their variograms simply add up, as the simulations of the nested structures are independent. This yields a final linear model of coregionalization for $Z_1(x)$ and $Z_2(x)$:

$Z_1(x)=Z0_1(x)+Z1_1(x)=(1)Y0_1(x)+(2.45)Y1_1(x)$

$Z_2(x)=Z0_2(x)+Z1_2(x)=(70)Y0_1(x)+(71.41)Y0_2(x)+(32.65)Y1_1(x)+(54.16)Y1_2(x)$

As a check on the total variances, we notice that the four directly simulated variables $Y0_1$, $Y0_2$, $Y1_1$, and $Y1_2$ are all independent, so:

$$Sill(Z_1) = 1^2+2.45^2 \simeq 7$$

$$Sill(Z_2) = 70^2+71.41^2+32.65^2+54.16^2 \simeq 14000$$

Similarly, the sill of the cross semivariogram is:

$$Sill(Z_1*Z_2) = (1)(70)+(2.45)(32.65) \simeq 150$$

Notice that these unconditionally simulated coregionalized variables have the proper direct and cross variograms, but their means are still zero. For a realistic unconditional simulation of porosity and permeability, we would need to add the sample mean values of porosity and permeability (regarded as estimates of the stationary expectations of Z_1 and Z_2) to all simulated values of the corresponding variables. This is unnecessary in the case of a conditional simulation, as conditioning with the real data values will impose the appropriate mean upon the simulated regionalization.

Notice also that, although the linear model used for the unconditional simulation of porosity and permeability involves only isotropic structures, the conditional simulation of porosity contoured in Figure 11 is clearly anisotropic. Actually this anisotropy is not evident at short distances. The overall anisotropy that is visible is entirely the result of conditioning with an anisotropic set of partly fictitious data[34], which imposed the observed large-scale anisotropy on the simulation.

3.3 UNCONDITIONAL SIMULATIONS: METHODS FOR GENERATING CORRELATED DATA

The production of a conditional simulation of coregionalization requires two types of input data: a set of coregionalized conditioning data, and several independently and unconditionally simulated sets of regionalized data, which are transformed into an unconditionally simulated coregionalization by means of a linear model. Ordinary kriging is

---

[34] The use of fictitious conditioning data and the importance of conditioning in general are discussed in Section 3.7.5.

usually employed for the conditioning. The organization of these steps will be described in more detail in Section 3.4.

The subject of this section is the methods used to produce the several independent unconditional simulations used as input to the linear model of coregionalization. A large literature has developed on this subject, and several remarkably diverse techniques have been proposed for unconditionally simulating correlated random data in one, two, and three dimensions. These methods fall into two main categories: space-domain (or time-domain, in one dimension) methods and frequency-domain (spectral) methods.

In space-domain simulations, the objective is to create a realization of a random function with a prespecified covariance structure. In frequency-domain simulations, the realization is characterized by a prespecified spectral density function. The covariance function and its Fourier transform, the spectral density function, convey the same information about the random process, but in different forms that are more or less natural in different applications. Because space-domain methods are easier to understand, generally just as fast, and probably more natural for most earth-science applications, the emphasis in this section will be placed on space-domain procedures. Section 3.7.6 summarizes the relative advantages and disadvantages of the methods and makes some general recommendations on their use.

Gaussian processes. All of the simulation methods to be discussed here produce approximate realizations of gaussian spatial random functions, commonly called simply "gaussian processes". A gaussian process (which has nothing to do with the gaussian semivariogram model in Figure 13) is defined as a random function Y(t) satisfying the

property that any finite set of random variables $\{Y(t_1),\ldots,Y(t_n)\}$ drawn from the random function will have a joint normal (gaussian) distribution (alternatively, any finite linear combination of the $Y(t_i)$'s is normally distributed). In a simulation, each simulated y(t) value is formed by summing up the contributions of several independent x(u) values, and such a sum is asymptotically normal by a central limit theorem as the number of terms in the sum becomes large.[35] Gaussian processes have some useful properties, derived from the properties of the multivariate normal distribution: the conditional expectation of an unknown y(t) is a linear function of the known conditioning values $\{y(t_i)\}$, hence a linear kriging estimate, which is the best unbiased linear estimate in terms of estimation variance, is also the best estimate overall; the mean and covariance function (mean vector and covariance matrix for a coregionalization) of a gaussian process characterize its spatial distribution completely; a zero cross-covariance function (at all h) between two gaussian processes indicates that the processes are completely independent, not just linearly uncorrelated; and a linear combination of gaussian processes, whether they are correlated or not, is also a gaussian process.

Unfortunately, real spatial phenomena do not always have the good taste to be gaussian. Adaptations of gaussian simulations to represent these phenomena are discussed in Section 3.5.

---

[35] The x(u) values can be drawn from any distribution with finite variance, but some distributions will result in slower convergence to normality for y(t). Values from the uniform distribution on the interval (0,1) can be generated most rapidly, but if a small number of x(u) values is being summed, it is safer to draw x(u) from a normal distribution to assure normality of y(t).

3.3.1 Space-Domain Approaches

3.3.1.1 Matrix Methods

The objective is to simulate a finite number N of values representing a realization of a stationary gaussian spatial process Y(t) at N grid points t_i, i = 1 to N. From the definition of a gaussian process, these values must be drawn from a multivariate normal distribution, which has a positive-definite symmetric NxN matrix C of covariances among the N random variables. Under second-order stationarity, each element c_{ij} of the covariance matrix is the covariance between the random variables $Y(t_i)$ and $Y(t_j)$, i.e. $c_{ij} = C(t_i - t_j) = C(h)$, where vector $h = (t_i - t_j)$. Thus one way to obtain an unconditional simulation of the process at N points is to use one of the standard methods for generating a multivariate normal random vector with a prespecified covariance matrix.

Factorization approach. A popular technique usually attributed to Scheuer and Stoller (1962) (although mentioned in passing by Hammersley and Nelder, 1955) is to factor the covariance matrix $C=AA'$ into the product of a matrix A and its transpose A'. A good choice of A is the Cholesky decomposition, described in Kennedy and Gentle (1980, p. 294), which produces a lower-triangular matrix A. The N-component random vector Y of simulated values can then be generated by

$$Y = A X + m$$

where X is an N-component vector of independent standard normal random numbers (mean=0, variance=1) and m is the desired vector of means. For a stationary gaussian process, all elements of m are equal. If conditioning is to be performed later, m can be set to zero. This approach and closely related ones are described in most basic simulation

and statistical-computing texts, e.g. Law and Kelton (1982, p. 269), Rubinstein (1981, p. 65), and Kennedy and Gentle (1980, p. 228). These texts also describe methods for generating the independent normal components of vector X.

Stepwise-conditional approach. Another, usually slower, approach relies on the expression of the multivariate distribution of Y as the product of nested conditional distributions. We begin by generating a realization $y(t_1)$ from the normal distribution with stationary mean=m and variance=C(0). Then we generate $y(t_2)$ from the conditional distribution of $Y(t_2)$ given $y(t_1)$, followed by generation of $y(t_3)$ given $y(t_2)$ and $y(t_1)$, and so forth. This approach is described by Kennedy and Gentle (1980, p. 229). Borgman (1982, p. 406) points out that it is possible to obtain a conditional simulation of a gaussian process directly by this sort of approach, using the well known expressions for the conditional mean ("regression function") and conditional covariance matrix (Anderson, 1958, p. 28, equations 5 and 6). Suppose we want to simulate a multivariate normal random vector Y, partitioned into two vectors Y_1 and Y_2. The mean vector and covariance matrix are similarly partitioned:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \quad m = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}' & C_{22} \end{bmatrix}$$

Now suppose that we are given a set of values $\{y_1\}$, drawn from Y_1, as conditioning data. We must obtain a realization of $Y_2|(Y_1=\{y_1\})$ by drawing a vector $Y_{2/1}$ from the multivariate normal distribution with mean and covariance:

$$m_{2/1} = m_2 + C_{12}'C_{11}^{-1}(y_1 - m_1)$$

$$C_{22/1} = C_{22} - C_{12}'C_{11}^{-1}C_{12}$$

Once the values for $m_{2/1}$ and $C_{22/1}$ are determined, any multivariate-normal generator can be employed to generate a realization from $Y_{2/1}$. Notice that the conditional mean vector $m_{2/1}$ is not necessarily zero even if the original vector m was taken to be zero.

Eigenvalue-eigenvector approach. Borgman (1982, p. 391) describes another multivariate-normal generator based on the eigenvector-eigenvalue relation

$$C\ B = B\ L$$

where B is a matrix whose columns are the eigenvectors of the covariance matrix C, and L is a diagonal matrix of eigenvalues of C.[36] The multinormal random vector Y is then expressed as

$$Y = B\ \sqrt{L}\ X + m$$

where $\sqrt{L}$ is the diagonal matrix of square roots of the eigenvalues. In the common situation where some of the $Y(t_i)$'s are highly correlated and some uncorrelated or only slightly correlated (i.e., when the variogram range is less than the size of the simulation domain) many of the eigenvalues will be zero or nearly so. If we partition the B and L matrices according to the sizes of the eigenvalues, i.e.,

$$L = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \text{ and } B = [B_1, B_2]$$

---

[36] This is the principal-components relation, discussed in more detail in Section 3.5.3.

where L_2 represents the near-zero eigenvalues, we then have a potentially faster way to generate Y:

$$Y = B_1 \sqrt{L_1}\, X_1 + B_2 \sqrt{L_2}\, X_2 + m$$

The second term in this expression for Y can be neglected in practice, because the elements of $\sqrt{L_2}$ are very small.

The factorization and eigenvalue formulations can be useful if many independent simulations at a fairly small number of grid points are to be performed. The matrix decomposition must be performed only once; thereafter the stored coefficients A or $B\sqrt{L}$ can be multiplied by any number of independent X vectors to get repeated unconditional simulations of the random function Y(t).

The big limitation of the matrix approach to simulation is the expense and impracticality of manipulating very large covariance matrices. For example, the triangular matrix A produced by the factorization method requires N(N+1)/2 storage locations[37] in the computer, unless time-consuming input-output operations are used. However, it is commonly possible to take advantage of special regularities in the covariance matrix to simplify the procedure. Borgman (1982, p. 396) points out that "the stationarity usually assumed for the time series [or spatial process] leads to the covariance function being only a function of time lag [or distance]. This behavior forces a regularity on the covariance matrix and reduces the whole concept of matrix multiplication to an equivalence with the filtering of white noise." In other words, we can regard the triangular matrix A as

---

[37] Some types of sparse matrices require fewer locations and are much easier and faster to manipulate. Matrix simulations taking advantage of such structures are described by Davis (1985a).

a re-expression of a moving-average weight function (filter) that transforms an uncorrelated input vector X (white noise) into a correlated output vector Y. A simple time-series example is shown in Figure 21, which depicts the matrix form of a simple moving-average process created by moving a constant-valued weight function over a time series of uncorrelated random numbers. Moving-average simulation methods are discussed in Section 3.3.1.2 in more detail.

Although impractical for very large simulations, matrix methods have been used to generate simulations of an "areal-average" (regularized) process for use in subsurface hydrology (Mantoglou and Wilson, 1981, p. 157). Matrix methods may be appropriate in such an application, in which many independent realizations, of only a few areal averages each, are needed for input to a fluid-flow simulation program. Borgman et al. (1984) also apply this approach to generate correlated data for use in a three-dimensional spectral simulation, described in Section 3.3.2.2.

3.3.1.2 Moving-Average Methods

Simulation by space-domain methods is greatly simplified if the random function can be modeled as a moving-average process. To obtain a moving-average process, we first define a random function X(u) at each point u in n-dimensional space R^n. X(u) is second-order-stationary with mean EX=0, covariance $C_X(h)$. We then define a new random function Y(t) at each point t in R^n as a "weighted average" of X(u):

$$Y(t) = \int_{R^n} f(t-u)X(u)du$$

The weight function f(t-u) applied to each value X(u) depends on the distance and direction (t-u) between point t and each location u. If we

$$[\,A\,] \bullet [\,X\,] = [\,Y\,]$$

$$\begin{bmatrix} 1 & & & & & & & & & \\ 1 & 1 & & & & & & & & \\ 1 & 1 & 1 & & & & & & & \\ & 1 & 1 & 1 & & & & & & \\ & & 1 & 1 & 1 & & & & & \\ & & & 1 & 1 & 1 & & & & \\ & & & & 1 & 1 & 1 & & & \\ & & & & & 1 & 1 & 1 & & \\ & & & & & & 1 & 1 & 1 & \\ & & & & & & & 1 & 1 & 1 \end{bmatrix} \bullet \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{bmatrix} = \begin{bmatrix} y_1 = x_1 \\ y_2 = x_1+x_2 \\ y_3 = x_1+x_2+x_3 \\ y_4 = x_2+x_3+x_4 \\ y_5 = x_3+x_4+x_5 \\ y_6 = x_4+x_5+x_6 \\ y_7 = x_5+x_6+x_7 \\ y_8 = x_6+x_7+x_8 \\ y_9 = x_7+x_8+x_9 \\ y_{10} = x_8+x_9+x_{10} \end{bmatrix}$$

$$[\,C\,] = [\,A\,] \bullet [\,A\,]'$$

$$[\,C\,] = \begin{bmatrix} 1 & 1 & 1 & & & & & & & \\ 1 & 2 & 2 & 1 & & & & & & \\ 1 & 2 & 3 & 2 & 1 & & & & & \\ & 1 & 2 & 3 & 2 & 1 & & & & \\ & & 1 & 2 & 3 & 2 & 1 & & & \\ & & & 1 & 2 & 3 & 2 & 1 & & \\ & & & & 1 & 2 & 3 & 2 & 1 & \\ & & & & & 1 & 2 & 3 & 2 & 1 \\ & & & & & & 1 & 2 & 3 & 2 \\ & & & & & & & 1 & 2 & 3 \end{bmatrix}$$

Covariance function representing the lower-right part of matrix C:

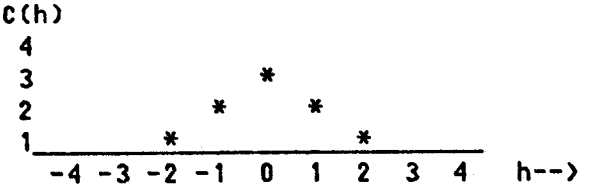


Figure 21: The relationship between the matrix and moving-average formulations of a discrete time series. Given a vector X of ten uncorrelated random numbers with mean=0, variance=1, we generate a discrete moving-average process of the form
$Y = X(t) + X(t-1) + X(t-2)$, for $t = 3,4,\ldots,10$.

represent the space R^n by a collection of discrete grid points in a simulation, Y(t) becomes

$$Y(t) = \sum_k f(t-u_k)X(u_k)$$

where the summation is carried out over all points u_k within the space, or within a sliding local neighborhood if $f(t-u)=0$ for large distances $|t-u_k|$. Clearly the expected value of Y(t) is also zero, and its covariance function $C_y(h)$ can be expressed in terms of $C_x(h)$ and a convolution of the weight function:

$$\begin{aligned} C_y(Y(t),Y(t+h)) &= E[Y(t)Y(t+h)] - EY(t)EY(t+h) \\ &= E\{\int f(t-u)X(u)du \int f(t+h-v)X(v)dv\} - 0 \\ &= \iint f(t-u)f(t+h-v)E[X(u)X(v)]dudv \\ &= \iint f(t-u)f(t+h-v)C_x(u-v)dudv = C_y(h) \end{aligned}$$

Now let's consider a very useful special case where the weight function has a constant value of 1 within a distance $a/2=|t-u|$ in any direction, and a value of zero at distances beyond a/2. (Notice that this amounts to a regularization of X(u) over a hyperspherical volume of diameter "a".) Suppose further that the random function X(u) is a standardized white-noise (pure-nugget-effect) process such that $C_x(0)=1$, $C_x(h)=0$ for $h>0$. Then:

$$\begin{aligned} C_y(h) &= \int_{|t+h-v|\leq a/2} 1 \cdot \int_{|t-u|\leq a/2} 1 \cdot C_x(u-v)dudv \\ &= \int_{\substack{|t+h-u|\leq a/2 \\ |t-u|\leq a/2}} du \end{aligned}$$

In other words, for this case the covariance between Y(t) and Y(t+h) is just equal to the volume of the <u>intersection</u> of two n-dimensional hyperspheres of diameter "a", one centered at t and the other at t+h. An illustration for n=1 and n=2 is provided in Figure 22.

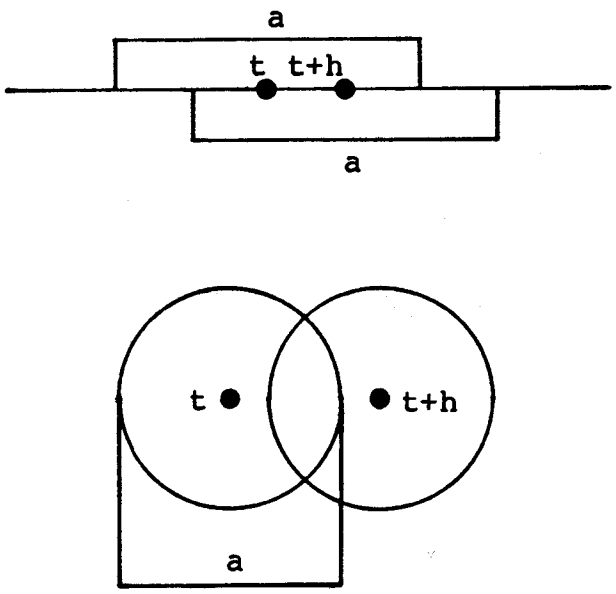


Figure 22: A white-noise random field averaged by a constant-valued hyperspherical moving weight function. In one dimension, the weight function is constant inside a moving line segment of length "a", and zero outside; in two dimensions, the weight function is constant inside a moving circle of diameter "a". The covariance between pairs of points separated by vector h is equal to the measure of the intersection of the two line segments or circles surrounding the respective points.

When the distance h exceeds the diameter of the hyperspheres, the two hyperspheres will no longer intersect, so Y(t) and Y(t+h) will be uncorrelated: $C_Y(h)=0$ for $h \geq a$. Thus the <u>diameter</u> of the hypersphere is the <u>range</u> of the variogram of the Y process. For h=0, the variance $C_Y(0)=\gamma_Y(\infty)=C_Y$ of the Y process is just the "volume" of the hypersphere, e.g.:

For n=1, $C_{y1} = a$ (length of line segment)

For n=2, $C_{y2} = \pi a^2/4$ (area of circle)

For n=3, $C_{y3} = \pi a^3/6$ (volume of sphere)

For distances r=|h| lying between 0 and a, the amount of overlap, $C_y(r)$, of the two hyperspheres can be calculated, or in the case of lower dimensions simply looked up in a mathematical handbook. For $0 \leq r \leq a$:

For n=1, $C_{y1}(r)$ = overlap of two line segments = a-r

For n=2, $C_{y2}(r)$ = overlap of two circles

$$= (a^2/2)\text{Arccos}(r/a) - (r/2)\sqrt{(a^2-r^2)}$$

$$= \pi a^2/4 - (a^2/2)\text{Arcsin}(r/a) - (r/2)\sqrt{(a^2-r^2)}$$

For n=3, $C_{y3}(r)$ = overlap of two spheres

$$= \pi a^3/6 - \pi a^2 r/4 + \pi r^3/12$$

We can continue with calculations of the semivariograms for n = 1, 2, or 3, using the relation $\gamma(r)=C(0)-C(r)$ for $0 \leq r \leq a$, $\gamma(r)=C(0)$ for $r \geq a$:

For n=1, $\gamma_{y1}(r) = [a]-[a-r] = r$

$$= [a] \cdot [r/a] \text{ for } 0 \leq r \leq a;$$

$\gamma_{y1}(r) = [a]$ for $r \geq a$.

For n=2, $\gamma_{y2}(r) = [\pi a^2/4]-[\pi a^2/4-(a^2/2)\text{Arcsin}(r/a)-(r/a)\sqrt{(a^2-r^2)}]$

$$= 0 + (a^2/2)\text{Arcsin}(r/a)+(r/2)\sqrt{(a^2-r^2)}$$

$$= [\pi a^2/4] \cdot [2/(\pi a^2)] \cdot [a^2\text{Arcsin}(r/a)+r\sqrt{(a^2-r^2)}] \text{ for } 0 \leq r \leq a;$$

$\gamma_{y2}(r) = [\pi a^2/4]$ for $r \geq a$.

For n=3, $\gamma_{y3}(r) = [\pi a^3/6]-[\pi a^3/6-\pi a^2 r/4+\pi r^3/12]$

$$= 0 + \pi a^2 r/4-\pi r^3/12$$

$$= [\pi a^3/6] \cdot \tfrac{1}{2} \cdot [3r/a-r^3/a^3] \text{ for } 0 \leq r \leq a;$$

$\gamma_{y3}(r) = [\pi a^3/6]$ for $r \geq a$.

The reader may recognize the semivariograms for n=2 and n=3 as the circular and spherical models, respectively, that were quoted in Section 3.1.3 and graphed in Figure 13, where the sill value C in those formulas has been scaled in this case to $C_{y2}=\pi a^2/4$ for the circular model and $C_{y3}=\pi a^3/6$ for the spherical model. For the case n=1, the analogous semivariogram is linear between r=0 and r=a; beyond r=a, the sill value of $C_{y1}=a$ is attained. A semivariogram of this "simple-transitive"[38] type is depicted in Figure 13, and the associated covariance function is shown in Figure 21 for a discrete time series. Matern (1960, p. 30) provides an expression for a hyperspherical covariance for n=5. Apparently the expression for n=4 is difficult to derive.

These simple models fit many natural phenomena very nicely. As they also can be represented so easily as a regularization of a white-noise process, it is easy to see a simple (though not necessarily fast) way to simulate such a process unconditionally. We simply generate a set of independent identically distributed random numbers (from any distribution with finite moments, uniform being the easiest) on a grid and then smooth them with a constant-valued linear, circular, or spherical moving window. We can also use a higher-dimensional model for a lower-dimensional process; e.g., an unconditional spherical simulation in two dimensions can be obtained (expensively) by averaging a three-dimensional grid of random data using a spherical window whose center travels over a single plane in the grid. Subroutine CS2D ("Circular Simulation, 2 Dimensions"), described in Appendix A, applies this method

---

[38] This is sometimes called the "triangular" model, as in Alfaro (1979).

to unconditional two-dimensional simulations of random functions with circular variograms.

In principle, any random function whose covariance can be expressed in terms of a convolution of a weight function (as shown on page 114) can be simulated by the moving-average method. However, the weight function may not be constant or restricted to a finite neighborhood as in the preceding example. Furthermore, even the very simple moving-average simulation applied in subroutine CS2D can become horribly time-consuming if the number of grid points inside the circular window is large and the number of circular-simulated data to be generated is also large. Simulation by this method requires a summation within 2n nested DO loops: one loop for each dimension inside the window, and one for each dimension of the simulation domain. Subroutine CS2D therefore is not recommended for really big simulation jobs, unless one's computer facility operates free of charge. Obviously a spherical simulation of this type would be excruciatingly slow.

In summary, we can now describe one simple method for unconditionally simulating a random function Y(t) on R^n that has a hyperspherical covariance function:

(1) Generate a grid of independent random numbers x(u) (e.g., uniform on the interval (0,1)) over a finite domain in R^n.

(2) Attribute to each grid node t in the simulation domain a value y(t) equal to the sum of all x(u) located inside a hypersphere of diameter a, centered at t. The simulation domain must be smaller than the x(u) domain by a distance "a" in each dimension.

Alternatively, we can simply generate the values x(u) one at a time at each point u, adding each x(u) value to all values y(t) located at points t within a hypersphere of diameter a, centered at u. This option avoids the necessity of storing and retrieving the whole series of x(u) values.

If the window covers many grid nodes, the original alternative can be speeded up by means of an "updating" formula. Instead of summing all x(u) values inside the hypersphere each time it is moved, the updating formula simply retains the previous sum, subtracts from it the fringe of x(u) values left behind the trailing edge of the hypersphere after it moves, and adds to it the new x(u) values picked up across the leading edge of the hypersphere. This method will not work for general moving-average simulations in which the moving weight function is not constant within the window.

The realization y(t) will have an isotropic n-dimensional hyperspherical variogram. The realization can be standardized to mean=0, variance=1, if desired. Figure 23(a) depicts the arrangement of points used for generating a circular realization on a rectangular spatial domain using a moving-average method.

3.3.1.3 Random-Average Methods

In the moving-average procedure, a hyperspherical volume is approximated by a set of discrete grid points, so the technique has something in common with the numerical quadrature procedures used for performing an approximate integration on a computer. It turns out that the alternative class of numerical integration procedures, the "Monte Carlo"

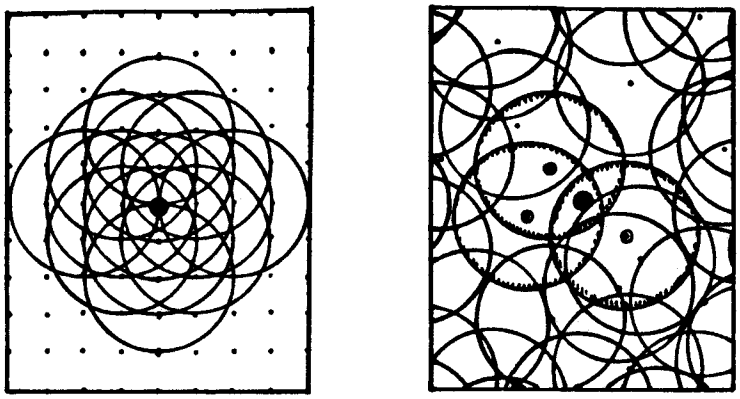


Figure 23: Simulation of a random function with a circular variogram by the moving-average and random-average methods. (The cheaper approach of centering the circles on the points u, adding the value x(u) to y(t) at all points t inside the circle, is depicted above. The equivalent approach of centering the circles on the points t is more intuitive and is used in the text to introduce the methods.) (a) Moving-average simulation: the central simulated value receives contributions from several surrounding x(u) values. The spacing between points u_i for the X(u) white-noise process must be less than or equal to the spacing between points t_j for the circular Y(t) process; otherwise adjacent y(t) values will occasionally be identical. The most efficient procedure is to use equal spacings for u_i and t_j, as in subroutine CS2D. (b) Random-average method: the central simulated value receives contributions from the x(u) values that happen to fall nearby (inside the shaded circles). The more dense the spacing of the Poisson points u_i, the less common will be the occurrences of adjacent identical y(t) values.

approach, has an analogy in simulations of random functions -- the "random-average" procedures. These procedures rely on the characteristics of the n-dimensional Poisson process. The Poisson process is a "point process" -- a random function describing the

distribution of discrete points in space.[39] (For some homely geologic examples, consider the locations of diamonds in a kimberlite pipe in three dimensions, the locations of kimberlite pipes on a map of Africa in two dimensions, or the times of kimberlite intrusions in one dimension.) The Poisson process is the simplest point process, defined in either of two ways:

(1) The random number K of points inside a volume of a given size follows a Poisson distribution with mean=variance=ν. This distribution has the discrete probability mass function $P(k)=P\{K=k\}=e^{-\nu}\nu^k/k!$ for integers k=0,1,2,...

(2) The numbers of points inside nonoverlapping volumes are independent; i.e., the point locations are "purely random".

A Poisson process is very easy to simulate, as the locations of the points are independent and uniformly distributed in n dimensions. If we wish to distribute Poisson points within a unit hypercube, we just draw independent vectors of n independent random numbers distributed uniformly on the interval (0,1), and assign a point to the coordinate location represented by each vector. For a rectangular hyperprism, we just rescale the range (0,1) of each coordinate to the arbitrary endpoints (a,b). For an irregular volume, we first enclose the volume inside a rectangular hyperprism, then discard all random points lying inside the prism but outside the volume.

Now what would be the covariance of the Poisson variable K(t), defined as the number of random points in a volume of size v, centered at t? If the volume v is translated in space, K(t) and K(t+h) will be

---

[39] Spatial point processes are described by Diggle (1983) and Ripley (1981, Chapters 7 and 8).

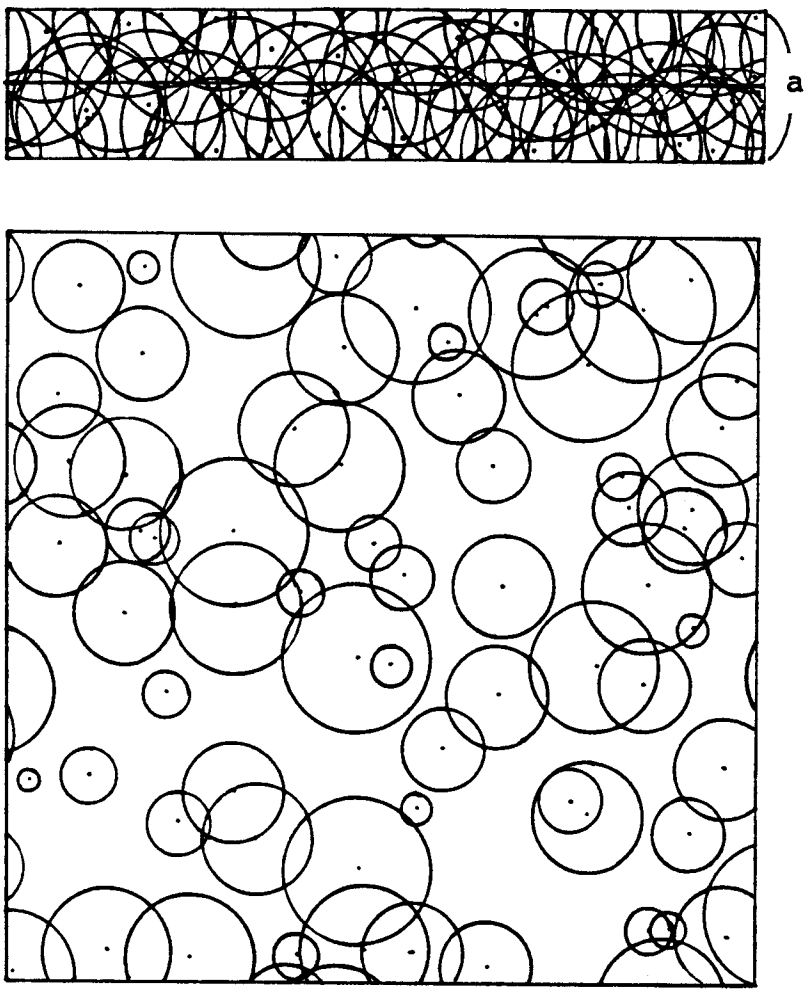


Figure 24: Random-coin simulation of a spherical random function on the plane. The upper illustration is a side view of a slice of vertical thickness "a" within the three-dimensional space, showing the locations of several random spheres of diameter "a" centered within the slice. The two-dimensional simulation plane is represented by the horizontal line through the middle of the slice. The lower illustration is a plan view of the simulation plane, showing the random-sized circular intersections of the constant-sized spheres with the plane. Notice that we need a lot of circles to avoid generating a "mosaic" simulation with locally constant values.

generated by drawing a random azimuth uniformly from the interval $(-\pi/2,+\pi/2)$ and a random distance from the center of the circle uniformly on (-R,+R). In three dimensions the domain is inscribed in a sphere. The distance from the center of the sphere to each simulated plane and the azimuth and plunge of the pole from the center to the plane are then drawn from (-R,+R), $(0,2\pi)$, and $(0,\pi/2)$, respectively.

3.3.1.4 Turning-Bands Method

The turning-bands method of Matheron (1973) is described in considerable detail by Journel and Huijbregts (1978) for three dimensions and by Mantoglou and Wilson (1981, 1982) and Bras and Rodriguez-Iturbe (1985) for both two and three dimensions, so only the essentials are presented here. This fast method builds a higher-dimensional simulation by averaging the contributions of several independent one-dimensional simulations oriented in several directions in space.

The method is geometrically easier to visualize in two dimensions. For a two-dimensional simulation (Figure 25), we first simulate (usually by a moving-average or spectral technique) several realizations from a one-dimensional random function with a prespecified covariance function $C_1(s)$, where s=|h| in one dimension. These realizations are attributed to lines that pass through the origin of the two-dimensional grid and are oriented in a regular fashion in the two-dimensional space (e.g., eight lines separated by angles of $\pi/8$).[42] Each regularly spaced

---

[42] Alternatively, the lines can be randomly oriented uniformly on the interval $(0,\pi)$. The distinction between regular and random line orientations is analogous to the distinction between the regular grid of white-noise points used in the moving-average method and the random collection of points used in the random-average method. The sample covariance functions calculated from turning-bands simulations

simulated point on each line is given a "band" of influence in the two-dimensional space. This band is perpendicular to the line and is delimited at the line by the midpoints between the simulated line point and the adjacent points on either side. To obtain a simulated value for the two-dimensional random function at some point on the two-dimensional grid, we ascertain which bands contain that grid point and simply sum up the simulated line values associated with those bands. If we wish the variance of the simulated point on the plane to be the same as the variance of the line processes, we just divide the sum by $\sqrt{L}$, where L is the number of lines. This two-dimensional realization will be drawn from some random function having an isotropic covariance $C_2(r)$, where $r=|h|$ on the plane. Given a desired covariance function $C_2(r)$, we must derive, by methods described below, the function $C_1(s)$ that should be specified for the independent line simulations.

In three dimensions the geometrical situation is analogous. We need a large number of lines regularly (or randomly) oriented in three-dimensional space. A good choice is fifteen lines connecting the opposite edges of a regular icosahedron (with twenty faces), as described by Journel and Huijbregts (1978, p. 503).

To see how the line covariance function $C_1(s)$ corresponds to a given two- or three-dimensional covariance C(r), consider the simple two-dimensional example in Figure 26. We generate two values $y(t_1)$ and $y(t_2)$ at two arbitrary points t_1 and t_2 on the plane, using contributions from only two (for simplicity) independent coplanar line

---

converge more rapidly (with fewer lines) to their theoretical covariances in the case of regularly oriented lines (Mantoglou and Wilson, 1981, p. 58).

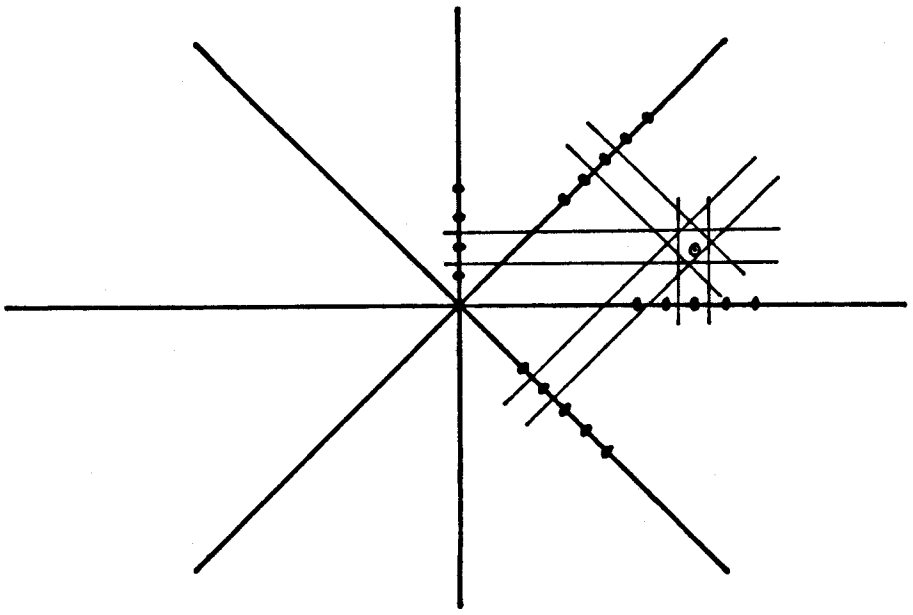


Figure 25: Illustration of the turning-bands method in two dimensions. Each simulated point on the plane is the average of all simulated line values whose bands of influence contain the point. For simplicity, only four line simulations are depicted here. Actually eight or sixteen lines would be preferable.

simulations -- $x(u_1)$ on line u_1 and $x(u_2)$ on line u_2. Each of these second-order-stationary line realizations has been generated independently with mean zero and covariance $C_1(s)$. The covariance between $Y(t_1)$ and $Y(t_2)$ is:

$$C_2(Y(t_1),Y(t_2)) = E\{[(X(u_{11})+X(u_{21}))/\sqrt{2}][(X(u_{12})+X(u_{22}))/\sqrt{2}]\}$$

Because the lines are independent:

$$= \{E[X(u_{11})X(u_{12})]+E[X(u_{21})X(u_{22})]\}/2$$

$$= [C_1(u_{11}-u_{12})+C_1(u_{21}-u_{22})]/2$$

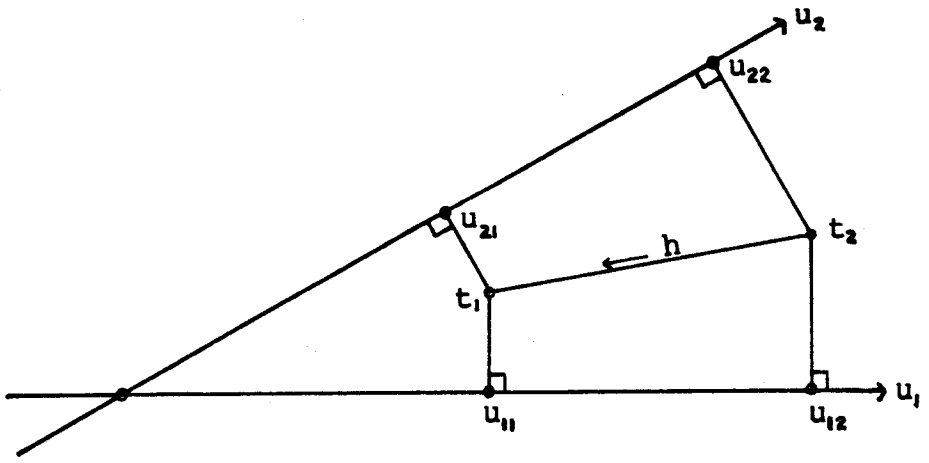


Figure 26: The relationship between covariances on the lines and on the plane. The two independent lines are u_1 and u_2. The two values to be simulated by turning bands are at t_1 and t_2.

We see that the covariance between the two Y(t) values is just the average of the covariances between the colinear X(u) values occurring closest to the projections of t_1 and t_2 onto the independent lines u_1 and u_2. Representing the projection of $h=(t_1-t_2)$ onto each line u_i as $\langle h,u_i\rangle$, the covariance in n-dimensional space of the Y's, when simulated from L independent line processes, is:

$$C_n(h) = (1/L)\sum_{i=1}^{L} C_1(\langle h,u_i\rangle)$$

If we could produce an infinite number of such lines with orientations varying uniformly over all directions in n-dimensional space, we would generate a process with an isotropic covariance $C_n(r)$, $r=|h|$,

$$C_n(r) = (1/S_n)\int_{S_n} C_1(\langle h,u\rangle)du$$

where S_n is the surface area of a hypersphere of radius=1. An explicit general expression for this relation in n dimensions is stated by Matheron (1973, p. 462). The solution for the case n=3 is very simple:

$$C_1(r) = d/dr\ [rC_3(r)]$$

This is usually easy to solve given any $C_3(r)$. This expression is derived in detail by Mantoglou and Wilson (1981, Chapter 3) and in less detail by Journel and Huijbregts (1978, p. 500). For n=2 the integral equation

$$C_2(r) = (2/\pi)\int_0^r [C_1(s)/\sqrt{(r^2-s^2)}]ds$$

(derived by Mantoglou and Wilson, 1981, Chapter 3) has a much more complicated solution, obtained by Brooker (1984, 1985):

$$C_1(r) = C_2(0)+\int_0^r [r/\sqrt{(r^2-s^2)}]\cdot[d/ds\ C_2(s)]ds$$

For the commonly used spherical model with sill=1, this solution yields

$C_1(r) = 1-(3\pi/4)[(r/a)-\frac{1}{2}(r/a)^3]$ for $r \leq a$

$= 1-(3/2)[(r/a)-\frac{1}{2}(r/a)^3]\arcsin(a/r)-(3/4)(r/a)^2\sqrt{[1-(a/r)^2]}$ for $r \geq a$

where "a" is the range of the spherical model.[43]

Before the solution for n=2 became known, it was difficult to obtain appropriate line realizations for the two-dimensional case, because there was no general expression for $C_1(s)$ in terms of $C_2(r)$. This problem was addressed at length by Mantoglou and Wilson (1981, 1982), who proposed two approaches: simulation of the line processes by spectral methods (the more versatile approach), and derivation of

---

[43] The expression for $r \leq a$ was also stated by Dagbert (1981).

one-dimensional covariances corresponding to a few particular two-dimensional covariance functions. They obtained a remarkably simple expression of the spectral density function $S_1(\omega)$ of the line process (defined in Section 3.3.2.1) in terms of the radial spectral density function $f(\omega)$ of the two-dimensional process,

$$S_1(\omega) = C_2(0)f(\omega)/2$$

where $C_2(0)$ is the variance of the two-dimensional process. The function $f(\omega)$ can be calculated (sometimes with difficulty) from $C_2(r)$. Knowing $S_1(\omega)$, one of the spectral simulation methods described in Section 3.3.2.2 can be used to simulate the line processes, without need of deriving any expression for $C_1(s)$.

Mantoglou and Wilson (1981, 1982) describe the convergence of the covariance of the simulated two-dimensional realization to the covariance of the model random function in terms of four factors: the number of simulated lines, the discretization length (bandwidth) along the lines, and the maximum frequency and number of harmonics used to represent the line spectral density. The most important of these variables is the number of lines, L. For equally spaced lines, the error around the true covariance is proportional to $1/L^2$. Mantoglou and Wilson (1981) also derive some rather complicated expressions for $C_1(s)$ for exponential and Bessel $C_2(r)$ functions, calculating the associated moving-average weight functions numerically.

For two-dimensional simulations, Mantoglou and Wilson (1981, 1982) report that their implementation of the turning-bands method is less expensive than the direct spectral methods of Shinozuka and Jan (1972) and Mejia and Rodriguez-Iturbe (1974), both discussed in Section

3.3.2.2, for the same accuracy. To generate values at N points on the plane, the cost is roughly proportional to $\sqrt{N}$ for turning bands with equally spaced lines, N for the spectral methods. It is also approximately proportional to N for the moving-average and random-average methods described earlier, but probably would be less accurate for the same cost using the random-average method. The moving-average and random-average methods also vary in cost in proportion to the number of points (regularly or randomly spaced) used to discretize the weight function.

Mantoglou and Wilson (1981, Chapter 4) provide a variation on the turning-bands method for simulating anisotropic random functions on the plane, by keeping the lines uniformly spaced but allowing their covariances $C_1(s)$ to vary with line orientation. In most practical situations, such an anisotropy can be more easily created by a simple transformation of the coordinate system. Mantoglou and Wilson (1981) also provide descriptions of turning-bands simulations of block grades (for regularized phenomena) and of some nonstationary phenomena with polynomial drifts ("intrinsic random functions"), which are described by Matheron (1973).

Subroutine TB3D ("Turning Bands, 3 Dimensions"), described in Appendix B, is a revised version of subroutine SIMUL, described in Journel and Huijbregts (1978, p. 537). Line processes are generated by the moving-average method, using discretized weight functions. This subroutine is specifically designed for three-dimensional simulations, using the fifteen-line icosahedron approach, and is thus not very efficient for the simulation of two-dimensional processes, for which

eight or sixteen coplanar lines separated by angles of $\pi/8$ or $\pi/16$, respectively, would be a better choice. Mantoglou and Wilson (1981, 1982) have developed and extensively tested a program for two-dimensional simulations using this arrangement.

3.3.1.5 Autoregressive Methods

Many discrete time series, and some very restrictive discrete multidimensional (lattice) phenomena, display a covariance structure consistent with an autoregressive (AR) model, which can be used to perform unconditional simulations. The general form of a time-series autoregressive model of order p, denoted AR(p), is

$$Y(t) = \alpha_1 Y(t-1) + \alpha_2 Y(t-2) + \ldots + \alpha_p Y(t-p) + X(t)$$

where X(t) is a stationary independent (white-noise) random variable with mean zero, usually considered to be normally distributed, and the coefficients α are constants subject to certain sufficient constraints to ensure stationarity (Chatfield, 1980, p. 48).[44] The processes Y(t) and X(t) are observed only at unit intervals of time t. This "one-sided" expression can be reformulated, if appropriate, to yield an autoregression on both past and future values. Autoregressive and moving-average (MA(q))[45] time-series models may be combined into ARMA(p,q) models and extended to nonstationary ARIMA(p,d,q) models, in

---

[44] The discrete $CaSO_4$ time series displayed in Figure 19(a) fits an AR(1) model (which has an exponential variogram for $0<\alpha<1$), but this model could not be used for simulating the intervening points between the data points. For simplicity, a simple-transitive model was used in the unconditional simulation and a linear model was used as a short-scale approximation in the kriging, instead of a discrete time-series model.

[45] Figure 21 illustrates a simple MA(2) process.

which the order-d differenced (i.e., dth-derivative) time series is stationary. Further extensions allow the inclusion of periodic components. This rich variety of discrete time-series models is discussed fully in Box and Jenkins (1976) and in numerous other basic texts in time-series analysis. A general comparison of the methods and goals of time-series analysis and geostatistics is provided by Solow (1984).

Clearly an autoregressive model is easy to simulate, given a starting set of p consecutive conditioning data: for each y(t), one simply generates a new independent "innovation" x(t) and obtains y(t) by applying the formula above. The variogram of an autoregressive process may be of exponential or hole-effect form, depending on the α coefficients. From inspection of the variogram one can decide at what time the simulated process will be nearly uncorrelated with its starting values (the "practical range" of the variogram). Beyond this time, one effectively has an unconditional simulation of the process.

The simplicity of generating one-dimensional autoregressive processes would seem to make them attractive for use in turning-bands simulations, but the higher-dimensional covariance functions corresponding to autoregressive line processes do not seem to be so simple. However, Mantoglou and Wilson (1981, p. 93) have derived one two-dimensional covariance function corresponding to a simple autoregressive line process.

Nearest-neighbor method. Spatial autoregressions (Ripley, 1981, p. 88) have received some limited application in the earth sciences. A particular case, the "first-order nearest-neighbor model", has been used

by Smith and Freeze (1979) to generate a two-dimensional field of hydraulic conductivities for an aquifer model, and a reformulation of their approach has been applied by Baker (1984) to the simulation of soil properties. The first-order nearest-neighbor model is

$$Y(u_1,u_2) = \alpha_1[Y(u_1-1,u_2)+Y(u_1+1,u_2)]+\alpha_2[Y(u_1,u_2-1)+Y(u_1,u_2+1)]+\mathcal{E}(u_1,u_2)$$

where u_1 and u_2 are the coordinates of a discrete two-dimensional grid point (with grid spacings scaled to 1), α_1 and α_2 are autoregressive coefficients associated with the u_1 and u_2 directions, and $\mathcal{E}(u_1,u_2)$ is a normal white-noise random variable. Smith and Freeze provide methods for selecting the variance of $\mathcal{E}$ and the appropriate values of α_1 and α_2 that will provide a stationary Y process with prespecified exponential covariance functions in the u_1 and u_2 directions. Their simulation is performed by a matrix method, first generating a field of uncorrelated standard normal variables represented by vector E, then solving for a vector Y of simulated zero-mean values,

$$Y = W\,Y + \eta\,E$$

or equivalently

$$Y = (I-W)^{-1}\,\eta\,E$$

where W is the weight matrix composed from the α's, I is the identity matrix, and η is a scalar multiplier used to control the variance of Y. The stationary mean of the process is added after the simulation. Notice that the second formulation above, which is the one actually used to simulate Y, is just a matrix formulation for a moving-average representation of the Y process.

Although nearest-neighbor methods apparently have been widely used to simulate subsurface hydrologic variables, Mantoglou and Wilson (1982, p. 1391) point out that this approach "provides a very particular

correlation structure that is not directly related to any isotropic two-dimensional covariance function, nor can the autoregressive parameters of the model be easily estimated directly from field data. . . . This approach yields an anisotropic field because care is taken to preserve the correlation structure only in two orthogonal directions." To apply this method in a conditional simulation, one could either approximate the semivariogram of the simulated random function by some simple model for use in kriging, or one could derive a covariance matrix $C=AA'$, where $A=(I-W)^{-1}\eta$, then use the conditional matrix method explained on page 109 to draw a realization of the unknown values from their conditional distribution, given the fixed values. However, the severe limitations of nearest-neighbor models as applied to real-world phenomena seem to recommend against their use in most practical situations. It would be easier and probably geologically more realistic to use an isotropic moving-average formulation that would resemble the weighting system $(I-W)^{-1}\eta$.

A major difficulty with the nearest-neighbor approach described above is the large size of the matrix $(I-W)^{-1}\eta$ that must be stored and manipulated. Sharp and Aroian (1985) have proposed an alternative "herringbone" method that generates the realization recursively, beginning with two or three one-dimensional AR(1) simulations positioned along the edges of the two- or three-dimensional simulation grid. This method is fast and avoids the generation of a huge weight matrix.

3.3.2 Frequency-Domain Approaches

Geologic phenomena do not commonly exhibit periodic behavior, so frequency-domain representations do not usually contribute much to an understanding of the phenomena. Nevertheless there are some limited advantages to the frequency-domain approach in simulation, particularly in the simulation of the line processes used in the turning-bands method. Unfortunately the mathematics can be relatively complicated and the notation used by different authors is wildly nonstationary. The discussion that follows uses notation similar to that of Mantoglou and Wilson (1981), who describe two approaches that have been used for direct frequency-domain simulations of spatial random functions.

Chatfield (1980, Chapters 6 and 7) provides a short, simple introduction to the frequency domain as it is applied in the analysis of discrete time series. Mantoglou and Wilson (1981, Chapter 2), Ripley (1981, Chapter 5), and Borgman et al. (1984) discuss some extensions into two and higher dimensions. A very brief summary for one and two dimensions is provided below.

3.3.2.1 Mathematics of the Frequency Domain

Consider a one-dimensional (time-series) stationary gaussian random function Y(t) with a covariance function C(h) satisfying:

$$\int_{-\infty}^{+\infty} |C(r)|\,dr < \infty$$

In the time-domain view of Y(t), the mean m(t)=m and the covariance function C(t,t+h)=C(h) are used to characterize the spatial distribution of this random function (and in the gaussian case, they characterize it

fully). In the frequency domain we use an alternate expression for the covariance,

$$C(r) = 2\int_0^\infty \cos(\omega r)S(\omega)d\omega$$

where $S(\omega)$ is the "spectral density function" of the time series, describing the contribution of variation at frequency ω to the total variance of the process. This expression has an inverse,

$$S(\omega) = (1/\pi)\int_0^\infty \cos(\omega r)C(r)dr$$

which is positive at all ω if the covariance function $C(r)$ is positive-definite. The frequency-domain approach relies on a model of the process $Y(t)$ as the sum of an infinite number of sinusoidal functions of the form

$$R\cos(\omega t+\phi) = R\cos(\phi)\cos(\omega t)-R\sin(\phi)\sin(\omega t)$$

where R is the amplitude, ω the frequency in radians, and ϕ the phase in radians. The amplitudes and phase angles associated with different values of ω are considered to be independent random variables, fixed for a particular realization of the process.

Because of the invertible relationship between $C(r)$ and $S(\omega)$,[16] it is clear that they contain the same information about the process but express it in different forms. The covariance $C(r)$ or the semivariogram $\gamma(r)=C(0)-C(r)$ expresses the variability at different <u>scales</u> or <u>distances</u> as compared to the total variability $C(0)=\gamma(\infty)$ of the process, whereas the spectral density $S(\omega)$ expresses the contribution of

---

[16] The expressions above for $C(r)$ and $S(\omega)$ constitute a Fourier transform pair in which $Y(t)$ must be a real-valued function.

variability at different frequencies to the total variability:[47]

$$C(0) = 2\int_0^\infty S(\omega)d\omega$$

In two dimensions, the analogous expression for C(h) is

$$C(h) = \int_{R^2} \cos(h\bullet\omega)S(\omega)d\omega$$

where h and ω are two-dimensional vectors and $(h\bullet\omega)$ is their inner product. Correspondingly:

$$S(\omega) = [1/(4\pi^2)]\int_{R^2} C(h)\cos(h\bullet\omega)dh$$

For isotropic processes, these functions can be expressed in terms of $r=|h|$ and $\omega=|\omega|$, leading to the following compact expressions for the isotropic covariance C(r) and "radial spectral density function" $f(\omega)$ (Mantoglou and Wilson, 1981, p. 26):

$$C(r) = C(0)\int_0^\infty J_0(\omega r)f(\omega)d\omega$$

$$f(\omega) = 2\pi\omega S(\omega)/C(0) = (\omega/C(0))\int_0^\infty C(r)J_0(\omega r)rdr$$

where C(0) is the variance of the process and J_0 is a Bessel function of the first kind, order zero. Mantoglou and Wilson (1982) provide expressions for $f(\omega)$ corresponding to two-dimensional isotropic exponential, Bessel, gaussian ("double-exponential"), spherical, and "Telis" covariances. Similar expressions (in quite different notation) are provided by Borgman et al. (1984) for the simple-transitive

---

[47] The factor of 2 in the expressions for C(r) and C(0) arises because we are considering only real processes, in which C(r) and $S(\omega)$ are real and symmetric functions that we may integrate over $(0,\infty)$ instead of $(-\infty,+\infty)$.

(positive-definite only in R^1), spherical, exponential, gaussian, and hole-effect covariances.

3.3.2.2 Spectral Simulation Methods

Method of Shinozuka and Jan (1972). One way to simulate a realization of Y(t) in one dimension is to begin with a discrete approximation to the covariance function on page 140,

$$C(r) \approx 2\sum_{k=1}^{K}\cos(\omega_k r)S(\omega_k)\Delta\omega$$

where again $2\sum S(\omega_k)\Delta\omega \approx C(0)$, the variance of the process. The discrete weights $S(\omega_k)\Delta\omega$ correspond to a histogram of K spikes representing K harmonics ω_k spaced at distances $\Delta\omega$, which together approximate the continuous spectral density function $S(\omega)$, as shown in Figure 27. The covariance can be reproduced by simulating a discretized sinusoidal model of the one-dimensional random function Y(t),

$$Y(t) \approx \sqrt{2}\sum_{k=1}^{K}\sqrt{[S(\omega_k)\Delta\omega]}\cos(\omega'_k t+\phi_k)$$

where ω'_k is the frequency ω_k perturbed by a small random variable added to avoid exact periodicities, and ϕ_k is an independent random phase angle uniformly distributed on $(0,2\pi)$. We simulate Y(t) at position t along the line by sampling a value of ϕ_k and ω'_k for each of the K harmonics ω_k and summing up their contributions according to the formula. The more harmonics ω_k that we use in the discretization, the more closely the simulated values will conform to the theoretical spectrum $S(\omega)$ and its associated covariance C(r). As in the simulation methods discussed previously, each simulated value is composed of the

sum of several random variables, resulting in a simulated gaussian process. This approach, suggested by Rice (1954), is presented more thoroughly by Mantoglou and Wilson (1981, p. 37).

Shinozuka and Jan (1972) propose a generalization to n dimensions, also discussed by Mantoglou and Wilson (1981, p. 36). In n dimensions, we must discretize the n-dimensional spectral density into $K_1 \cdot K_2 \cdot \cdot \cdot K_n$ spikes of an n-dimensional histogram:

$$S(\omega 1_{k1}, \omega 2_{k2}, \ldots, \omega n_{kn}) \Delta\omega 1 \cdot \Delta\omega 2 \cdot \cdot \cdot \Delta\omega n$$

For each Y(t) value that we simulate, we must draw n values of ω' and one of ϕ corresponding to each of the $K_1 \cdot K_2 \cdot \cdot \cdot K_n$ spikes. Mantoglou and Wilson suggest that $K_1 = K_2 = 20$ harmonics (400 altogether) be used to get sufficient "earth-science" accuracy for a typical two-dimensional simulation with an exponential covariance. For a large grid of simulated Y(t) values, such a simulation could be rather expensive.

Mantoglou and Wilson concern themselves primarily with exponential covariances and others of similar appearance. The accuracy of the simulation for a given density of discretization depends on the type of covariance being modeled (Journel, 1974a, p. 42) -- particularly on its behavior near the origin, $r \to 0$. For models with parabolic behavior near the origin, $S(\omega) \to 0$ quickly as $\omega \to \infty$. For instance, $S(\omega)$ for the gaussian semivariogram model (Figure 13) is a gaussian (normal) density function. Models with linear behavior near the origin show less continuity at short distances, thus more high-frequency variation, so $S(\omega) \to 0$ slowly as $\omega \to \infty$. For example, the exponential covariance model yields a Cauchy density for $S(\omega)$. To discretize such a long-tailed density, we must either sacrifice some detail in the discretization by making the

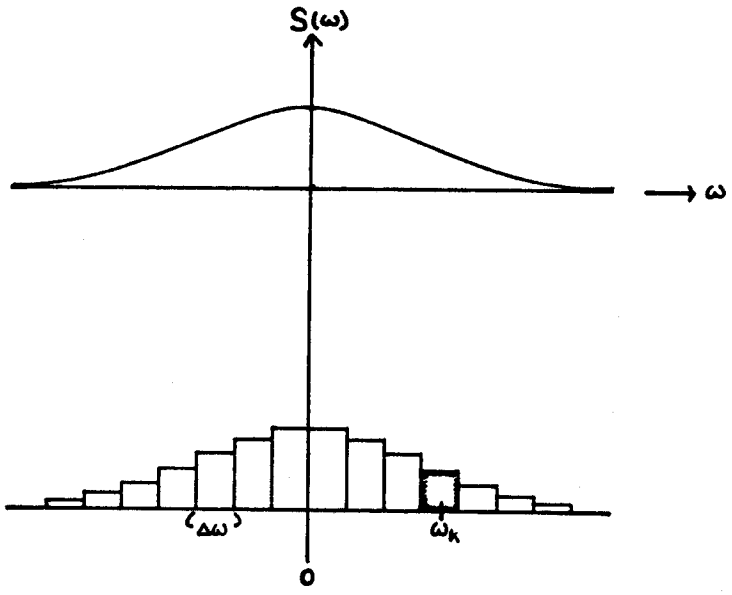


Figure 27: Discretization of a one-dimensional spectral density function. The weight of the shaded spike is $S(\omega_k)\Delta\omega$. For real-valued processes, only the positive frequencies have to be considered.

increment $\Delta\omega$ large, or use a large number K of increments, driving up the cost of the simulation.[48] So in general, spectral methods are cheaper to implement in the case of very continuous (low-frequency) random functions, which unfortunately are not so commonly encountered in geostatistical practice.

---

[48] This would not be a problem if the density of the simulation grid were very coarse. In that case frequencies higher than one cycle per two grid spacings (i.e., the Nyquist frequency -- Chatfield, 1980, p. 131) could not be observed and should not be simulated, in order to avoid "aliasing" (Chatfield, 1980, p. 156).

Method of Mejia and Rodriguez-Iturbe (1974). This method, also summarized by Mantoglou and Wilson (1981, p. 35) and by Bras and Rodriguez-Iturbe (1985, p. 288), samples randomly from the spectral density function, whereas the method of Shinozuka and Jan (1972), described above, discretizes the function with a regular grid and weights the cosine terms in the simulation equation according to the associated density. Briefly, in one dimension the model of Mejia and Rodriguez-Iturbe is

$$Y(t) \simeq \sqrt{(2C(0)/K)}\sum_{k=1}^{K}\cos(\omega_k t+\phi_k)$$

where $C(0)$ is the variance of the process, ϕ_k is a random phase angle uniform on $(0,2\pi)$, and ω_k is an independent random variable sampled from the probability density function $S(\omega)/C(0)$. For n dimensions, ω_k and t become vectors. Two-dimensional isotropic simulations can make use of the radial spectral density function. For one-dimensional processes, Mantoglou and Wilson (1981) state that the covariance of this type of simulation converges to the model covariance as $1/\sqrt{K}$, much more slowly than $1/K^2$ for the method of Rice (1954).[49]

Applications in turning-bands simulations. Because of the simple relationship (Mantoglou and Wilson, 1981, p. 61)

$$S_1(\omega) = C_2(0)f(\omega)/2$$

---

[49] We have seen this distinction in two other contexts: the moving-average method versus the random-average method, and evenly spaced lines versus randomly oriented lines in the turning-bands method. It appears that in all of these approaches the discretization methods, analogous to an integration by numerical quadrature, converge more rapidly to the theoretical covariance than the Monte Carlo methods.

between the spectral density $S_1(\omega)$ of a turning-bands line process and the radial spectral density $f(\omega)$ of the associated two-dimensional process, spectral methods turn out to be very useful in the simulation of line processes for two-dimensional turning-bands simulations. Using this relationship and the radial spectral densities provided by Mantoglou and Wilson (1982) and Borgman et al. (1984), one can obtain line spectral densities $S_1(\omega)$ corresponding to a variety of two-dimensional processes.[50] Mantoglou and Wilson use the method of Rice (1954), as modified by Shinozuka and Jan (1972), to generate line processes for their two-dimensional turning-bands simulations.

Method of Davis et al. (1981). An approach applicable to one-dimensional simulations for which the one-dimensional covariance is known is presented by Davis et al. (1981), who provide a FORTRAN computer program employing the finite Fourier transform of Cooley and Tukey (1965). Their simulation algorithm employs a numerical Fourier transformation of an arbitrary discretized one-dimensional covariance function. Because the one-dimensional covariance must be known, this approach has received little application, but it is more valuable now that Brooker's (1985) solution for one-dimensional turning-bands covariances is available.

---

[50] For instance, the commonly used exponential covariance function $C(r)=\exp(-r/a)$ (for sill=1) has a radial spectral density function $f(\omega)=a^2\omega/(1+a^2\omega^2)^{3/2}$. For the line spectrum, one simply divides this expression by 2. Unfortunately, the even more commonly used spherical covariance has a decidedly complicated $f(\omega)$, explicitly stated in Mantoglou and Wilson (1982, p. 1381). For spherical simulations in two dimensions, it would probably be easier to use Brooker's (1985) solution for the one-dimensional spherical covariance in combination with the finite-Fourier-transform simulation approach of Davis et al. (1981), discussed in the next paragraph.

Method of Borgman et al. (1984). This technique simulates a three-dimensional realization as a stack of two-dimensional realizations, each created using two-dimensional radial spectra and a two-dimensional simulation method similar to that of Davis et al. (1981). The three-dimensional structure is built from models of the two-dimensional radial spectral density of the individual planes and a matrix of direct and cross spectra among the planes. This spectral matrix model is used to generate correlated normal random variables by the eigenvalue method (discussed in Section 3.3.1.1) for use in the separate planar simulations. The difficulty of simulating long vectors of random data by matrix methods effectively limits the extent of the simulation domain in the third dimension.

3.4 CONDITIONAL SIMULATION OF COREGIONALIZATION FOR GAUSSIAN PROCESSES

This short section summarizes the essential steps in a conditional simulation of a stationary gaussian coregionalization -- the only kind of phenomenon that can be simulated with the methods described so far. In this simplest possible (but highly restrictive) case, we simulate k•N values $\{z_i(x_j),\ j = 1 \text{ to } N,\ i = 1 \text{ to } k\}$ drawn from a k•N-variate normal distribution. At each point x_j, the k cross-correlated random variables $Z_i(x_j)$ have a k-variate normal distribution characterized by the stationary mean vector $m = (m_1, m_2, \ldots, m_k)'$ and the kxk positive-definite symmetric covariance matrix:

$$C(0) = \begin{bmatrix} C_{11}(0) & & & \\ C_{21}(0) & C_{22}(0) & & \\ \vdots & \vdots & \ddots & \\ C_{k1}(0) & C_{k2}(0) & \cdots & C_{kk}(0) \end{bmatrix}$$

Similarly, all k•N random variables taken together have a k•N-variate normal distribution with a k•N-component mean vector consisting of $(m_1,m_2,\ldots,m_k)'$ repeated N times, and a (k•N)x(k•N) positive-definite symmetric covariance matrix consisting of the kxk C(h) matrix (shown above for h=0) repeated NxN times (for all possible vectors h separating the locations x_j).

Unfortunately, real-world earth-science data seldom fit this model, as we will see in Section 3.5.1. Consequently, we must consider a variety of transformations, to be introduced in Section 3.5, for converting other types of data into a form acceptably close to this model. Once we have imposed an approximation of multinormality on the transformed data set, we proceed with the simulation using the following steps:

(1) Perform a variographic (structural) analysis of the (usually transformed) data set to infer a set of positive-definite semivariogram models, with a positive-definite matrix of sills or covariance functions, as described in Sections 3.1.3 and 3.2.

(2) Formulate a linear model of coregionalization, as described in Section 3.2.

(3) Perform independent unconditional simulations of the independent components of the linear model, using an appropriate method selected from those described in Section 3.3 (and heeding the caveats of Section 3.7).

(4) Combine the unconditional simulations as directed by the linear model to obtain an unconditional simulation of coregionalization.

(5) Using the locations from the original data set, perform two ordinary krigings (or cokrigings, for some cases described in Section 3.7) -- one using the original (usually transformed) data set, and one using simulated data from the corresponding locations.

(6) Combine the two krigings and the unconditional simulation of each coregionalized variable according to the formula

$$z_{sc}(x) = z^*_{0k}(x) + [z_s(x) - z^*_{sk}(x)]$$

as described in Section 3.1.7, page 93, to obtain the final conditional simulation of coregionalization.

(7) Check the statistics and variograms of the final simulation to confirm agreement with the important characteristics of the data set.

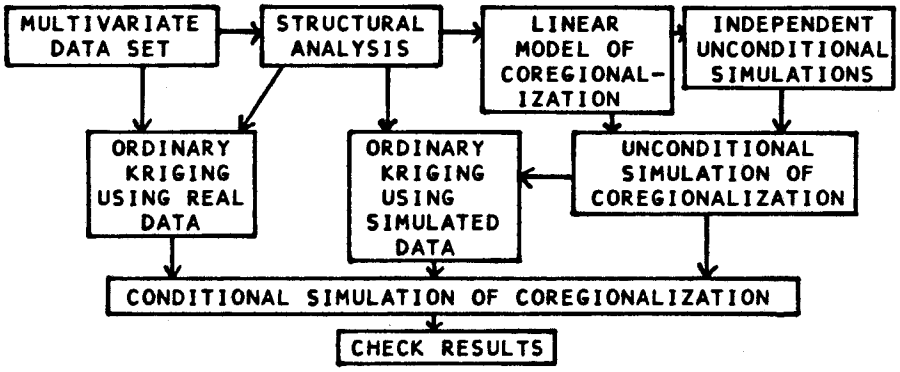


Figure 28: Steps in a conditional simulation of coregionalization for a gaussian process.

These steps are summarized in a flow chart in Figure 28. If the original data were transformed prior to the structural analysis in

step (1) above, an inverse transformation is performed after step (7) to obtain a simulation of the original phenomenon. Step (7) should be repeated after the inverse transformation to confirm that the important characteristics of the original data have been preserved. The importance of transformations is discussed in the next section.

3.5 TRANSFORMATIONS OF COREGIONALIZED DATA

3.5.1 Nongaussian Processes in the Earth Sciences

All of the unconditional simulation methods reviewed in Section 3.3 produce realizations of gaussian processes. Because linear combinations of gaussian processes also are gaussian, an unconditional simulation of coregionalization, created by combining gaussian realizations according to a linear model, is a multivariate gaussian process. If the conditioning data are drawn from a gaussian process, the conditional simulation will be gaussian as well.

But it is evident that few earth-science phenomena can be realistically modeled as gaussian processes. Gaussian distributions are symmetrical and unbounded, whereas many earth-science variables have markedly skewed distributions or distributions with domains bounded by maximum or minimum constraints. Trace-element and permeability data are usually positively skewed and are generally considered to be approximately lognormally distributed. All percentage data are bounded by 0% and 100%, although percentage data with small variances and means far from the constraints, such as sandstone porosity data and some major-element chemical data, may look gaussian. Some major-element data commonly are negatively skewed, e.g. calcium in limestone or iron in iron ore.

Even if the univariate distributions of a set of coregionalized variables appear to be roughly gaussian, there may be multivariate relationships that clearly are nongaussian. For example, most complete chemical analyses are supposed to sum to approximately 100%, and the sum of any subset of the chemical components in an analysis should have an upper bound of about 100%. Such a constraint cannot hold if the variables have a multivariate normal distribution, even if the marginal distributions of the chemical variables appear to be approximately normal.

Finally, even if both the univariate and multivariate sample distributions of a set of coregionalized data appear to be gaussian, the data still need not be drawn from a gaussian process. Normality of the histograms and scattergrams of k coregionalized variables $\{z_i(x), i = 1 \text{ to } k\}$ does not imply the joint normality of all random variables $\{Z_i(x_j)$ for all i and all j$\}$, or even the joint normality of all pairs of data separated by a vector h, i.e. $\{z_i(x), z_j(x+h)\}$.

Clearly we must either find new methods for simulating a wide variety of nongaussian spatial processes directly, or find methods to transform realizations of nongaussian processes into a form that is at least approximately gaussian. Very little is known about the former approach. The infinite variety of nongaussian processes that are possible seems to preclude a general method for their direct simulation.[51] Furthermore,

---

[51] However, specific kinds of processes may be directly simulated as nonlinear functions of one or more gaussian processes; e.g., Miller and Borgman (1985) provide a method for simulating exponentially distributed properties as the sum of squares of two independent gaussian simulations, and Alfaro (1979) describes the properties of other random functions obtained from squares (gamma process), products (Bessel), and exponentiations (lognormal) of gaussian processes.

their distributions in general would not be preserved in linear combinations, so conditioning by kriging and the use of linear models of coregionalization would be restricted. However, the second approach is more promising. We first need to find an invertible transformation of the data values that will convert the original process into a gaussian process. We can then perform a conditional simulation of the gaussian process, which will reproduce both its mean and its covariance (or mean vector and covariance matrix, in the case of a coregionalization). Because the mean and covariance of a gaussian process determine its entire spatial distribution, this conditionally simulated gaussian process will reproduce the entire spatial distribution of the transformed conditioning data. Then we need only to invert the transformation originally applied to the data to convert the conditional gaussian simulation to a conditional simulation of the original process. Notice that this procedure assumes that a transformation from the original process to a gaussian process indeed exists, and that only the data <u>values</u> (not their locations) need to be transformed to create a gaussian-like realization. Section 3.7.8 points out some common natural phenomena for which this assumption is not valid and also summarizes some other practical hazards encountered in the use of transformations.

It is important to realize that, in theory, we must transform the original process <u>exactly</u> to a gaussian process in order to reproduce the entire spatial distribution of the original phenomenon <u>exactly</u>. Only in the case of a gaussian process will reproduction of the mean and covariance function assure reproduction of the entire spatial distribution. If the distribution is not reproduced exactly, even the

covariance function (or the variogram) of the original process will not necessarily be reproduced exactly by the back-transformed simulation, for the covariance

$$C(x_1,x_2) = E[Z(x_1)\cdot Z(x_2)] - EZ(x_1)\cdot EZ(x_2)$$

involves an integration (within the expectation $E[Z(x_1)\cdot Z(x_2)]$) of the cross product $Z(x_1)\cdot Z(x_2)$ over the bivariate distribution of $Z(x_1)$ and $Z(x_2)$. Therefore, a very useful (but not sufficient) check on the accuracy of a simulation is a check on the reproduction of the original variograms or covariances.

To apply the approach described above rigorously, we would have to infer the spatial distribution of the original random function. This is fundamentally impossible for earth-science data, because we have only one realization of the random process, thus only one sample from the distribution of each random variable $Z_i(x_j)$ at each x_j. But <u>if</u> we can assume strict stationarity and ergodicity (Section 3.7.2), then samples from different points x_j can be combined to construct the histograms and k-variate scattergram of $\{z_i(x), i = 1 \text{ to } k\}$. Furthermore, the joint distributions of all finite sets of random variables $\{Z_i(x_j)$ for all i and j$\}$ will depend only on the distances among the points. Then, in principle, we can infer at least the bivariate distributions between pairs such as $\{Z_1(x),Z_2(x+h)\}$ for various distance and direction classes of vector h by examining the "h scattergrams" (Omre, 1984, Journel, 1984a, and Journel, 1984c) of all pairs $\{(z_1(x_1),z_2(x_2), x_2-x_1=h\}$ that can be found in the data set.

There is an easy way to obtain a quick (but not definitive) check on the binormality of h scattergrams. Computer programs that calculate

sample variogram functions must identify all pairs separated by vector h and accumulate the sum of their squared differences. It is easy to alter such a program to save a vector of the values of these individual differences. If the data come from a stationary gaussian realization, if the spatial domain of the data is large in comparison to the range of the variogram (to assure ergodicity -- see Section 3.7.2), and if the data pairs cover the spatial domain evenly (e.g., randomly or on a grid), then these differences will be normally distributed with mean zero, variance $2\gamma(h)$. If we define the realization within this domain to be a "population" represented by the data, then we can perform a goodness-of-fit (e.g., chi-square) test for normality on the sample distribution of differences saved by the modified program. Checks of this and similar types have been illustrated by de Oliveira Leite (1983).

The "curse of dimensionality". In practice we still cannot reliably infer the multivariate distribution of many coregionalized variables at a large number of vectors h using these simple checks. The difficulty lies in the limited amount of data available. If a given number "n" of data is considered necessary to assess the properties of a univariate distribution via its sample histogram, then assessment of the properties of an m-variate distribution with marginals of the same general type will require roughly n^m data to achieve the same accuracy.[52] This

---

[52] To see why this is so, visualize a histogram of one variable obtained by dividing the domain of the sampled distribution into 10 histogram classes. To obtain a reliable histogram, we must sample until we obtain some minimum number of data falling within each class. For a bivariate scattergram, the domain is a rectangular area, which we would probably want to divide into 10x10 rectangular classes, each filled with at least the same minimum number of data. In general, if n data are adequate to characterize a univariate histogram,

situation is an instance of Bellman's (1961, p. 94) famous "curse of dimensionality". An assessment of the multivariate distribution underlying a coregionalization of k variables, where pairs of data are broken into j distance-and-direction categories (representing various classes of distances and directions of vector h) would thus seem to require about n^{jk} data. Actually the situation is not quite so bad, because in construction of the j sets of h scattergrams, most data will be used several times, and because more efficient methods for multivariate density estimation than the histogram method suggested in the footnote above are available (some references are in the next paragraph). Nevertheless, in a typical two-dimensional (i.e., simple) mining situation we would be very surprised to find more than a few thousand analyses (and usually fewer than that), yet we might have k=10 variables and want n=100 data per variable and j=(8 lags x 4 directions)=32 categories of h, thus requiring between $n^k=10^{20}$ and $n^{jk}=10^{640}$ well placed data for a good assessment of the complete distributional properties of the phenomenon.

The multigaussian hypothesis. It should now be eminently clear that knowledge of the complete spatial distribution of a natural phenomenon is not obtainable strictly from the data. We must either use our knowledge of the physical processes that created the phenomenon to obtain some theoretical insight into the distribution, or just make some "reasonable" hypotheses about the multivariate distribution based upon analysis of a few projections or combinations of the multivariate data. Numerous methods are available for assessing the properties of marginal

---

approximately n^m data will be adequate for an m-dimensional histogram.

distributions and multivariate distributions with relatively few variables (e.g., Section 5.4 of Gnanadesikan, 1977, Chapter 2 of Hand, 1981, and Friedman et al., 1982).

One natural hypothesis to make about a coregionalized spatial phenomenon is that the multivariate distribution among different classes of vector h is either "similar to" or "more normal than" the distribution at h=0. For instance, the scattergram (for h=0) of a set of multivariate chemical analyses might exhibit some univariate or joint lognormal behavior that could be removed by transforming the data to their logarithms. If this transformation succeeded in making the scattergram at h=0 look normal, then it could be hypothesized (especially after a few spot checks) that the multivariate distribution among various values of $h \neq 0$ would be normal as well. Here we have hypothesized that the multivariate distribution at $h \neq 0$ is "similar to" the distribution at h=0.

An analysis at h=0 might also be subject to the constraint that the sum of any number of chemical components in a single chemical analysis cannot exceed 100%. However these same components, some analyzed at one location x and some at another location x+h, would not be strictly subject to this constraint (although they might obey it approximately, especially at small distances |h|), so their joint distribution could be considered "more normal than" (here meaning "less constrained than") the distribution at h=0. We could hypothesize that a transformation removing the constraint at h=0 would have a similar effect at $h \neq 0$, thus transforming the whole spatial distribution into one that is "more normal".

Section 3.5.2 describes a few particular transformations designed to impose normality, or a semblance of it, onto univariate histograms and multivariate scattergrams at h=0. The efficacy of these methods in imposing multivariate normality onto a spatial random process at all h depends upon the validity of the "similar-to" or "more-normal-than" hypothesis. The general hypothesis of normality at h≠0 given normality at h=0 is called the "multigaussian" or "multinormal" hypothesis by Verly (1984a, 1984b) and is the basis for the application of multigaussian (MG) nonlinear estimation methods.

3.5.2 Transformations of Nongaussian Data

> God does not usually tell us from what distribution the data come. [Bratley et al., 1983, p. 124]

3.5.2.1 Requirements

Two requirements must be met before a transformation can be applied to convert a nongaussian data set into an approximately gaussian data set suitable for simulation purposes. (Note that in some cases no such transformation may exist; see Section 3.7.8.)

First, we must know or infer a few necessary characteristics of the data distribution. Some transformations require more knowledge of the distribution than others. Transformations with fixed mathematical forms work properly only if the original data follow particular distributions (e.g., logarithmic transformations are useful only for lognormal-like distributions). Transformations to remove constraints on the domain of a distribution can be applied only if the constraints are known. The stepwise-conditional gaussian transformation described in Section

3.5.2.3 requires knowledge of the complete multivariate density at h=0. Even when only the sample distribution of the raw data is used, as in the univariate normal-scores or "graphical" transformation of Section 3.5.2.2, we must still know or assume that this distribution is representative of the phenomenon to be simulated.

Second, we must insure that the transformation is invertible, or one-to-one. From a practical point of view, this is necessary so that the conditional simulation can be transformed back to a form approximating the original process and the conditioning data can be restored to their original values. From a theoretical point of view, a correct restoration of the simulated process to a form resembling the original process is assured because the original and transformed multivariate distributions determine each other exactly only in the case of a one-to-one transformation. If we transform the original k-variate distribution $F(z_1,z_2,\ldots,z_k)$ to a k-variate gaussian distribution $G(y_1,y_2,\ldots,y_k)$ by applying the transformation $y_i = t_i(z_1,z_2,\ldots,z_k)$, $i = 1,2,\ldots,k$, and then in the simulation we reproduce this transformed distribution, $G_s(y_1,y_2,\ldots,y_k) \cong G(y_1,y_2,\ldots,y_k)$, then upon applying the inverse transformation $z_j = s_j(y_1,y_2,\ldots,y_k)$, $j = 1,2,\ldots,k$, to the simulated data y_i, we will obtain a simulation of the original phenomenon with distribution $F_s(z_1,z_2,\ldots,z_k) \cong F(z_1,z_2,\ldots,z_k)$ that is very close to the original distribution, F.

A small practical difficulty in constructing one-to-one transformations occurs when the sample distribution contains one or more "spikes". A spike is a vertical jump in the cumulative histogram caused by several "tied" data with exactly the same value. This is a common

occurrence when values are rounded off to a small number of digits, when the data are in the form of integer "counts", or when several "zero" or "below-detection-limit" values are present. Tied data can be a problem because some transformations (particularly the normal-scores transformation discussed in Section 3.5.2.2) implicitly involve ranking of the data (independently for each coregionalized variable), followed by a transformation of the ranks. If there is no unique ranking of the original data, there can be no unique transformation to a gaussian distribution, which has no spikes. In practice we must adopt a convention for assigning distinct ranks to the tied data. At one extreme, we can assign the ranks randomly, probably creating a transformed process with an overly high nugget component. At the other extreme, we can assign ranks by kriging each tied value (after deleting it) and then ranking the tied values according to the ranks of their kriged values. This is probably less dangerous, but it runs the risk of creating an overly smooth transformed process. A simpler, and probably less smooth, approximation to the kriging approach has been suggested by Verly (1984a, 1984b). This technique, employed in G. W. Verly's (1984b) subroutine DESPIK, ranks the tied data according to the average of data within a prespecified "zone of influence" of fixed size centered on each tied data point.

In applying the transformations discussed below to a coregionalization, we transform only the multivariate distribution at h=0 (or in the case of univariate gaussian transformations, just the marginal distributions), relying on the multinormal hypothesis of Section 3.5.1 (or on a much stronger hypothesis that marginal

transformations at h=0 effect the proper multivariate transformation) to assure the correct transformation among all h.

3.5.2.2 Univariate Gaussian Transformations

Normal-scores (graphical) transformation: The only method for transforming a nongaussian process to approximate normality that is now in widespread use is the univariate gaussian transformation (gaussian anamorphosis), which is described in Journel and Huijbregts (1978, p. 508). The simplest type of univariate gaussian transformation is the familiar normal-scores transformation, which is employed in some nonparametric or rank-order statistical tests. It corresponds to the "graphical transform" described by Journel and Huijbregts (1978, p. 476). To apply this transformation, we begin with a set of N data $\{z_0(x_i), i = 1 \text{ to } N\}$ from some regionalized variable $z_0(x)$, with sample cumulative distribution function $F^*(z_0)$. These data are assumed to be representative of the population being modeled.[53] The univariate normal-scores transformation simply associates each observed value $z_0(x_i)$ with a standard[54] normal value $u_0(x_i)$ (called the "normal score" of $z_0(x_i)$) drawn from the standard normal distribution function $G(u_0)$, such that

$$F^*(z_0) = G(u_0)$$

---

[53] One must be very careful in making this assumption. Sample distributions can be very sensitive to "domain errors" and "estimation errors", as described in Sections 3.7.2 and 3.7.3. Furthermore, the data may be clustered, the sampling may have been done preferentially, or the support may be heterogeneous. A method for "declustering" data is described by Journel (1983, Appendix A).

[54] "Standard" typically means mean=0, variance=1. For small data sets, the normal scores actually have a variance different from 1.0, because the continuous distribution has been discretized. This is unimportant in the present application.

as illustrated in Figure 29. A conditional simulation based upon the transformed data set is then performed, and finally the simulated values are back-transformed (by a technique described below) to reproduce the original distribution F*. It is emphasized that the distribution being reproduced is the sample distribution F*, not the underlying F. Any peculiarities that happen to appear in F* simply because of sampling variability will be reproduced in the simulated distribution.

The graphical transformation and inverse transformation between data sets having the two distributions F* and G can be performed by Verly's (1984b) subroutines GINV and LINT, possibly with the assistance of his subroutine DESPIK to dissolve spikes in the distribution F*. The forward transformation (GINV) is a straightforward calculation of $u_0=G^{-1}F^*(z_0)$. The inverse of the gaussian distribution function, G^{-1}, is calculated by an approximation formula (Abramowitz and Stegun, 1972, p. 933, Equation 26.2.23). $F^*(z_0)$ is equal to $\text{rank}(z_0)/(N+1)$, N being the number of z_0 data. The smallest z_0 value has rank=1. If many z_0 data are available to define the function F*, and if F* is fairly continuous (no big gaps in the z_0 values), then a very simple linear interpolation (LINT), described below, is adequate for the inverse transformation of the simulated standard gaussian data $u_s(x)$.[55] For a particular value of $u_s(x)$ equal to one of the original data values $u_0(x_i)$, the inverse transformation $z_s=F^{*-1}G(u_s)$ is simply $z_s(x)=z_0(x)$. If the value of $u_s(x)$ does not equal an original value $u_0(x_i)$ but lies between two ordered values, $u_0(x_i)<u_s(x)<u_0(x_j)$, then a linear

---

[55] A modification of this procedure that may speed up execution when large numbers of real and simulated data are involved is demonstrated in Section 4.2.2.6.

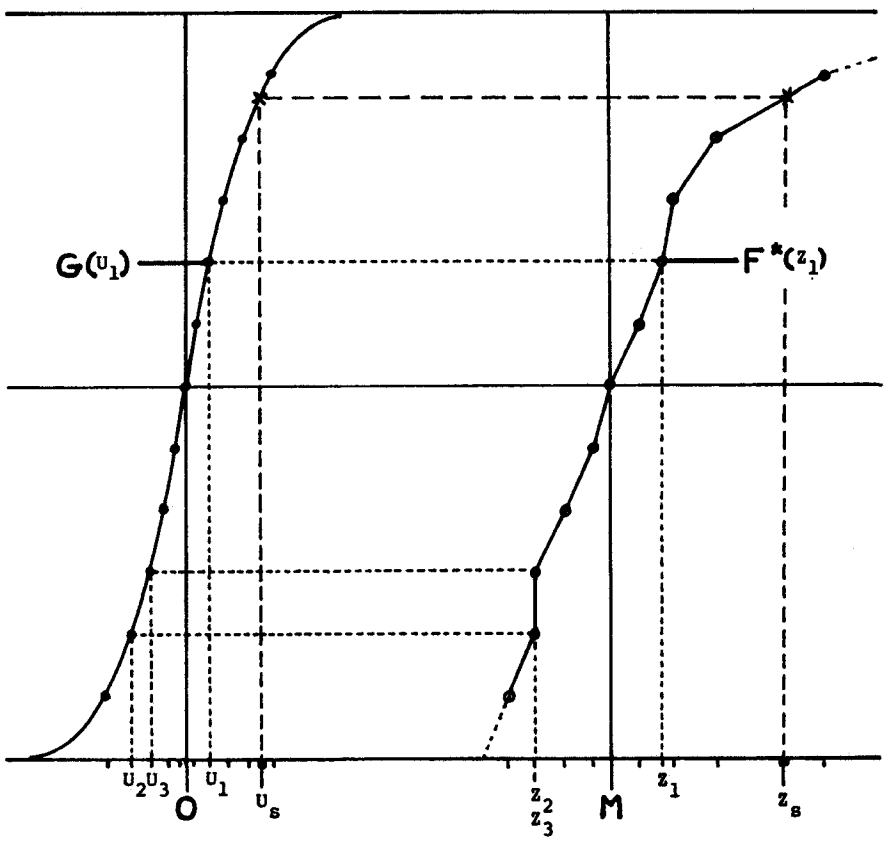


Figure 29: Normal-scores (graphical) univariate gaussian transformation. A straightforward transformation is illustrated by the relationship between datum z_1 and its normal score u_1. Data z_2 and z_3 are tied, so some convention must be adopted to decide which of the scores u_2 and u_3 is assigned to each z. Simulated score u_s could be transformed back to z_s by the means illustrated here, but the linear-interpolation method described in the text effectively substitutes a linear approximation to distribution G in the region of u_s. The median M of the data set is transformed to zero, the median of the standard normal distribution.

interpolation is performed to obtain $z_s(x)$:

$$[z_s(x)-z_0(x_i)]/[z_0(x_j)-z_0(x_i)] = [u_s(x)-u_0(x_i)]/[u_0(x_j)-u_0(x_i)]$$

Extreme values of $u_s(x)$, lying above or below all values of the transformed data set $\{u_0(x_i)\}$ can be (as an approximation) transformed

to the maximum or minimum values of the data set $\{z_0(x_i)\}$ or, preferably, can be interpolated using two prespecified extreme transform pairs, (min z_0, min u_0) and (max z_0, max u_0), chosen such that the probability that a simulated value $u_s(x)$ will fall outside these extremes is very small. Selection of appropriate values for min z_0 and max z_0 may be a very subjective undertaking that requires one to imagine what the extreme tails of the original distribution F(z) might look like.

Other transformations. A few other approaches to univariate gaussian transformations are available. If one can reliably infer the mathematical form of the underlying distribution F(z), it may be possible to transform $z_0(x)$ directly into a gaussian $u_0(x)$, as in the case of the lognormal distribution, for which $u_0(x)=\log z_0(x)$ is normally distributed.[56] This type of transformation avoids the possibly undesirable reproduction in the simulation of the sampling-induced peculiarities of F* that is inherent in the graphical method.

An alternative to the graphical transformation is to fit a Hermite-polynomial expansion to the transformation function $G^{-1}F^*$, as described by Marechal (1976), Journel and Huijbregts (1978, Sections VI.B.3 and VI.B.4), and Dowd (1978, Chapter 3). However, this approach seems needlessly complicated for simulation purposes. A brief illustration of its application in a simulation of coregionalization is provided by Dagbert (1981), and a lengthy example is provided by Dowd (1978, Chapter 3).

---

[56] Journel and Huijbregts (1978, pp. 525-526) describe the mathematical relationships between the structures of u_0 and z_0 for the lognormal case.

There are still other "data-directed" transformations in the literature that might be applied to transform an arbitrary sample distribution to one that is "more normal". An example is the power transformation of Box and Cox (1964). The application of this transformation to multivariate data is described by Gnanadesikan (1977, p. 137).

Hypotheses. If we are simulating only one regionalized variable $z(x)$, the hypothesis underlying the use of a univariate gaussian transformation is that a transformation of the univariate histogram of $\{z_0(x_i)\}$ will result not only in the univariate normality of $\{u_0(x_i)\}$ but also in joint normality among $\{u_0(x+h_j)\}$ for all x and all possible vectors h_j, as we would find in a realization of a gaussian process. This is exactly Verly's (1984a, 1984b) multinormal hypothesis. Although this is a strong hypothesis, practical simulations incorporating this transformation generally seem to have achieved acceptable reproduction of at least the univariate distribution and the variogram of the original data.

If this transformation is applied in a simulation of coregionalization, in which each coregionalized variable $z_j(x)$ is independently transformed to an associated $u_j(x)$, a stronger multinormal hypothesis must be made. We must assume that the univariate normality of all of the marginal regionalized variables $u_j(x)$ implies not only the joint normality of $u_j(x)$ and $u_j(x+h)$ for each j and all h, but also the joint normality of $u_j(x)$ and $u_i(x+h)$, $j \neq i$. Joint normality for $j \neq i$ is a very strong assumption, especially for $h=0$, so it is a good idea to check for joint normality (e.g., using one of the methods described by

Gnanadesikan, 1977, Section 5.4), at least at h=0, before proceeding with a simulation of coregionalization using this approach. The approach is likely to be particularly unsatisfactory at h=0 if multivariate constraints (such as $\sum z_j(x) \leq 100$, for percentage data) exist among the original variables at h=0. Univariate transformations of the marginal distributions may distort, but not eliminate, such constraints, so the transformed variables cannot be multivariate normal. It is preferable to remove these constraints by a prior transformation (Section 3.5.2.4).

3.5.2.3 Multivariate Gaussian Transformations

Because the use of univariate gaussian transformations of only the marginal distributions of a set of coregionalized data requires such a strong hypothesis of multinormality, an obvious next step is to transform the entire multivariate scattergram at h=0 (not just the marginal histograms) to a multivariate normal scattergram. Then we need only to hypothesize that multinormality among the transformed coregionalized variables $u_j(x)$ at h=0 will result in approximate multinormality among $u_j(x+h_i)$ for all regionalized variables u_j and all possible vectors h_i.

Stepwise-conditional transformation. If we know (or can reliably estimate) the multivariate distribution at h=0 of the original data, a multinormal transformation of the scattergram can be performed exactly, using the stepwise-conditional approach proposed by Rosenblatt (1952). This transformation is illustrated in Figure 30. If the original coregionalized random function $Z=(Z_1,Z_2,\ldots,Z_k)$ has a k-variate distribution (at h=0)

$$F(z_1,z_2,\ldots,z_k) = P\{Z_1 \leq z_1, Z_2 \leq z_2, \ldots, Z_k \leq z_k\}$$

then the following stepwise transformations (taken in any order among the k variables)

$$u_1 = F(z_1) = P\{Z_1 \leq z_1\}$$

$$u_2 = F(z_2|z_1) = P\{Z_2 \leq z_2 | Z_1 = z_1\}$$

$$\ldots$$

$$u_k = F(z_k|z_{k-1},z_{k-2},\ldots,z_1)$$

will yield a random function $U=(U_1,U_2,\ldots,U_k)$ that is uniformly distributed (at h=0) on the k-dimensional unit hypercube; i.e., the k components of U are uniformly and independently distributed over the interval [0,1]. Now all we need to do is to apply the standard gaussian inverse transformation G^{-1} (using the method in subroutine GINV) to obtain a new k-variate random function

$$V = (V_1,V_2,\ldots,V_k) = (G^{-1}(U_1),G^{-1}(U_2),\ldots,G^{-1}(U_k))$$

consisting of k independent (at h=0) standard normal variables. If the multinormal hypothesis is valid, this will transform the k coregionalized random functions $Z_1,Z_2,\ldots,Z_k$ with an arbitrary multivariate distribution (at all vectors h) into k standard gaussian random functions $V_1,V_2,\ldots,V_k$. If we are lucky, the independence of the transformed variables at h=0 may further result in independence of the k transformed regionalized variables at all vectors h, making them very easy to simulate. It is important to realize that independence at h=0 may be only "apparent": the multivariate transformation above does not take spatial dependencies into account, so a check of the cross variograms of the components of V is needed. Gaussian random functions with zero cross variograms are strictly independent. (The effects of

transformations on spatial dependence are discussed further in Section 3.5.3.)

This method is so nice that there must be some reason why no one seems to have tried it (although Chiles, 1984, has used a similar approach, described below). Probably the reason is that, in practice, the original multivariate distribution F is not given but must be estimated from the data, and reliable nonparametric multivariate density estimates based only on sample scattergrams tend to require a lot of data (because of the curse of dimensionality -- see Section 3.5.1) to avoid seriously biased estimates.[57] Hand (1981) summarizes several methods of multivariate density estimation, some of which require fewer data than the simple histogram method described in a footnote in Section 3.5.1, but all of which work best with either a lot of data or only a few variables. The best method for the many-variable case may be the projection-pursuit method of Friedman et al. (1982), which was specifically designed to sidestep the "curse".

If we can obtain a reliable estimated density F* for the original k-variate data set, we can apply the transformations described above to obtain samples from k regionalized variables $v_1(x), v_2(x), \ldots, v_k(x)$, which we hope will be multinormal among all h. We then perform a structural analysis of the transformed variables and proceed with a classical conditional simulation of coregionalization. (It is likely

---

[57] Bias (difference between estimated and true density) increases as the number of data decreases. For instance, in the histogram method of density estimation (Section 3.5.1), we divide the domain of the distribution into cells and count the number of data falling within each cell. If there is a shortage of data, we must increase the cell size to get adequate counts. Then any local details of the original density will be smeared out across the large cells, producing a bias in the estimate.

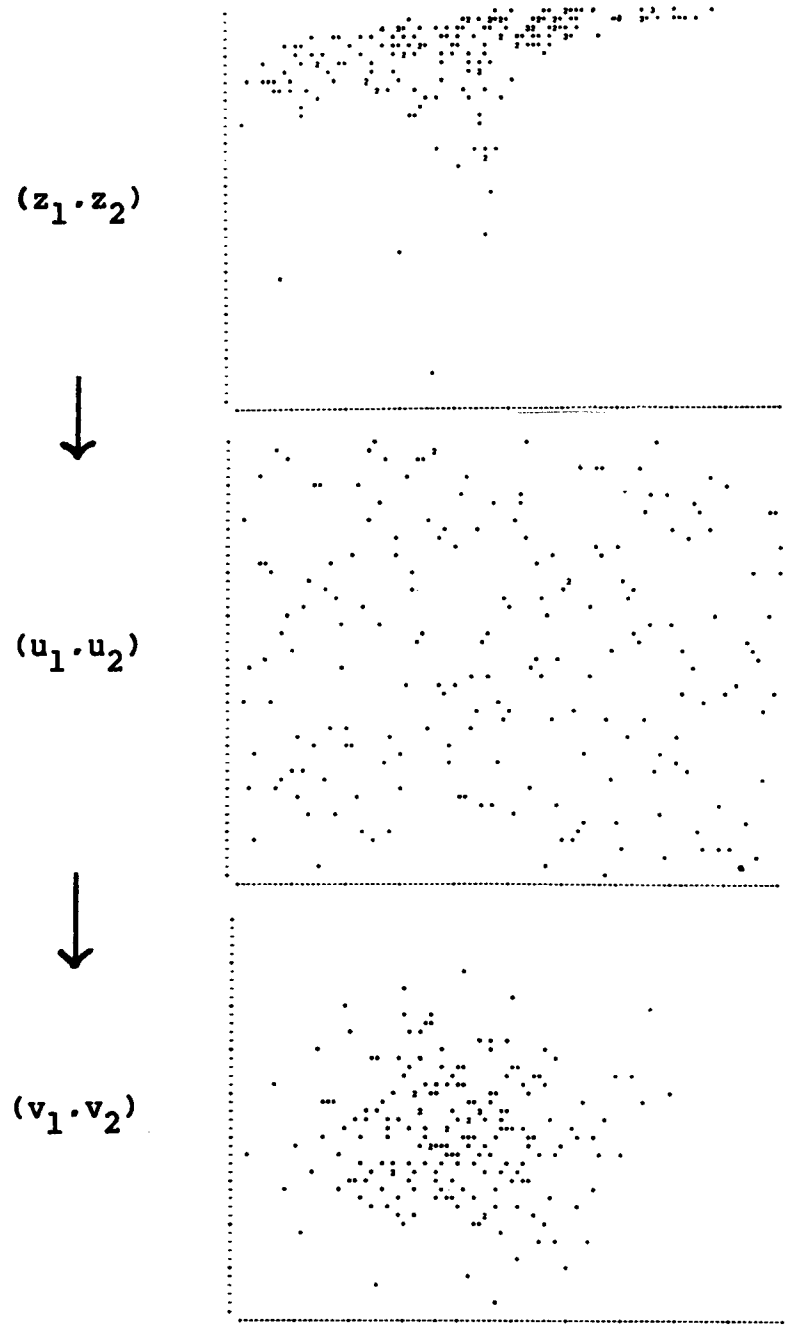


Figure 30: Illustration of the stepwise-conditional multivariate gaussian transformation for some bivariate data. This is a fictitious idealized illustration created by first simulating pairs of independent uniform (0,1) data and attributing their ranks to (u_1, u_2), and then applying arbitrary transformations to get (z_1, z_2) and normal-scores transformations to get (v_1, v_2).

that the linear model of coregionalization will be rather simple, as independence of the transformed data at h=0 may suggest, but not guarantee, zero cross variograms among some variables.) Having simulated the k realizations $v_{js}(x)$, we then apply the G transformation to obtain $u_{js}(x)=G(v_{js}(x))$, and finally invert Rosenblatt's transformation in a stepwise manner (in which the variables u_{js} might be back-transformed in any order, regardless of the order of the forward transformation):

$$z_{1s}(x) = F^{*-1}_{z1}(u_{1s}(x))$$

$$z_{2s}(x) = F^{*-1}_{z2/z1}(u_{2s}(x))$$

. . .

$$z_{ks}(x) = F^{*-1}_{zk/zk-1,..,z1}(u_{ks}(x))$$

Implementation of this approach would involve a lot of computational effort and detail to obtain F* and its associated conditional distributions and transformations, and to assure strict restoration of the values of the original conditioning data. The only advantage is elimination of our reliance on a hypothesis of multinormality at h=0, so in most practical situations, univariate gaussian transformations, accompanied by some checks on multinormality at h=0, will probably suffice. Nevertheless, multivariate transformations of this type seem to be a promising topic for research and may be necessary in the occasional instances in which univariate transformations do not work well.

Other transformations. Some other, less comprehensive, multivariate transformations have been described. A generalization of the Box and Cox (1964) univariate power transformation to simultaneous

transformations of multivariate data is summarized and illustrated with some examples in Gnanadesikan (1977, pp. 137-150 and 241-247), and Chiles (1984) has proposed a linear-regression approach that is similar in spirit to the stepwise-conditional transformation, described above. Chiles' approach, which he applies to a simulation of coregionalization of thicknesses and grades in a nickel-laterite deposit, is particularly useful if the variance of one coregionalized variable appears to depend on the value of another "leading" variable (a dependency that does not exist in a multivariate normal distribution). To transform a bivariate scattergram, Chiles first designates one of the coregionalized variables $z_1(x)$ as the leading variable. He independently transforms $z_1(x)$ and $z_2(x)$ into $u_1(x)$ and $u_2(x)$, respectively, using the univariate normal-scores transformation described previously. $u_2(x)$ is then transformed further to

$$u_2'(x) = [u_2(x) - m(u_1(x))] \,/\, \sigma(u_1(x))$$

where $m(u_1(x))$ is the linear regression of u_2 on the transformed leading variable u_1, and $\sigma(u_1(x))$ is the conditional standard deviation (a function of u_1) of the residuals $[u_2-m(u_1)]$. In practice u_1 and u_2' will be roughly independent at $h=0$, and the variance of u_2' will be independent of u_1. If the cross variogram of u_1 and u_2' is zero for all h and the multinormal hypothesis is valid, these variables can be simulated independently.

3.5.2.4 Transformations to Remove Constraints

Univariate transformations. Most earth-science phenomena cannot rigorously be normally distributed because, among other reasons, the

data are not allowed to vary over the range $[-\infty,+\infty]$ as required for a normal random variable. Percentage data, for instance, are confined to the range [0,100].[58] If the constraints apply only to individual regionalized variables and not to functions (e.g., sums) of several coregionalized variables, they can be neatly removed in one of two ways. (1) A univariate normal-scores transformation will automatically remove the constraints from the transformed variable. To ensure that the interpolated back-transformed simulated data obey the original constraints, simply set the minimum constraint equal to the minimum interpolation-class boundary "min z_0" and the maximum constraint to "max z_0". (2) Before (or perhaps in place of) the gaussian transformation, one can perform a logarithmic or logit transformation on the original data. If z(x) is a positive regionalized variable with minimum constraint "a" and maximum constraint "b", then:

$u(x) = \log(z(x)-a)$	has no minimum constraint,
$u(x) = -\log(b-z(x))$	has no maximum constraint, and
$u(x) = \log[(z(x)-a)/(b-z(x))]$	has neither.

If some of the raw data are <u>exactly</u> equal to the constraints, it will be necessary to adjust the constraints to $[a-\epsilon,b+\epsilon]$, where ϵ is a very small number, to avoid some obvious difficulties with the transformations. Mosteller and Tukey (1977, Chapter 5) provide a guide to these and several other useful transformations (which they call "re-expressions") for various types of raw data.

---

[58] Actually, there are some familiar chemical examples where this is not really true. For instance, a complete chemical analysis of a rock in which all iron is reported as Fe_2O_3 and all sulfur as SO_3 will probably have a total above 100% if the rock contains pyrite (FeS_2). Sampling and analytical errors can also result in totals above 100%.

Multivariate transformations. Many coregionalized variables also have multivariate constraints. The most familiar of these are the constraints [0,100] on sums of percentage data. Complete chemical analyses may be constrained to sum exactly to 100% (perhaps plus a small allowance for error). This type of constraint is treated in detail later in this section. There may also be mineralogical constraints built into chemical analyses of rocks. For instance, the case study of Section 4.2 involves the simulation of a limestone deposit that, on a microscopic level, consists of essentially three phases: calcite ($CaCO_3$), dolomite ($CaMg(CO_3)_2$), and a mixture of clays. The rocks are analyzed only for weight percentages of $CaCO_3$ and $MgCO_3$, so there are four multivariate constraints,

$$0\% \leq CaCO_3 + MgCO_3 \leq 100\% + \epsilon_1, \text{ and}$$

$$0 \leq (MgCO_3)/(CaCO_3+MgCO_3) \leq 0.457 + \epsilon_2,$$

where ϵ_1 and ϵ_2 are tolerances for sampling and analytical errors. The second pair of constraints marks the permissible range in compositions for mixtures of stoichiometric calcite and dolomite. (All analyses contain some $CaCO_3$, so the denominator of the ratio above never reaches zero.) These constraints can be removed in two steps: (1) convert the original variables $z_1=CaCO_3$ and $z_2=MgCO_3$ to two new variables $u_1=CaCO_3+MgCO_3$ and $u_2=MgCO_3/u_1$; (2) apply univariate gaussian transformations to u_1 and u_2 to remove the constraints $[0,100+\epsilon_1]$ and $[0,0.457+\epsilon_2]$, respectively, where $0.457 = MW(MgCO_3)/MW(CaMg(CO_3)_2)$, and "MW" stands for "molecular weight". These transformations are one-to-one, so if the gaussian transformations of $u_1(x)$ and $u_2(x)$ form a multinormal coregionalization (multinormality should be checked at least

at h=0), then a straightforward conditional simulation of coregionalization followed by inversion of the gaussian (first) and constraint (second) transformations should reproduce the entire spatial distribution of z_1 and z_2.

This example illustrates the general approach for removal of multivariate constraints: transform the original coregionalized variables (e.g., $CaCO_3$, $MgCO_3$) into a new invertible set of coregionalized variables consisting of the functions that are to be constrained (e.g., $CaCO_3+MgCO_3$, $MgCO_3/(CaCO_3+MgCO_3)$), then remove the constraints either through the normal-scores back transformation or through some other transformation, such as a logarithmic or logit transformation. As long as all transformations are invertible and inverse transformations are performed in reverse order to the forward transformations, any number of nested transformations can be imaginatively applied to convert a troublesome initial distribution to approximate normality.

Transformations to remove constraints may become more difficult to apply in an unusual situation where the number of constraint functions exceeds the number of coregionalized variables, as the resulting one-to-many multivariate transformation could not be inverted after simulation to recover the original variables. In such a case, we should first make sure that no constraints are redundant (e.g., $CaCO_3 \leq 100\%$ is redundant given $CaCO_3+MgCO_3 \leq 100\%$). Then we might combine some transformations in a piecewise manner, as illustrated by the extreme example in Figure 31. Such a fixup becomes hard to visualize in more than two dimensions (in the variable space), but it seems unlikely that many high-dimensional

coregionalizations of geologic phenomena would embody more recognizable constraint functions than coregionalized variables. The piecewise nature of this transformation might make the transformed realization look nonstationary, as its behavior would change abruptly at locations where the original realization crossed the boundaries between the piecewise zones of influence in Figure 31. A further gaussian transformation probably would smooth out this behavior.

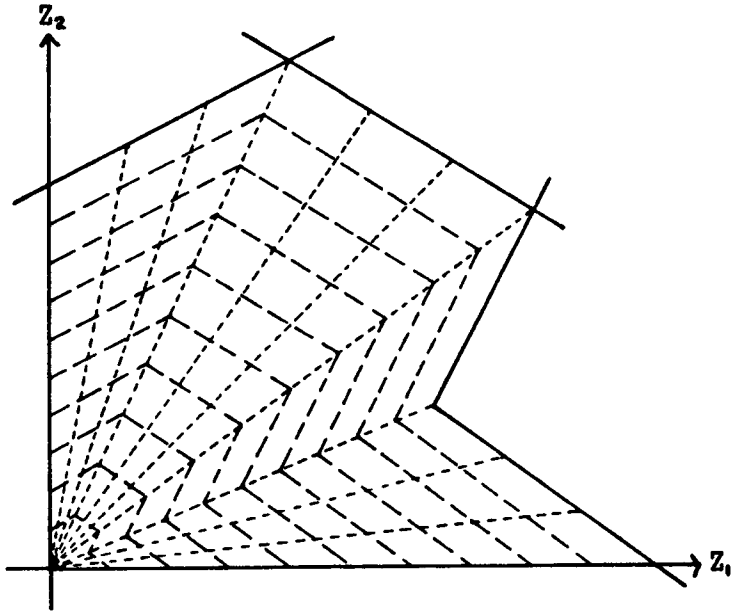


Figure 31: Piecewise transformations for use when constraint functions outnumber the variables. The dashed lines represent the contours of two transformed variables.

Transformations for constant-sum data. The most common constraint encountered in practice is the constant-sum or "closed-array" constraint, which has received an embarrassment of attention in the

literature of mathematical geology and petrology. (One may begin exploring this literature by looking up the references cited by Aitchison (1981), Butler (1981), and LeMaitre (1982).) This constraint applies to percentages (or proportions) that must sum exactly to 100% (or to 1), although the statistical behavior of such data is essentially unchanged if the constraint is only approximate (e.g., because of analytical errors). A set of k-variate percentage data subject to this constraint actually contains only k-1 degrees of freedom, as the kth percentage can be determined by subtracting the sum of the other k-1 percentages from 100%. Geometrically, the data must plot on a (k-1)-dimensional simplex connecting the points (100,0,0,...,0), (0,100,0,...,0), ..., (0,0,...,0,100) in k-dimensional space. As the k percentage variables are effectively "competing for space", an increase in one variable requires a corresponding decrease in one or several of the others. This competition tends to induce negative correlations among the percentages. Most of the many authors who have written on this subject over the years have been groping for ways to decide when observed correlations among the variables are caused by some substantive association among the variables other than this "closure property". The usual approach is to transform the data into some "open" form and examine the correlations among the transformed variables to see if any petrologically significant relationships are revealed.

In conditional simulations we usually do not care <u>how</u> the multivariate associations among a set of coregionalized variables are produced, so long as the relationships that do exist are recognized and reproduced by the simulation. Nevertheless, some transformations that

have been proposed to remove the closure constraint may be useful in transforming closed data into a more nearly normal (unconstrained) form for simulation purposes. For example, Aitchison (1981) has proposed the following transformation: given a k-dimensional set of nonzero percentage data $z=(z_1,z_2,\ldots,z_k)$ where $z_1+z_2+\ldots+z_k=100\%$, we transform the first k-1 data into

$$u_i = \log(z_i/z_k) = \log z_i - \log z_k, \quad i=1,2,\ldots,k-1$$

which have inverses:

$$z_i = 100\ \exp(u_i)/[1+\sum_{j=1}^{k-1}\exp(u_j)], \quad i=1,2,\ldots,k-1$$

These k-1 invertible transformations produce k-1 variables, u_i, that are distributed over $[-\infty,+\infty]$, as required for normality. If the u_i's (or subsequent transformations of the u_i's) are approximately multivariate normal at h=0, and the multinormal hypothesis for $h\neq 0$ is valid, we can conditionally simulate these k-1 regionalized variables, then invert the transformations by the inversion formula above to obtain simulated values for $z_1,z_2,\ldots,z_{k-1}$. The last "simulated" variable, z_k, is calculated by difference and will always be greater than zero, as required, for

$$\sum_{i=1}^{k-1} z_i = 100 \sum_{i=1}^{k-1}\exp(u_i)/[1+\sum_{j=1}^{k-1}\exp(u_j)]$$

$$= 100\ \mathrm{SUM}/(1+\mathrm{SUM})$$

which is less than 100.

Aitchison's transformation is most easily applied if all k original variables z_i are approximately lognormally distributed (an unlikely case, considering the constraint on their sum), for in that case the

marginal distributions of all $u_i = \log z_i - \log z_k$ are normal, so further transformations to impose normality at h=0 are not required. In practice, one could pick z_k to be a lognormal-looking variable, as it appears in every u_i, and then apply a further transformation to normalize any resulting u_i that did not appear to be normal.

In cases where the sums of the original percentage data are not exactly 100%, Aitchison's transformation can still be used by adding a fictitious "remainder" variable if the sum is always less than 100%, or by adding the remainder variable and then raising the constant sum to some value safely above 100%, perhaps 105%, if the actual sum varies both above and below 100%.

There is a simpler alternative to Aitchison's transformation that might be called a "successive-remaining-space" transformation. We begin with a vector z of constant-sum data such that $\sum z_i = 100$, as before, and apply the following k-1 transformations in succession:

$$u_1 = z_1$$

$$u_i = 100\, z_i / \left(100 - \sum_{j=1}^{i-1} z_j\right), \quad i = 2, 3, \ldots, k-1$$

These u_i's are also percentage data, but they are percentages of the "remaining space" in the analysis, after accounting for the i-1 previous variables, rather than the original space of the z_i's. Being percentage data, the u_i's are not likely to be normal-looking and will have to be further transformed such that the percentage constraints $0 \leq u_i \leq 100$ (or some narrower range suggested by inspection of the transformed data) are observed. The sole function of this transformation is to ensure that the final back-transformed simulated data strictly obey the original constant-sum constraint on the z_i's. The simulated u_i's are invertible,

$$z_1 = u_1$$

$$z_i = u_i(100-\sum_{j=1}^{i-1} z_j)/100, \quad i=2,3,\ldots,k-1$$

and the final "simulated" z_k, calculated by difference, must be nonnegative, for each successive simulated z_i is calculated to fill only a fraction $(u_i/100)$ of the remaining space

$$\left(100 - \sum_{j=1}^{i-1} z_j \right)$$

that is left after summation of the previously inverted z_j's, $j<i$.

A potential practical difficulty with this transformation is that each transformed variable u_i includes contributions from all previously transformed z_i's. The many individual nested structures that might end up in the variograms of the last few u_i's would be difficult to observe in the sample variogram plots, with the possible result that the regionalizations of the final z_i's would not be well modeled (see Section 3.7.8). Because the sequence of transformations may influence the quality of the model, it might be a good idea to transform the most "important" variables (economically or geologically) first.

Dowd (1978, 1984) used a remaining-space type of transformation (which he called his "second method") to simulate a coregionalization of bed thicknesses in a petroleum reservoir. The problem in that case was that the individual simulated bed thicknesses had to sum to the simulated total thickness of the section. This was accomplished by first simulating the proportion of the total thickness attributed to each bed by the method described above, then multiplying these proportions by the simulated total thickness to obtain the simulated individual bed thicknesses.

3.5.3 Transformations to Simplify the Covariance Matrix

In the construction of a linear model of coregionalization for k coregionalized variables, it is necessary to fit (k)(k+1)/2 model semivariograms and cross semivariograms and to ensure that all matrices of nested semivariogram functions with the same general structure (i.e., the same shape except for a multiplicative factor, such as the sill, corresponding to the vertical scale of the semivariogram) are positive-definite. This can be a formidable task if k is large. For instance, complete chemical analyses of rocks may include nine or more oxides, requiring the fitting of at least 45 models, possibly with complicated nested structures and anisotropies. Even assuming (outrageously) that sufficient data are available to estimate all of these models and their associated parameters simultaneously and accurately, one must still consider the labor involved in fitting the functions, verifying positive definiteness, and constructing a linear model. Hence it is worthwhile to investigate whether transformations might be made that could simplify the covariance structure that must be modeled and reproduced.

Transformations to remove multivariate constraints (Section 3.5.2.4), particularly constant-sum constraints, commonly will effect some simplification in the covariance structure by removing cross-associations induced by the constraints. But in most real situations, cross correlations among the variables are likely to persist even in gaussian-transformed data, although the stepwise-conditional gaussian transformation introduced on page 165 has the nice property of removing cross correlations among the original variables at h=0. This transformation unfortunately requires a very good estimate of the

multivariate distribution at h=0, and if the number of variables is large this estimate may be as hard (or harder) to obtain than the variogram estimates for all h. Moreover, independence at h=0 does not imply independence for all h, although in practice some of the cross variograms among the transformed variables probably would be practically zero, thus simplifying the linear model of coregionalization.

Principal-components transformation. A simpler transformation with a similar effect at h=0 is the principal-components transformation. This is an invertible multivariate linear transformation of the original variables z_i, i = 1 to k, of the form

$$y_j = \sum_{i=1}^{k} a_{ij} z_i, \quad j=1,2,\ldots,k$$

in which the coefficients a_{ij} are selected in such a way that the cross covariances $C_{\mu\nu}(0)$, $\mu \neq \nu$, among the transformed variables are zero. As these covariances are the cross-semivariogram sill values of the new variables, many of the cross semivariograms $\gamma_{\mu\nu}(h)$ probably will be close to zero for all h.[59] If the original scattergram at h=0 is approximately k-variate normal, the multinormal hypothesis for h≠0 would suggest that each transformed variable y_μ is independent of any other transformed variable y_ν for which $\gamma_{\mu\nu}(h)=0$ for all h. If this were true among all k coregionalized variables, we could just simulate each transformed variable independently, requiring the modeling of only k

---

[59] Principal-components computer programs normally use the sample variances and covariances of the data to obtain the components. For geostatistical applications, it is usually preferable to work with a matrix of vertical-scale parameters of the variograms (the semivariogram sills, if they exist) estimated from the data rather than the sample variances and covariances, which are actually only estimates of the dispersion variances and covariances of the data within the finite spatial domain that has been sampled.

direct semivariogram functions instead of (k)(k+1)/2 direct and cross semivariograms. In general it will not be true for all $\gamma_{\mu\nu}$, but some simplification of the covariance structure and its associated linear model is still likely. The effects of principal-components transformations on semivariogram models are discussed in more detail further on in this section.

The principal-components transformation and its associated inverse actually constitute one particular linear model of coregionalization, in which only the matrix of total covariances (not those of the nested structures)[60] is used to derive the coefficients of the model. In theory, this model should provide a poorer fit than a painstakingly constructed linear model of all of the various nested structures involved in the coregionalization. In practice, the results might be just as good, but only if the structure is simple (such as an intrinsic coregionalization) or if deficiencies in the data will not allow a comprehensive set of linear models to be confidently constructed.

Principal-components analysis and the related techniques of factor analysis and correspondence analysis are described in many textbooks on multivariate statistics. These techniques have a long history of applications in the earth sciences, many of them reviewed by Joreskog et al. (1976). Geostatistical applications are more recent and still few in number. Borgman and Frahme (1976) used principal components to approximate a coregionalization of eleven measured properties of Wyoming bentonites using only five roughly independent regionalizations of the five largest components; Bryan and Roghani (1982) reduced twelve

---

[60] But see the last paragraph in this section for a refinement.

variables to five components for use in uranium reserve estimation; and Davis and Greenes (1983) kriged the principal components of ash, BTU, and sulfur in a study of coal quality. Dagbert (1981) simulated a coregionalization of seven major-element concentrations in a quartz-diorite pluton using seven principal components, one of which alone accounted for 68% of the sum of variances in the original data. The marginal distributions of the original variables were normalized by univariate gaussian transformations (using the Hermite method) before application of the principal-components transformation, and conditioning was performed directly on the simulated components using the components of the original conditioning data.

Many statistical computer packages containing principal-components routines are widely available. Davis (1973, p. 494) and Cooley and Lohnes (1971, p. 116) also provide listings of simple programs and associated subroutines to perform principal-components transformations.

To transform a random vector $Z=[z_1,z_2,\ldots,z_k]'$ to its principal components $Y=[y_1,y_2,\ldots,y_k]'$, we multiply Z by the transpose of a kxk matrix $A=[A_1,A_2,\ldots,A_j,\ldots,A_k]$ of column vectors A_j, so $Y=A'Z$. Each y_j has the form:

$$y_j = a_{1j}z_1+a_{2j}z_2+\ldots+a_{kj}z_k = A_j'Z$$

If A is an orthogonal matrix, such that $A'A=AA'=I$ (where I is the identity matrix), then the elements a_{ij} of each column vector A_j are scaled such that

$$A_j'A_j = \sum_{i=1}^{k} a^2_{ij} = 1$$

and $A'_\mu A_\nu = 0$, $\mu \neq \nu$. This particular scaling preserves the sum of variances of the data,

$$\sum_{i=1}^{k} \mathrm{Var}(z_i) = \sum_{j=1}^{k} \mathrm{Var}(y_j)$$

while leaving the y's uncorrelated.

The result of this transformation is that the k original variables Z, with positive-definite symmetric covariance matrix

$$C = \begin{bmatrix} c_{11}(0) & & & \\ c_{21}(0) & c_{22}(0) & & \\ \cdots & \cdots & \cdots & \\ c_{k1}(0) & c_{k2}(0) & \cdots & c_{kk}(0) \end{bmatrix}$$

are transformed into k new variables Y having a simpler covariance structure

$$L = \begin{bmatrix} \lambda_1(0) & & & \\ 0 & \lambda_2(0) & & \\ \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & \lambda_k(0) \end{bmatrix}$$

and satisfying the convenient condition that:

$$\sum_{j=1}^{k} \lambda_j(0) = \sum_{i=1}^{k} c_{ii}(0)$$

The variances λ_j of the principal components y_j are the eigenvalues of the covariance matrix C. By convention, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$. The matrix A is the associated matrix of eigenvectors (the column vectors A_j), which form an orthogonal basis for the k-dimensional variable space if the λ's are distinct.

The components y_j and their associated eigenvectors A_j can be calculated sequentially. To find $y_1 = A_1'Z$, we choose A_1 so that

$\lambda_1=Var(y_1)=A_1'CA_1$ is maximized, subject to the constraint $A_1'A_1=1$. If Z is a multinormal random vector, we are actually finding the direction (eigenvector) in the space R^k of the variables that corresponds to the largest axis of the hyperellipsoidal point cloud formed by the k-dimensional data. A Lagrange procedure similar to the one used to solve for kriging weights in Section 3.1.4 can be used here to solve for A_1. The second principal component $y_2=A_2'Z$ must have the largest possible variance λ_2 of all possible combinations $A_2'Z$ that are orthogonal to Y_1 and satisfy the scaling $A_2'A_2=1$. The Lagrange solution to this problem thus maximizes $Var(y_2)=A_2'CA_2$ subject to the scale constraint $A_2'A_2=1$, and subject to the orthogonality constraint that $Cov(y_2,y_1)=A_2'CA_1$ is zero, or more simply, that the eigenvectors A_1 and A_2 are orthogonal: $A_2'A_1=0$. For each subsequent principal component that we extract, we must add another orthogonality constraint. The principal components $y_1,\ldots,y_k$ that eventually emerge from this process are uncorrelated and have progressively decreasing variances, which add up to the sum of variances of the original Z data. The covariance matrices C and L are related by $L=A'CA$ or equivalently $C=ALA'$ (because $AA'=I$). If Z is a multinormal vector, the k eigenvectors A_j will correspond to the k axial directions of the hyperellipsoidal point cloud of Z, and the standard deviations $\sqrt{\lambda_j}$ will be proportional to the lengths of the axes.

The simple transformation $Y=A'Z$ in general will result in principal components Y that do not have zero means. Most programs actually perform the transformation $Y=A'(Z-M)$, where M is the vector of sample means of Z. After simulation of the components, each simulated vector

Y_s will have the form $Y_s=A'(Z_s-M)$. Recalling that $AA'=I$, we can back-transform the simulated Y_s values as follows:

$$A\ Y_s + M = A\ A'\ (Z_s - M) + M = Z_s - M + M = Z_s$$

Notice that a small eigenvalue λ_j corresponds to a component y_j with very small variance, so little information may be lost in replacing such a y_j by its mean, which is zero if the transformation $Y=A'(Z-M)$ has been adopted. This means that, if the last few eigenvalues λ_j, $m<j\leq k$, are close to zero, we need only to simulate the first m components, Y_{ms}, filling in the remaining k-m elements with zero, and later apply the inverse transformation:

$$Z_{ms} = A\ Y_{ms} + M \simeq Z_s$$

This will restore the proper mean to the simulated values, but the variances will be decreased slightly because the last few (tiny) components were deleted. Also, the inverse transformation will impose an exact linear dependency among the Z_{ms}, instead of the near dependency implied by the near-zero eigenvalues. This transformation is not one-to-one, so the distribution is not being reproduced exactly by the simulation, even in the multigaussian case. Moreover, if one of the deleted components happens to be weighted heavily in the expression of an economically important variable, the simulation of that variable may be unacceptably smooth. Consequently, although principal-components analysis is used in many disciplines primarily to reduce the dimensionality of the variable space, for simulation purposes it is better to retain all principal components unless simulation costs are a major consideration.

The principal-components transformation constructs new variables whose cross covariances are constrained to be zero only at h=0, whereas for simulation purposes it is preferable for the cross covariances to be zero at <u>all</u> h. Unfortunately, a zero cross covariance at h=0 does not generally imply a zero cross covariance at all h -- not even for transition models, in which the cross covariance at h=0 is the cross-semivariogram sill. To see this, consider a coregionalization of variables z_α, α = 1 to k, with a matrix of (possibly nested) transition-model semivariograms, which are denoted $\gamma_{\alpha\beta}(h)$. If we construct principal components of the form

$$y_\mu = \sum_{\alpha=1}^{k} a_{\alpha\mu} z_\alpha \qquad \mu = 1,2,\ldots,k$$

then the components will have semivariograms of the form

$$\gamma_{\mu\nu}(h) = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} a_{\alpha\mu} a_{\beta\nu} \gamma_{\alpha\beta}(h)$$

which we would like to be zero at all h, for $\mu \neq \nu$, even though only $C_{\mu\nu}(0) = \gamma_{\mu\nu}(\infty) = 0$, $\mu \neq \nu$, is assured by the transformation. Because the structures $\gamma_{\alpha\beta}$ are transition models, the components will have transition models as well (usually nested), and will have cross semivariograms $\gamma_{\mu\nu}$ with the general appearance depicted in Figure 32. The function in Figure 32 is zero by definition at h=0, and it has also been set to zero at h=∞ by the transformation. For this function to be zero at <u>all</u> h, it must not not exhibit any hole effects. Unfortunately, a coregionalization of arbitrary monotonic transition models does not necessarily yield principal components whose cross variograms are free of hole effects, as the following simple example will illustrate.

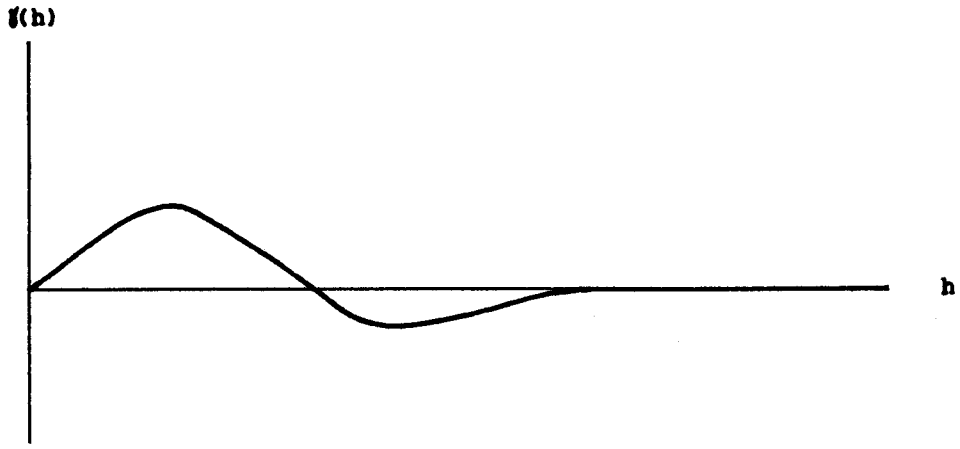


Figure 32: A cross semivariogram between two principal components constructed from transition-model variables.

Suppose we have the simple situation illustrated in Figure 33(a): two coregionalized time series z_1 and z_2 with nested simple-transitive semivariogram models. The positive-definite sill matrices are given below:

Variable	Range a_1		Range a_2		Total Sills	
z_1	5	-6	5	0	10	-6
z_2	-6	8	0	2	-6	10

We shall perform a principal-components transformation using the matrix C of total sills. For this simple matrix the eigenvalues and eigenvectors are easily obtained without resorting to the Lagrange procedure described previously. To get the eigenvalues, we solve for the two values λ_1 and λ_2 for which the determinant of $(C-\lambda I)$ is zero:

$$\begin{vmatrix} (10-\lambda) & -6 \\ -6 & (10-\lambda) \end{vmatrix} = (10-\lambda)^2 - (-6)^2 = \lambda^2 - 20\lambda + 64$$

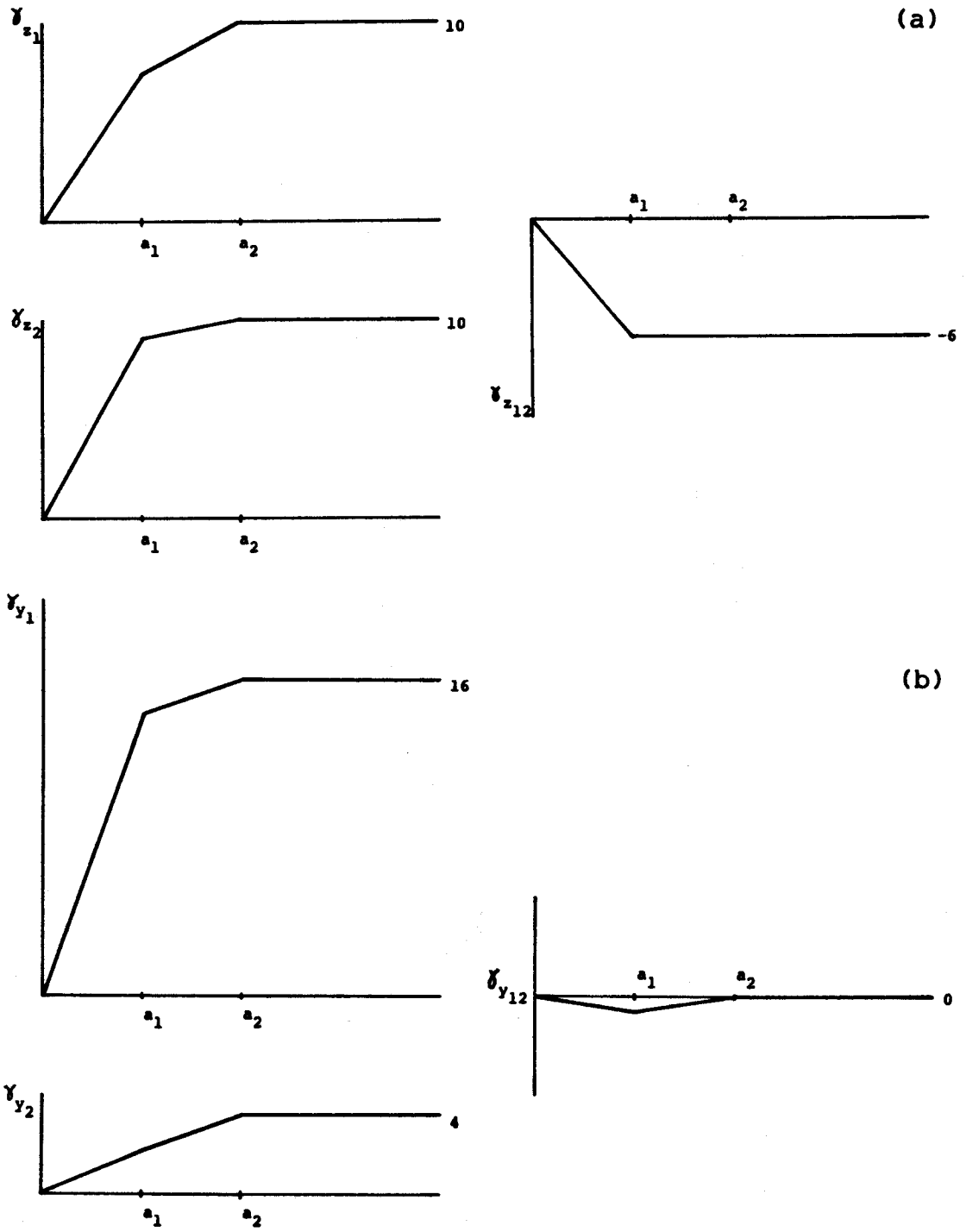


Figure 33: Application of a principal-components transformation to a simple coregionalization of transition-model variables. (a) Semivariograms of the original variables. (b) Semivariograms of the principal components.

So the eigenvalues are the two roots of $\lambda^2-20\lambda+64 = 0$; i.e., $\lambda_1=16$ and $\lambda_2=4$. These roots sum to the sum of variances as required: 16+4 = 10+10 = 20. Each eigenvalue λ_j has an associated eigenvector A_j for which $A_j'A_j=1$. The eigenvector is obtained by solving $(C-\lambda_j I)A_j=0$ for A_j, subject to $A_j'A_j=1$:

$$\text{For } \lambda_1=16,\ (C-\lambda_1 I)A_1 = \begin{bmatrix} -6 & -6 \\ -6 & -6 \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

yielding $A_1'=[a_{11},a_{21}]$ = any vector of the form $[x,-x]$. We want a solution for which $A_1'A_1=1$, i.e. $A_1'=[\sqrt{\frac{1}{2}},-\sqrt{\frac{1}{2}}]$.

$$\text{For } \lambda_1=4,\ (C-\lambda_2 I)A_2 = \begin{bmatrix} 6 & -6 \\ -6 & 6 \end{bmatrix} \cdot \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

yielding $A_2'=[a_{12},a_{22}]$ = any vector of the form $[x,x]$. We want a solution for which $A_2'A_2=1$, i.e. $A_2'=[\sqrt{\frac{1}{2}},\sqrt{\frac{1}{2}}]$. The two eigenvectors are orthogonal: $A_2'A_1=0$. Furthermore the matrix $A=[A_1,A_2]$ is orthogonal: $AA'=I$.

The principal components are thus

$$y_1 = a_{11}z_1+a_{21}z_2 = \sqrt{\tfrac{1}{2}}\ z_1 - \sqrt{\tfrac{1}{2}}\ z_2$$

$$y_2 = a_{12}z_1+a_{22}z_2 = \sqrt{\tfrac{1}{2}}\ z_1 + \sqrt{\tfrac{1}{2}}\ z_2$$

and their direct and cross semivariograms, applying the formula on page 186, are:

$$\gamma_{y1}(h) = \tfrac{1}{2}\gamma_{z1}(h) + \tfrac{1}{2}\gamma_{z2}(h) - \gamma_{z12}(h)$$

$$\gamma_{y2}(h) = \tfrac{1}{2}\gamma_{z1}(h) + \tfrac{1}{2}\gamma_{z2}(h) + \gamma_{z12}(h)$$

$$\gamma_{y12}(h) = \tfrac{1}{2}\gamma_{z1}(h) - \tfrac{1}{2}\gamma_{z2}(h) + 0$$

The positive-definite sill matrices of the principal components are given below, and their semivariograms are presented in Figure33(b).

Variable	Range a_1		Range a_2		Total Sills.	
y_1	$12\frac{1}{2}$	$-1\frac{1}{2}$	$3\frac{1}{2}$	$1\frac{1}{2}$	16	0
y_2	$-1\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	$3\frac{1}{2}$	0	4

Notice that the cross semivariogram γ_{y12} displays a slight hole effect that might, on a sample plot, be difficult to recognize. Notice also that the individual matrices of nested components have been made more complicated by this transformation -- the opposite of our original intention.

If the structures $\gamma_{\alpha\beta}(h)$ are in intrinsic coregionalization (Journel and Huijbregts, 1978, p. 174), i.e., $\gamma_{\alpha\beta}(h)=b_{\alpha\beta}\gamma_0(h)$ for all α and β, where $[b_{\alpha\beta}]$ is a positive-definite matrix of constants, then

$$\gamma_{\mu\nu}(h) = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} a_{\alpha\mu} a_{\beta\nu} b_{\alpha\beta} \gamma_0(h)$$

for all μ and ν. Then if we set $\gamma_{\mu\nu}(h)$ to zero for all h, the form of the common factor $\gamma_0(h)$ is irrelevant, as it can be divided out of the equation. Thus, in the particular case of intrinsic coregionalization, we can use principal components to reduce the cross variograms to zero at all h, simply by working with the matrix of coefficients $b_{\alpha\beta}$. Even a linear γ_0 would suffice, as the matrix of semivariogram slopes $b_{\alpha\beta}$, rather than the matrix of sills $\gamma_{\alpha\beta}(\infty)$ used for transition models, would be subjected to the eigenvalue analysis. The slopes of the linear cross semivariograms of the principal components would be set to zero. There is commonly a great deal of subjectivity involved in fitting a matrix of direct and cross semivariograms to sample plots, so it pays to fit an intrinsic-coregionalization model where possible.

Intrinsic coregionalization is only a sufficient condition for zero cross variograms among principal components: conceivably, more complicated sets of nested structures could also have this property, at least approximately. One does not have to depend on inspection of noisy sample semivariograms of the components to judge whether their models should be zero, as the formula on page 186 can be used to plot the component models precisely.

Applications in structural analysis. Principal-components analysis has a useful application in structural analysis and the formulation of a linear model for a coregionalization. If we first perform just the initial steps of a structural analysis by hand -- i.e., we separate the matrix of total semivariogram functions into its additive matrices of nested structures -- we can then use a principal-components program to do the rest automatically. It will check the positive definiteness of the individual sill matrices (by checking that all eigenvalues are positive), and it will derive a linear model from the eigenvectors. Because the individual matrices of nested structures are intrinsically coregionalized, the cross variograms of the nested structures will be zero at all h. This linear independence (zero cross variograms) among the components of the linear model is the essential property that allows us to construct simulations of coregionalization from sets of independently simulated basic structures. In the presence of "good conditioning", only the short-scale structures of the direct and cross semivariograms must be modeled carefully for simulation purposes, as explained in Section 3.7.5. Thus in many cases it should be possible to reduce the number of nested intrinsic coregionalizations that must be

modeled to one matrix of nugget structures and another matrix of transition models to represent variogram behavior near the origin.

3.6 CONDITIONAL SIMULATION OF COREGIONALIZATION FOR TRANSFORMED PROCESSES

This short section summarizes some extensions to the procedure outlined in Section 3.4 for the simulation of gaussian processes. Several useful transformations of nongaussian and many-variable data have been proposed in Section 3.5, all with the aim of enabling one to simulate very complicated multivariate processes using the simple unconditional gaussian simulation methods of Section 3.3. Now we must establish how these transformations relate to the procedure outlined in Figure 28 of Section 3.4, and how they relate to one another.

There is a fairly unambiguous hierarchy of transformations that determines the proper order in which they should be carried out. Figure 34 depicts an expanded general procedure for simulating nongaussian spatial processes. The hierarchy of transformations can be summarized as follows:

(1) Transformations of the raw data to remove constraints from the scattergrams at h=0 should usually be performed first. Otherwise, other transformations (particularly nonlinear transformations with no definite algebraic form, such as the univariate normal-scores or multivariate stepwise-conditional transformations) might distort the existing constraints beyond recognition, making their subsequent removal difficult.

(2) Transformations to normalize nongaussian scattergrams should be performed next. Multivariate transformations that better assure

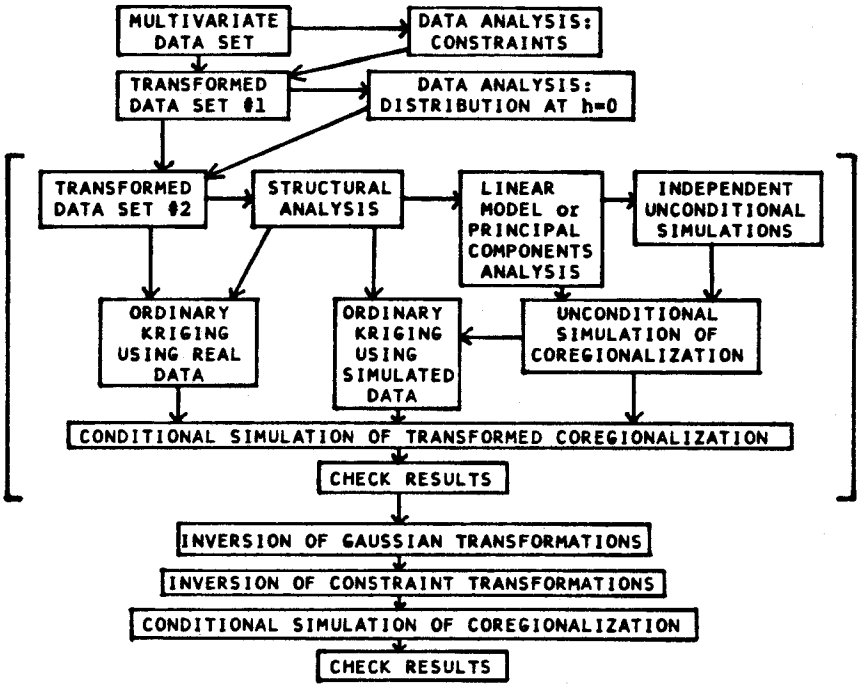


Figure 34: Conditional simulation of coregionalization for transformed processes. This chart shows the steps involved in a conditional simulation of several coregionalized nongaussian random functions. In this version, conditioning is done on transformed (approximately gaussian) cross-correlated data using ordinary krjging. Figure 35 shows an alternate arrangement of the central (bracketed) part of this chart (corresponding to all of the chart in Figure 28).

multivariate normality of the scattergrams are preferable in theory but are computationally messy and commonly suppose either better knowledge of the starting distribution than is likely to exist or more data for estimation of the distribution than are likely to be available. Univariate normal-scores transformations of the marginal distributions, followed by some informal checks on multivariate normality, are likely to be the method of choice for most practical simulations.

(3) Principal-components transformations to simplify the overall covariance structure are worthwhile only when the number of variables is large (say at least four) and it is difficult to infer a linear model from the starting data. Any transformation to impose gaussian behavior onto a nongaussian data set should be performed before the principal-components transformation, as the components are geometrically interpretable in the gaussian case. Furthermore, if the lack of correlations among the components at h=0 also produces zero cross variograms at all h, the components will be strictly independent in the gaussian case, making the use of independent simulations easier to justify. If the principal-components transformation does not appear to reduce all cross variograms to zero, it is at least likely to reduce several of them to zero, so that a linear model accounting for the cross structures that remain will be easier to construct. A principal-components analysis applied to the individual matrices of nested structures can be used to construct a linear model automatically, as described at the end of Section 3.5.3.

After conditioning and checking of results for the conditional simulation of coregionalization of the gaussian-transformed process, the normalizing and constraint-removal transformations should be inverted in reverse order. Finally, the variograms, scattergrams, histograms, constraints, and any other noteworthy characteristics of the original data should be compared with the corresponding characteristics of the completely back-transformed conditional simulation of coregionalization to see that all important features of the phenomenon have been preserved.

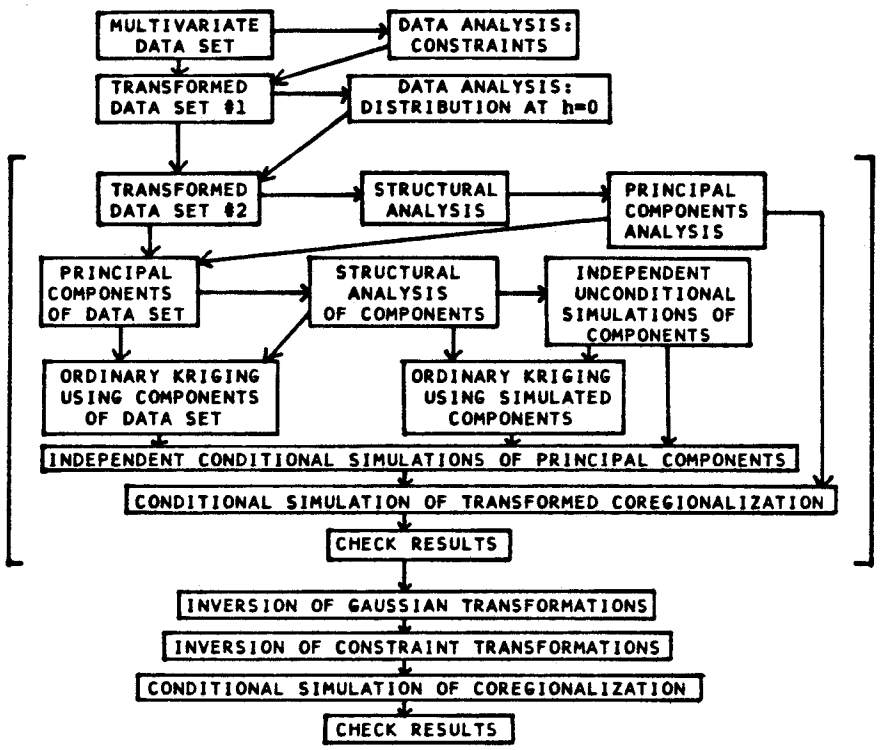


Figure 35: Conditional simulation of coregionalization with conditioning on principal components. The central part of this chart is a revision of the central part of Figure 34 to allow conditioning of the principal-component scores by ordinary kriging.

Conditioning should always be done using fully transformed conditioning data, as the entire distribution of the phenomenon is preserved during conditioning only in the multigaussian case. Conditioning may be done either on principal-component data or on the reconstituted dependent gaussian data. Conditioning using the principal components of the conditioning data by ordinary kriging actually

incorporates some of the potential advantages of conditioning the dependent data by cokriging, as the principal components include contributions from all of the original coregionalized variables. However, the precise cokriging model usually is not being duplicated here, as only the total covariances C(0) of the scattergram at h=0, not the whole cross-variogram structure of the coregionalization, are being used to determine the contributions of other variables to the estimation of the kriged variable. Figures 34 and 35 show both options for conditioning: conditioning after reconstitution of the cross-correlated gaussian variables, and conditioning on principal components.

3.7 GUIDELINES FOR AVOIDING BIAS IN SIMULATIONS

> A certain [simulation] model becomes more accurate as it becomes more expensive. [Mantoglou and Wilson, 1981, p. 35]

This section provides some practical advice on appropriate choices of models, methods, and parameters for conditional simulations, and offers some comments on the sensitivity of simulation results to these choices. For simplicity, much of the discussion that follows makes reference only to simulations of a single regionalized variable $z_s(x)$.[61] Simulations of several coregionalized variables deserve much the same advice, plus a few additional caveats that are brought up mainly in Section 3.7.5. The discussion in Sections 3.7.1 through 3.7.7 is directed primarily toward the minimization of errors that can occur in the creation of a gaussian conditional simulation. A host of additional errors, described mostly

---

[61] If we are considering a conditional simulation, the notation $z_{sc}(x)$ that is used in Section 3.1.7 corresponds to the shorter notation $z_s(x)$ that is used in Section 3.7.

in Section 3.7.8, may creep into a nongaussian simulation if the multigaussian hypothesis is not valid for the transformed data.

3.7.1 The Total Simulation Error and Its Components

The objective of a conditional simulation of a gaussian process is to reproduce the characteristics of a real regionalized phenomenon $z_0(x)$ (or of a coregionalization) in a simulated phenomenon $z_s(x)$. We depend almost entirely on a finite set of data $\{z_0(x_i), x_i \text{ in } V_0\}$, collected within a finite continuous spatial domain V_0 in the real space R^n, to infer the properties of $z_0(x)$ (or at least to infer the mean, variogram, and histogram) and to condition the simulation.[62] The continuous spatial domain V_s in R^n of the simulated data set $\{z_s(x_j), x_j \text{ in } V_s\}$ may not coincide exactly with the domain of observation V_0: it may be a subset of V_0, a larger domain surrounding V_0, a partly overlapping domain, or even a disjoint domain. The boundaries of the continuous domains V_0 and V_s are somewhat arbitrary, although it is convenient in many situations to define them as the convex hulls around $\{x_i \text{ in } V_0\}$ and $\{x_j \text{ in } V_s\}$, as in Figure 36. This definition is not always appropriate in the case of V_0, as there may be large regions inside the convex hull that contain no data.

We would like the properties of $\{z_s(x_j), x_j \text{ in } V_s\}$ to be as close as possible to the properties of $\{z_0(x), \text{all } x \text{ in } V_s\}$. In other words, if Q is any quantity or function of interest (e.g., the mean over some volume, the local variogram function calculated within that volume, the

---

[62] In some situations the data sets used for estimation and conditioning may not be exactly the same. In this section they are assumed to be the same, for simplicity.

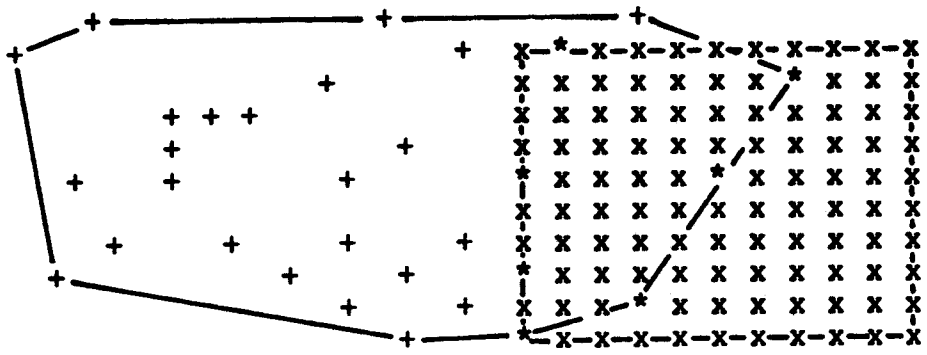
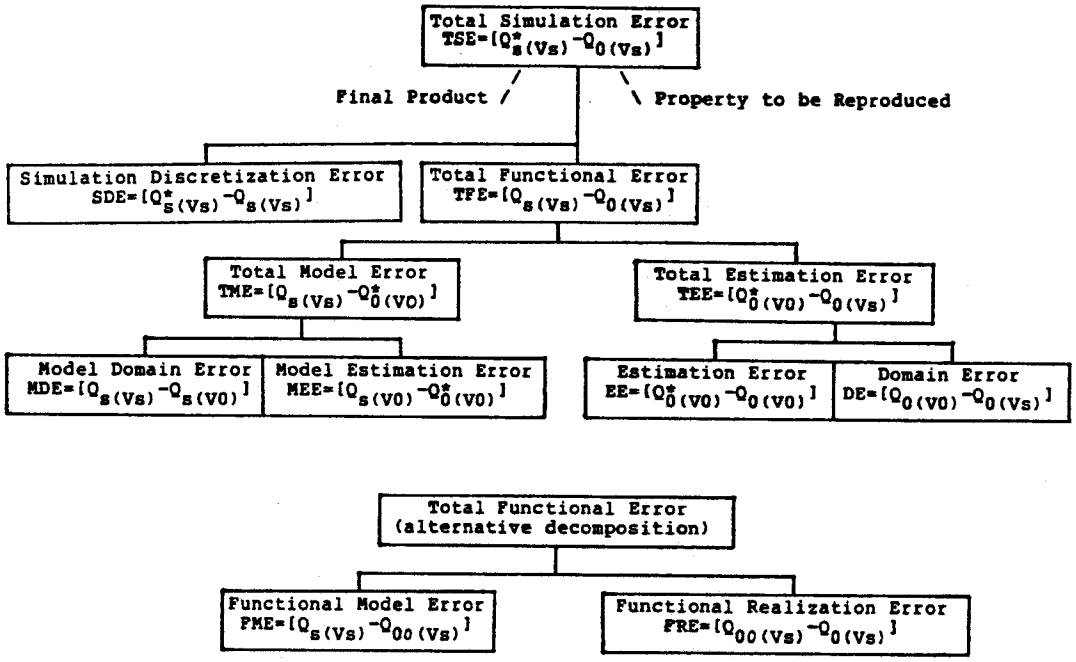


Figure 36: Spatial domains of conditioning data and simulated data. Continuous domains V_0 and V_s of the conditioning data set $\{z_0(x_i),\ x_i \text{ in } V_0\}$ and simulated data set $\{z_s(x_j),\ x_j \text{ in } V_s\}$, for an example in the real cartesian space R^2. $\{x_i \text{ in } V_0\}$ are marked by "+"; $\{x_j \text{ in } V_s\}$ are marked by "x"; $\{x_i = x_j \text{ in } V_0 \text{ and } V_s\}$ are marked by "*". For convenience, V_0 and V_s are defined to be the regions bounded by the convex hulls around $\{x_i\}$ and $\{x_j\}$, respectively.

probability density function of the regionalized variable z(x) within the volume, or the realization z(x) itself), we would like the simulated $Q^*_{s(vs)}$, measured from the discrete (denoted by *) simulated (denoted by $_s$) data $z_s(x_j)$ in volume V_s (denoted by $_{(vs)}$) to be as close as possible to the true but unknown quantity $Q_{0(vs)}$ that would be measured from the real (denoted by $_0$) phenomenon $z_0(x)$. The error $[Q^*_{s(vs)} - Q_{0(vs)}]$ will be called the "total simulation error", or TSE. It is composed of several nested subsidiary errors, summarized in Figure 37. Continual reference to this figure will be helpful during the reading of the following paragraphs.



Additive relationships:

TSE = SDE + TFE

= SDE + [TME + TEE] = SDE + [FME + FRE]

= SDE + [(MDE+MEE) + (EE+DE)]

Notation:

- Q is some quantity (parameter, function) of interest;
- * denotes a discrete sample from a continuous realization;
- 0 denotes the real regionalized phenomenon z_0, regarded as a realization of the random function Z;
- 00 denotes another regionalized phenomenon z_{00}, regarded as another realization of the random function Z;
- s denotes the simulated regionalized phenomenon z_s, similar to but not exactly the same as a realization of the random function Z;
- (v0) denotes the continuous domain of observation (and conditioning) V_0;
- (vs) denotes the continuous domain of simulation V_s.

Figure 37: Sources of error contributing to differences between a real gaussian phenomenon and a simulation of it.

First, we can decompose TSE into two additive components. One is the "simulation discretization error" SDE, which is the difference between the quantity $Q^*_{s(vs)}$ calculated from the grid of discrete simulation data and the ideal $Q_{s(vs)}$ that would be obtained from an infinitely dense grid simulated by an "exact" method (e.g., a method that does not crudely approximate the covariance function or some other characteristic). The nature of this error will be explained below in more detail. The other component of TSE is the "total functional error" TFE, which is merely the difference between this ideal $Q_{s(vs)}$ and the unknown true quantity $Q_{0(vs)}$. In theory, TFE should be merely the error committed by substituting the simulated realization $z_s(x)$ of the random function $Z(x)$ for the actual realization $z_0(x)$; this is the only difference that is contemplated in the basic theory of conditional simulation presented in Section 3.1.7. But in practice, the simulated phenomenon $z_s(x)$ may not be a realization of $Z(x)$, because we actually do not know the true properties of $Z(x)$, or even of $z_0(x)$. We have only estimated these properties from a finite data set $\{z_0(x_i), x_i \text{ in } V_0\}$, and then fitted a usually biased or oversimplified model to the estimate. Thus, in practice, the error between $Q_{s(vs)}$ and $Q_{0(vs)}$ includes much more than just the difference between two realizations of the same $Z(x)$.

We can decompose TFE in two different ways, both illustrated in Figure 37. The "functional realization error" FRE is just the difference that can occur between one realization $z_0(x)$ of the random function $Z(x)$ and another realization $z_{00}(x)$ of the same random function. The remaining error is the "functional model error" FME,

which accounts for the additional departure from $z_0(x)$ caused by estimation and fitting errors. The statistical properties of FRE characterize the minimum obtainable TFE, corresponding to perfect estimation and modeling of the random function $Z(x)$ underlying the realization $z_0(x)$. Unfortunately the properties of FME and FRE cannot be measured in practice, so this decomposition of TFE is not suited to much further analysis.

To obtain a more useful decomposition of TFE, we can examine the procedure that is actually used to obtain a model for generating $z_s(x)$. We first obtain estimates of the relevant properties of $\{z_0(x), \text{ all } x \text{ in } V_s\}$, e.g. the property $Q_{0(vs)}$, by examining the corresponding properties of the data set, $\{z_0(x_i), x_i \text{ in } V_0\}$, e.g. the property $Q^*_{0(v0)}$. The difference between these two is the "total estimation error" TEE. This error is composed of two components: the "estimation error" $EE=[Q^*_{0(v0)}-Q_{0(v0)}]$, which is simply the error committed when replacing the parameter $Q_{0(v0)}$ by a statistic $Q^*_{0(v0)}$ calculated from a finite sample, and the "domain error" $DE=[Q_{0(v0)}-Q_{0(vs)}]$, which is the error made when the domain of observation (and conditioning) V_0 and the domain of simulation V_s do not coincide.[63]

Once we have obtained an estimate $Q^*_{0(v0)}$, involving the commission of a total estimation error, we still must construct a simulation model, entailing the commission of a "total model error" $TME=[Q_{s(vs)}-Q^*_{0(v0)}]$,

---

[63] If the quantity Q is the semivariogram function $\gamma(h)$, the errors EE and DE are conceptually similar to the errors associated with the "variance of estimation" and "fluctuation variance" of local semivariograms. The only difference is that the domain V_s is replaced by the entire space R^n in the definition of fluctuation variance (Journel and Huijbregts, 1978, p. 192).

which is also composed of two errors. The "model estimation error" $MEE=[Q_{s(v0)}-Q^*_{0(v0)}]$ represents the error made in fitting an ideal model (such as a smooth positive-definite semivariogram model) to an estimate from a discrete sample. The "model domain error" $MDE=[Q_{s(vs)}-Q_{s(v0)}]$ is analogous to the domain error $[Q_{0(v0)}-Q_{0(vs)}]$ considered earlier, but it is not exactly the same, because $z_s(x)$ may not differ from one spatial domain to another in the same manner as $z_0(x)$.

Notice that in decomposing the total functional error TFE, we encounter estimation and domain errors twice: once in a continuous-to-discrete direction (TEE) as we construct a sample estimate $Q^*_{0(v0)}$ of $Q_{0(vs)}$, and once in a discrete-to-continuous direction (TME) as we infer a continuous model $Q_{s(vs)}$ from the estimate $Q^*_{0(v0)}$. Both of the associated estimation errors, EE and MEE, can be reduced by obtaining more closely and evenly spaced data within the domain of observation V_0. Both of the domain errors, DE and MDE, can be eliminated completely if the domain of observation V_0 completely contains, or at least coincides with, the domain of simulation V_s.[64] In summary, the total functional error TFE, which is probably the more important component of the total simulation error TSE in most practical situations, can be reduced to insignificance by following one simple (but expensive) rule: collect a densely spaced data set, evenly covering all of the domain of simulation V_s.

---

[64] For good conditioning, V_0 should extend beyond V_s in all directions. However, the properties of $\{z_0(x)$, all x in $V_s\}$ are best measured using only data within V_s, if enough such data are available.

The other component of the total estimation error is the "simulation discretization error" SDE. Ideally, this inevitable error would be attributable solely to the effects of discretizing the continuous domain V_s into a finite grid of points. In this case, the error could be reduced by merely decreasing the grid spacings and, of course, by extending the grid over the whole domain V_s. However, there can be another component of discretization error: the error that is made by using the short cuts or approximations built into almost all simulation procedures. This component of SDE also arises primarily because of coarse discretization (but of a different type) and is difficult to extricate from the more straightforward part of SDE explained above, as both usually involve interrelated choices of parameters for the unconditional-simulation step. Approximations that could give rise to noticeable errors of this type include the use of too few Poisson points to perform a random-average simulation, the use of too few lines in a two-dimensional turning-bands simulation, the use of too few discrete harmonics in a spectral simulation, or even the use of too small a kriging neighborhood at the conditioning stage.

Now that the main sources of error in the simulation of a gaussian process have been identified, let's examine exactly <u>how</u> the departures occur in order to develop ways for selecting appropriate simulation methods and parameters.

3.7.2 Domain Errors (DE and MDE)

To appreciate the importance of the domain error DE, consider the time series depicted in Figure 19, Section 3.1.7. Suppose that we divided the time domain shown in Figure 19(a) into its first, middle, and last thirds, and looked at the statistics of each third compared to those of the whole. Clearly the mean of the first third is much higher than that of the last, and the middle third is somewhere in between. The middle third also shows a marked trend, so its semivariogram would probably exhibit a linear or concave-upward form with no visible sill, whereas the semivariograms of the first and last thirds might have sills at fairly low levels (although the tiny number of sample points involved would probably result in messy semivariograms in all three cases -- contributing to a high estimation error EE[$\gamma(h)$]). The histograms, and any other summarization of the three samples that we could make, would also differ strongly from one another. Obviously then, a simulation over one of these domains would poorly reproduce the characteristics of the time series in that domain if we relied solely on the statistics (and conditioning data) of another domain.

The difficulty here is that the three time domains, in addition to being disjoint and having a small number of discrete sample points (thus producing high estimation errors), are also very short in comparison to the range of the variogram for this process. In fact the range is over half the length of each subdomain, based on the sample semivariogram in Figure 20(a), Section 3.1.7. (The semivariograms in Figure 20 are calculated over a longer domain than the one covered by the sample plots in Figure 19.) Within these small subdomains the random process does

not have room to exhibit its "average" properties in a single realization. A finite realization observed over such a restricted domain is dominated by local "hills and valleys" in the realization. Means will be unstable and variograms will tend to exhibit drifts and hole effects not characteristic of the real process Z(x) or of any realization z(x) observed over a large domain. If we observed for a period of time that was a large multiple of the range, the behavior of the process (if it is stationary at least up to the order of the statistics we are calculating) would be much better estimated by the sample.

Ergodicity. This leads us to the concept of "ergodicity". Informally, a random process is said to be ergodic if the statistics of a single realization of the process, observed over a finite spatial domain (time domain for the example used here), converge to their corresponding expected values as the size of the domain of observation increases. For instance, if the example in Figure 19 represents a realization of an ergodic process, then the sample mean, variogram, and other statistics that we calculate from the data will converge to the true parameters (stationary expected value, positive-definite variogram function, etc.) of the process as the time domain of observation increases.

The convergence of different types of sample functions may be guaranteed by different kinds of requirements on the random process. For instance, a sequence of sample means

$$m(T) = (1/T) \sum_{t=1}^{T} z(t)$$

for a discrete time series Z(t) is said to be ergodic if

$$\lim_{T\to\infty} \mathrm{Var}[m(T)] = 0$$

(Parzen, 1962, p. 73). A sufficient condition for this to be true is that the covariance function C(h) exists and tends to zero as h tends to $\pm\infty$. (In other words, the variogram must have a sill.)

Convergence of the sample variogram is a more difficult problem, requiring knowledge of the fourth moments of the random function Z(x). In a simulation project, it is vitally important that the variogram of the simulated data,

$$2\gamma^*_{s(vs)}(h) = (1/N_s(h))\sum_{j=1}^{N_s(h)}[z_s(x_j+h)-z_s(x_j)]^2$$

where $N_s(h)$ is the number of data pairs $[(x_j+h),(x_j)]$ located within the domain of simulation V_s, be as close as possible to the unknown "local variogram" of the phenomenon of interest,

$$2\gamma_{0(vs)}(h) = (1/V_s(h))\int_{V_s(h)}[z_0(x+h)-z_0(x)]^2dx$$

where $V_s(h)$ is the intersection of the domain V_s with its translation by vector -h (Figure 38). In other words, we want to keep the total simulation error $TSE[\gamma(h)] = [\gamma^*_{s(vs)}(h)-\gamma_{0(vs)}(h)]$ as small as practically possible. If the semivariograms $\gamma_0(h)$ and $\gamma_s(h)$ that we observe and simulate, respectively, differ greatly between domains V_0 (observation and conditioning) and V_s (simulation), this error may be unacceptably large because of large domain errors $DE[\gamma_0(h)]$ and $MDE[\gamma_s(h)]$. These errors are closely related to the "fluctuation variances" (Journel and Huijbregts, 1978, p. 192) of the two semivariogram models,

$$\mathrm{Var}[\gamma_{(v)}(h)] = E\{[\gamma(h)-\gamma_{(v)}(h)]^2\}$$

where $\gamma_{(v)}(h)$ is the true local semivariogram over the continuous domain V, and $\gamma(h)$ is the true semivariogram for the theoretical underlying random function (or for a regionalized variable observed over a practically infinite domain, if the process is ergodic). $\gamma(h)$ is the expected value of $\gamma_{(v)}(h)$.

The variances associated with domain errors depend on the spatial proximity of the two domains; e.g., under second-order stationarity, $DE[\gamma_0(h)] = [\gamma_{0(v0)}(h)-\gamma_{0(vs)}(h)]$ has a variance equal to:

$$\mathrm{Var}\{DE[\gamma(h)]\} = \mathrm{Var}\,[\gamma_{0(v0)}(h)] + \mathrm{Var}[\gamma_{0(vs)}(h)] - 2E\{[\gamma_{0(v0)}(h)-\gamma_0(h)][\gamma_{0(vs)}(h)-\gamma_0(h)]\}$$

The last expectation above is a "fluctuation covariance". If the two domains are about the same size and they essentially coincide, this last term will be approximately equal to the sum of the first two, so $\mathrm{Var}\{DE[\gamma(h)]\} \approx 0$. If they are about the same size but lie far apart, the last term will be roughly zero, and $\mathrm{Var}\{DE[\gamma(h)]\} \approx 2\ \mathrm{Var}\{DE[\gamma_{0(v0)}(h)]\} \approx 2\ \mathrm{Var}[\gamma_{0(vs)}(h)]$.

The domain error and fluctuation variance of the semivariogram are functions of the vector h. In practice, <u>unconditional</u> simulations tend to reproduce their model semivariogram functions rather poorly for large values of h, particularly if the distance |h| is large in comparison to the size of the simulation domain. This should not be surprising, as the domain of integration $V_s(h)$ for the local variogram (Figure 38) is very small in this case, thus providing a poor estimate of the behavior of $\gamma_s(h)$ for the random function, and thus a large fluctuation variance. A similar problem occurs in estimation of the variogram function from a data set, resulting in the "practical rule" (Journel and Huijbregts,

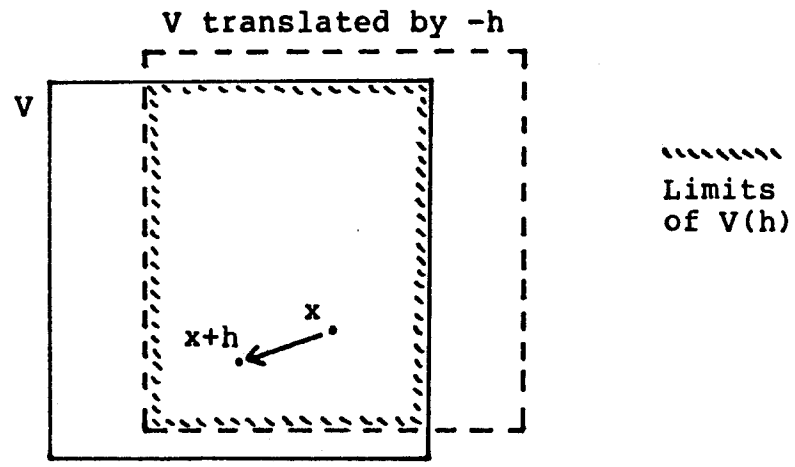


Figure 38: A continuous spatial domain V and the corresponding domain of integration V(h) over which a local variogram can be defined. Integration can be carried out over all points x for which both x and x+h lie within V. These points x define domain V(h), which is equivalent to the intersection of V with its translation by -h.

1978, p. 194) that inferences of semivariogram models for random functions (but not local semivariograms) are only valid for $|h| < L/2$, where L is the dimension of the sampled volume in the direction of h.

To see how sensitive to domain or fluctuation errors a local variogram can be, compare the sample semivariograms plotted in Figure 40. The sample semivariogram in Figure 40(a) was calculated from a series of 320 simulated time-series data, shown in Figure 39(a). The semivariogram model used for this unconditional moving-average simulation was simple-transitive with range=10, sill=10, so the simulation involved merely generating a series of independent standard normal data and performing a running sum of groups of ten data. As this

is a densely and evenly spaced data set, the estimation error $[\gamma^*_{0(320)}(h)-\gamma_{0(320)}(h)]$ over the field of 320 data should be very small, except for values of h close to 320, where the number of data pairs would be small. (More about this in Section 3.7.3.) So most of the fluctuation of the sample semivariogram about the model should be attributable to the fluctuation error $[\gamma_{s(320)}(h)-\gamma_s(h)]$. The model, plotted as a solid line in Figure 40(a), is well reproduced near the origin but rather poorly reproduced at large h, where there are obvious hole effects. Even though the domain is 32 times the range, the local variogram is nevertheless affected by local hills and dales in the simulated realization.

This effect becomes clearer when we cut the realization of 320 data into two halves, consisting of the first 160 consecutive data, and then the rest. Now each half covers only sixteen times the range. The semivariograms of the two halves are presented in Figure 40(b). Both show "pseudoperiodicities", but they are quite different in detail. Still, the behavior near the origin is well reproduced.

Finally, Figure 40(c) shows four sample semivariograms representing the first, second, third, and fourth groups of eighty consecutive data. Each point on these plots still is estimated by at least twenty data pairs, so the marked departures from the model are largely the result of fluctuation errors. Now consider the unfortunate consequences of unconditionally simulating the local realization in the second group of eighty data using only a model inferred from the sample semivariogram of the first group of eighty data. The very large differences $[\gamma^*_{s(1st\ 80)}(h) - \gamma^*_{s(2nd\ 80)}(h)]$ would be mostly the result of domain error, $[\gamma_{s(1st\ 80)}(h) - \gamma_{s(2nd\ 80)}(h)]$.

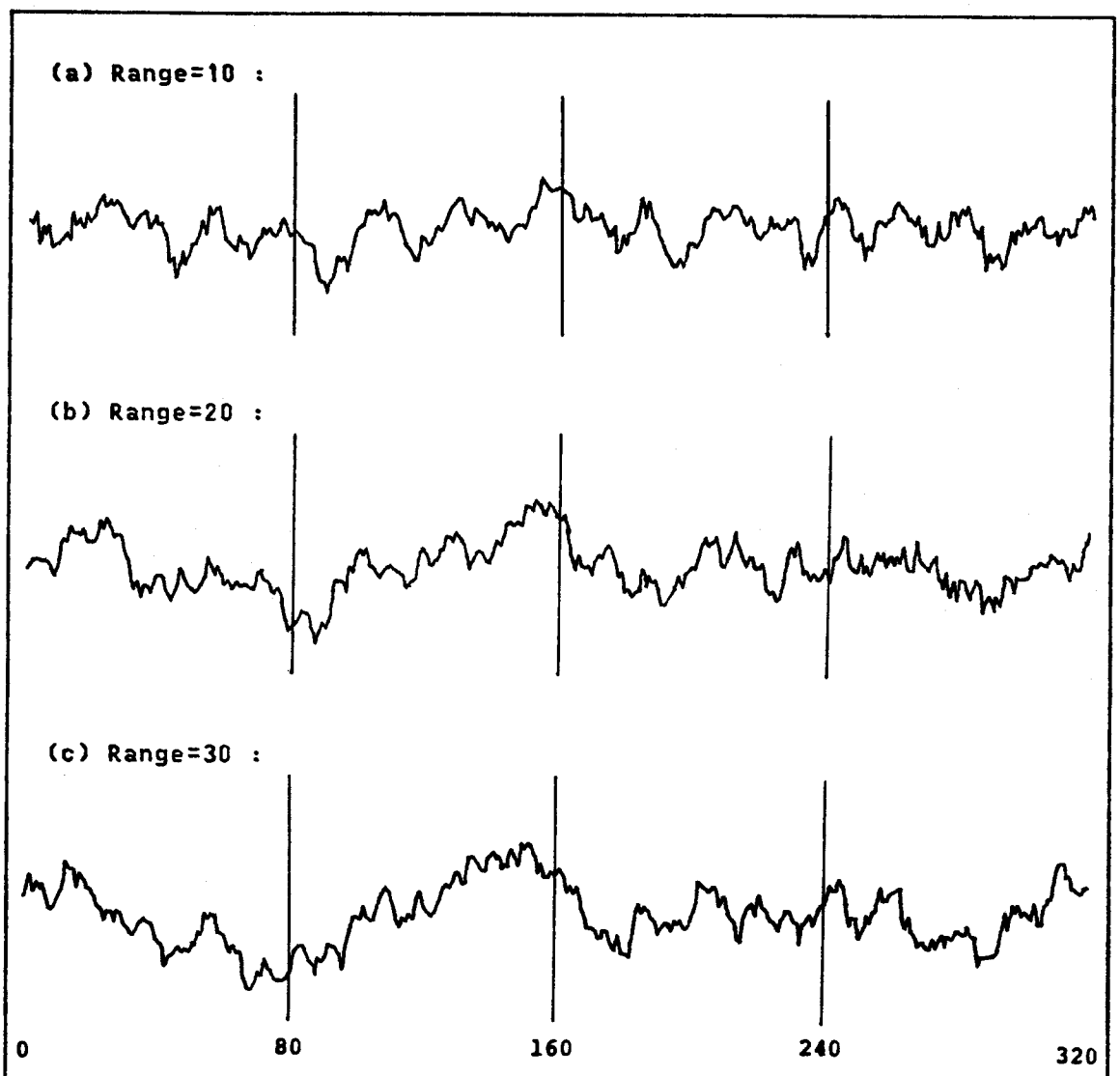


Figure 39: Three simulated time series with variogram ranges of ten, twenty, and thirty. Semivariograms and histograms of these series and of subsets of them are displayed in the next four figures. All semivariogram models are simple-transitive.

The simulations depicted in Figure 39(b) and (c) were performed using the same set of independent random numbers (but necessarily somewhat longer sequences) that was used for simulation (a), but running sums of

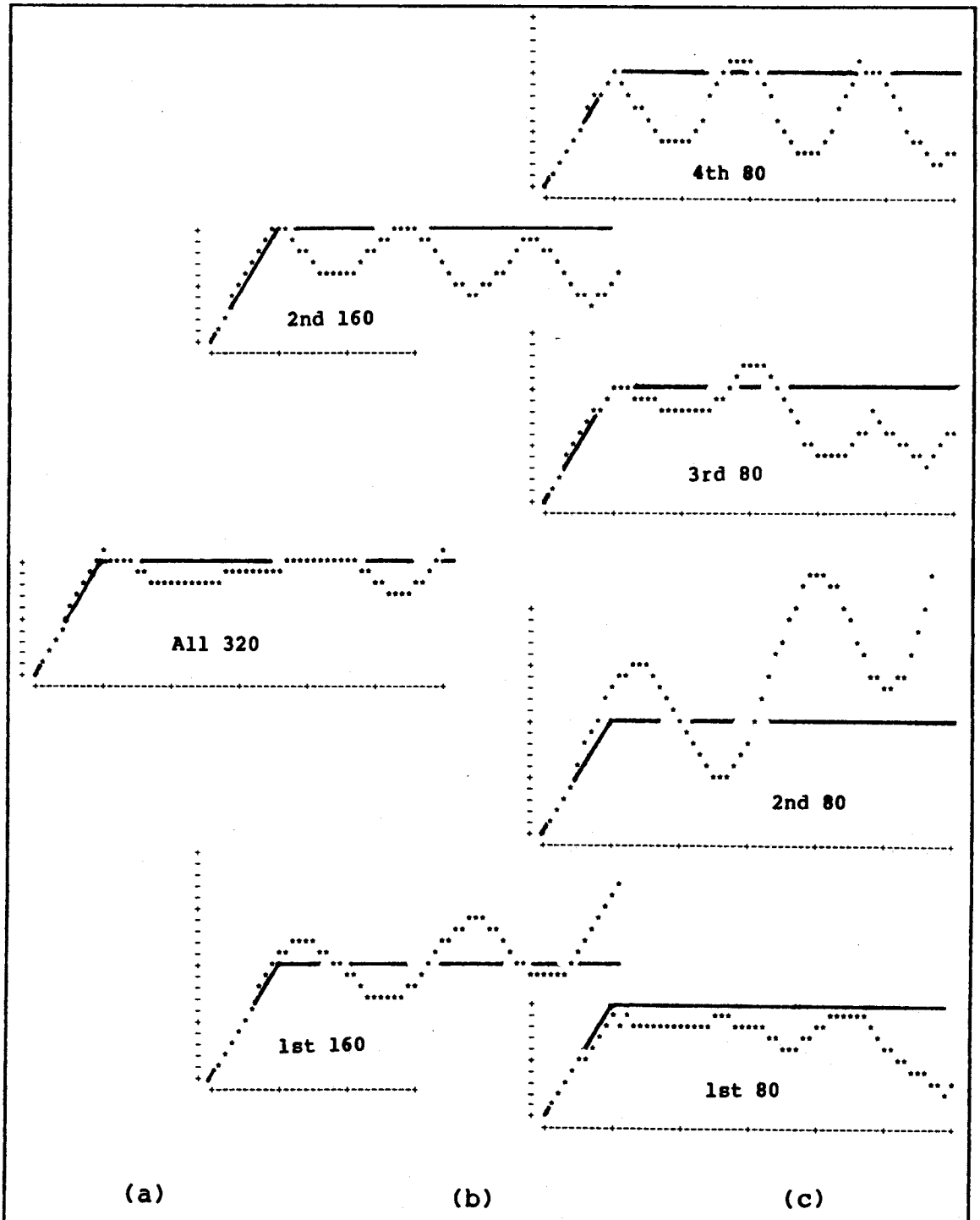


Figure 40: Semivariograms of parts of the simulated time series in Figure 39(a). (a) Semivariogram of all 320 data. (b) Semivariograms of the first and second series of 160 data. (c) Semivariograms of the first, second, third, and fourth series of 80 data.

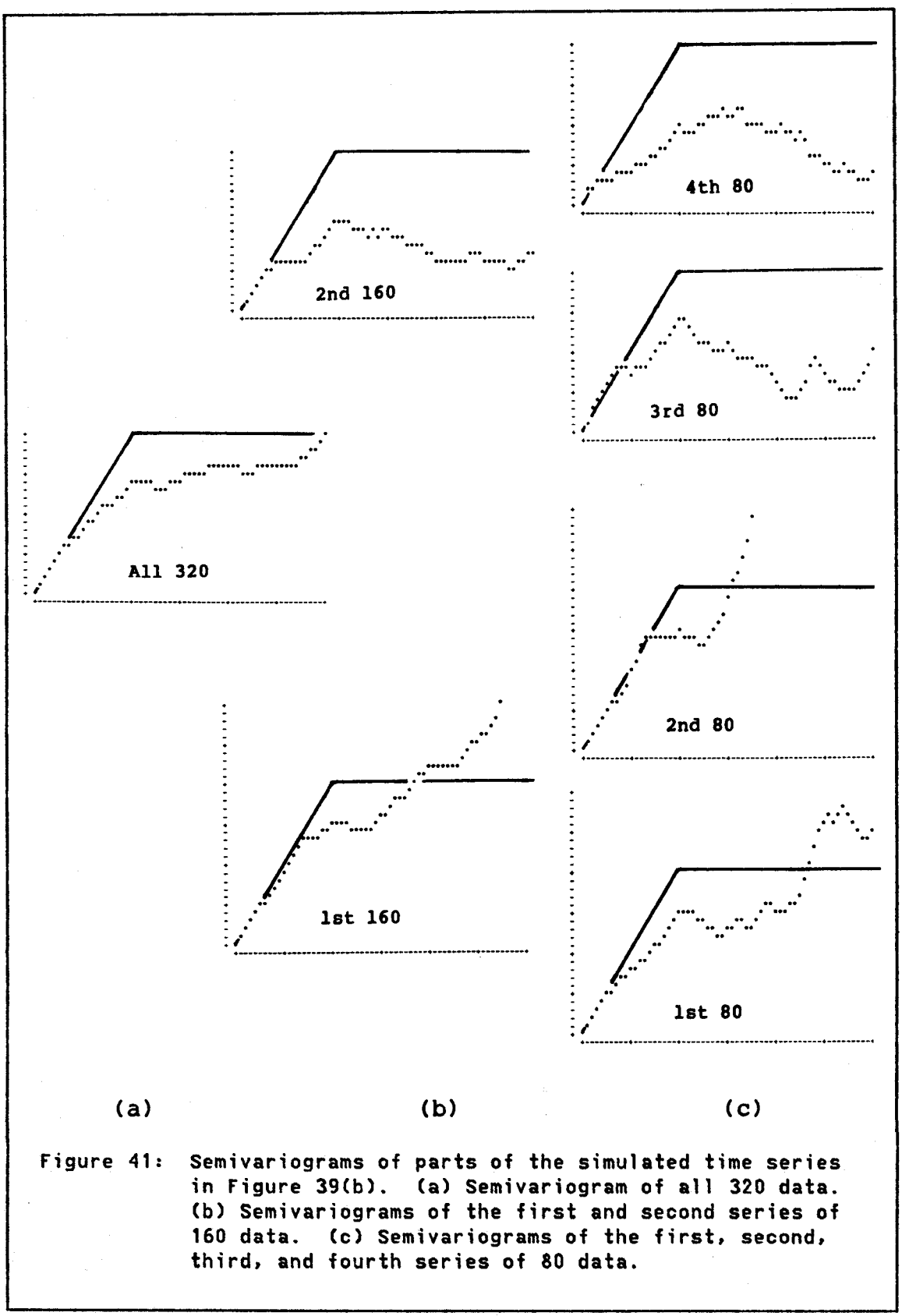


Figure 41: Semivariograms of parts of the simulated time series in Figure 39(b). (a) Semivariogram of all 320 data. (b) Semivariograms of the first and second series of 160 data. (c) Semivariograms of the first, second, third, and fourth series of 80 data.

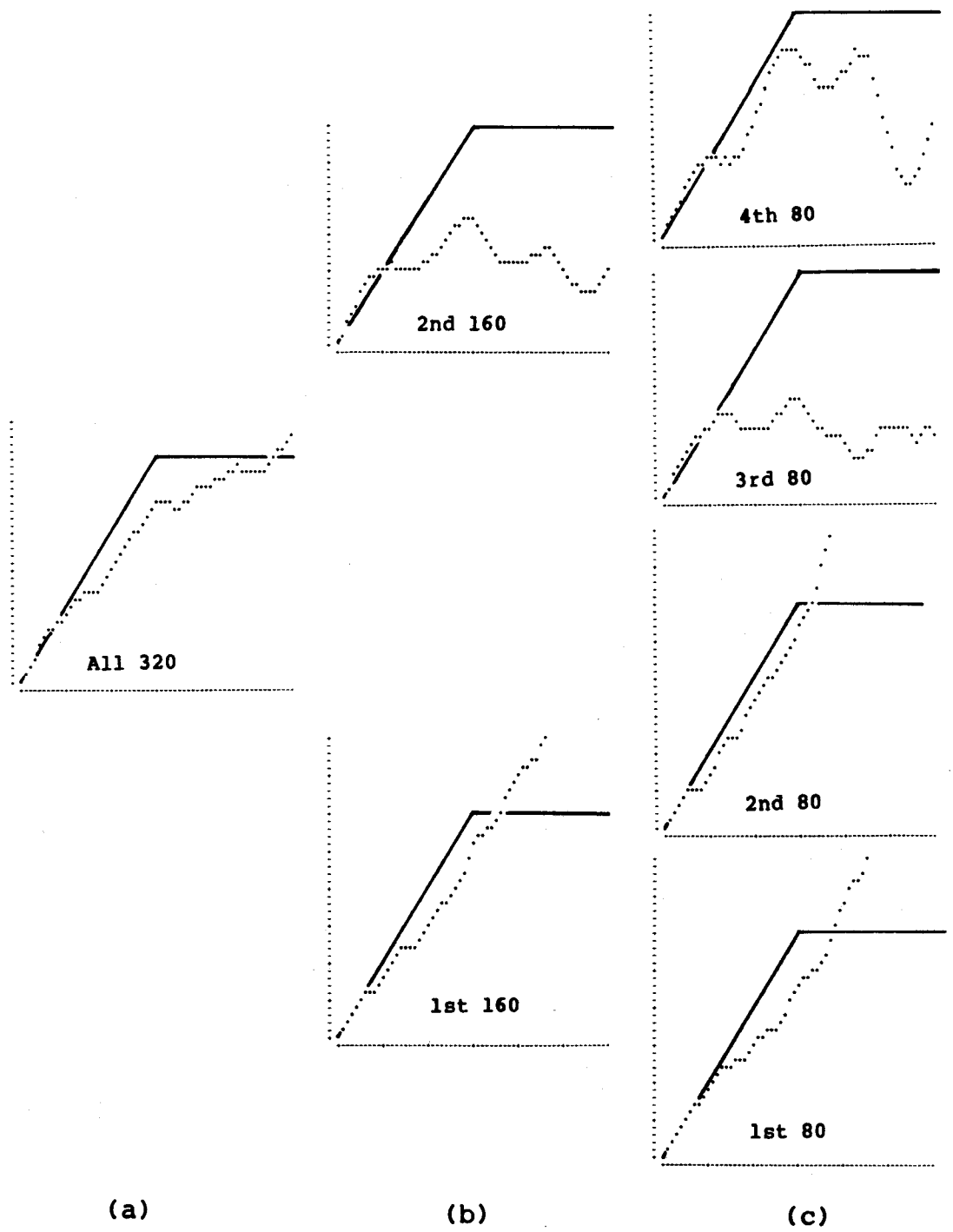


Figure 42: Semivariograms of parts of the simulated time series in Figure 39(c). (a) Semivariogram of all 320 data. (b) Semivariograms of the first and second series of 160 data. (c) Semivariograms of the first, second, third, and fourth series of 80 data.

twenty and thirty data, respectively, were performed to obtain realizations of simple-transitive processes with ranges and sills of twenty and thirty. The semivariograms corresponding to these plots are displayed in Figures 41 and 42. Notice that the "periods" of the local hole effects in these plots tend to increase with the ranges of the variograms. This should not be surprising, as the widths of the local hills and dales in a realization should naturally widen as the range of its variogram lengthens.

Although the examples in Figures 40, 41, and 42 depict only one-dimensional realizations, the problems of variogram domain errors persist into higher dimensions, where they might appear as unstable local anisotropies as well as local drifts and hole effects. However, because variograms in any one direction can be calculated over several lines of grid points instead of just one, the very severe fluctuations that are evident in the one-dimensional examples shown here become progressively less severe in progressively higher dimensions.

The magnitude of the fluctuation variance of a local variogram depends on three factors: the spatial probability law obeyed by the random function, the dimensionality of the spatial domain, and the characteristics (range, sill, functional form) of the semivariogram model. Alfaro (1979) has conducted an extensive investigation into these factors and their relevance in simulation. For <u>unconditional</u> simulations, his major results can be summarized as follows: (1) Fluctuation variance is comparatively small for gaussian processes, and notably huge for lognormal processes. (2) For power-model semivariograms ($\gamma(h)=A|h|^{\nu}$, $0\leq\nu<2$), fluctuation variance increases

with ν. In general, more continuous phenomena fluctuate more widely. (3) Fluctuation variance increases with increasing $|h|$; i.e., reproduction of variogram structure is better at shorter lags. For $|h| \to 0$, the fluctuation variance approaches zero as well. (4) Fluctuation variance decreases with decreasing variogram range (or horizontal scale parameter). (5) Fluctuation variances within linear domains of length L in R^1 are much greater than those within square domains of size LxL in R^2. For gaussian processes in one dimension, the "relative fluctuation variance" $\mathrm{Var}[\gamma_{(v)}(h)]/[\gamma(h)]^2$ is is equal to $\frac{2}{}$ for h=L, regardless of the semivariogram model. This explains the poor behavior of the semivariograms in Figures 40-42 at large $|h|$. For two dimensions, the relative fluctuation variance at h=L depends on the form of the variogram but is always much less than 2. (6) Covariances fluctuate more widely than variograms, particularly at small $|h|$. The fluctuation variance of a covariance does not approach zero as $h \to 0$. This is not surprising, as local covariances depend on local means.

Figure 43 shows histograms, means, and standard deviations corresponding to the realizations in Figure 39(c) and the semivariograms of Figure 42. It is apparent that domain errors among these univariate statistics can also be large, particularly when the range is large compared to the size of the domain. One must be careful when comparing the histograms, as the effects of estimation (in this case, simulation discretization) errors are also important where there are few data per histogram class. However, it is clear that the overall shapes of the histograms vary greatly from one subdomain to the next.

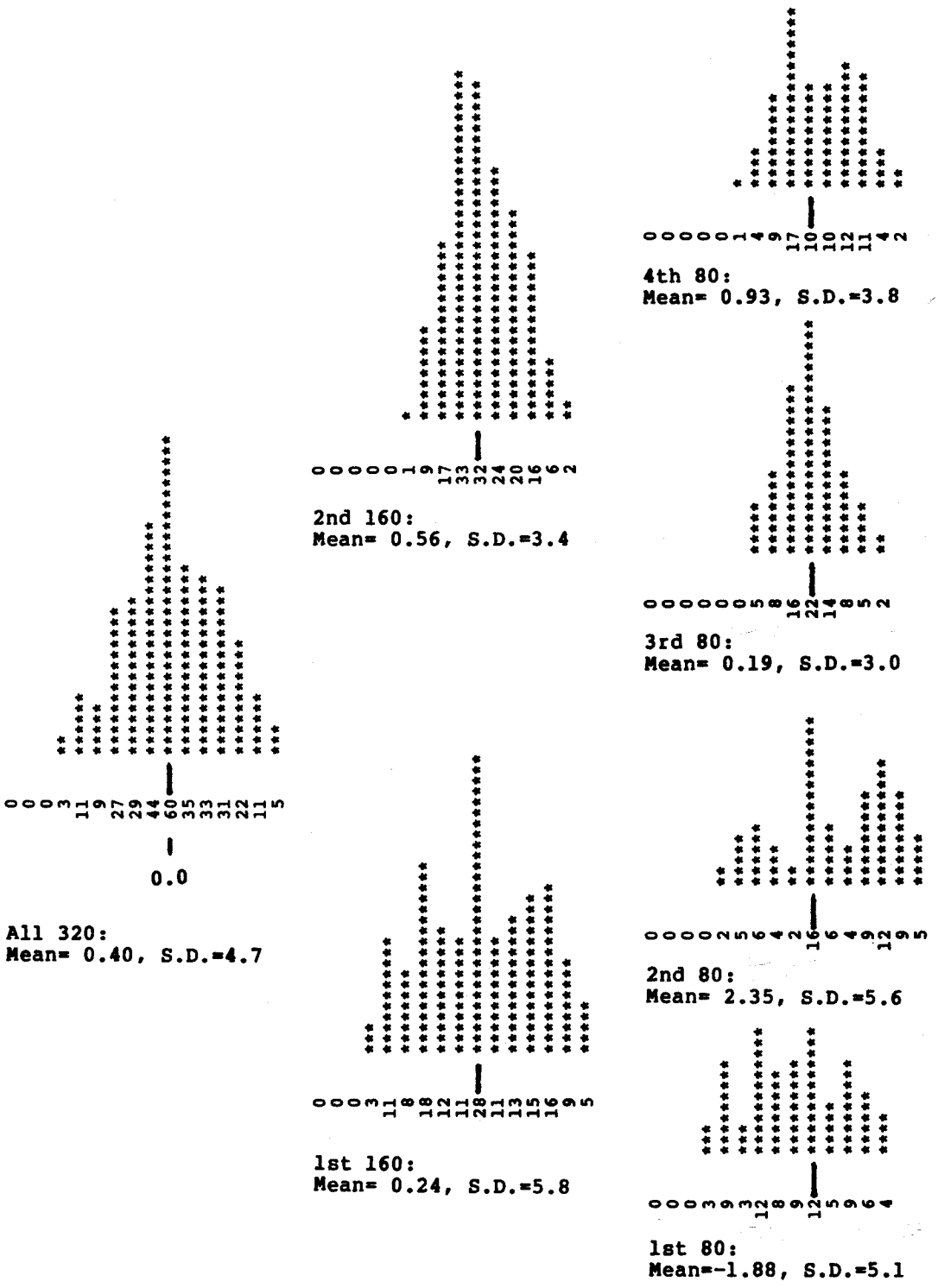


Figure 43: Histograms, means, and standard deviations of parts of the simulated time series in Figure 39(c).

In conditional simulations, the large-scale behavior of the true local semivariogram $\gamma_{0(vs)}(h)$ over the simulation domain can generally be very well reproduced in the simulated semivariogram $\gamma_{s(vs)}(h)$, provided that conditioning data are evenly spread over the whole domain and tend to be closer together than the distance |h| that is of concern. The advantage of good conditioning is evident in the time-series example in Figure 20, Section 3.1.7. The semivariogram of the unconditional simulation (20(b)) departs somewhat from the shape of the sample semivariogram (20(a)), especially at large h, but the semivariogram of the conditional simulation (20(d)) reproduces the idiosyncrasies of the sample plot very well. The conditioning in this example was exceptionally good, as regularly spaced data were available over the whole simulation domain. However, even a few conditioning data are capable of vastly reducing variogram fluctuations among simulations, especially at large |h|. For instance, Alfaro (1979) obtained theoretical results for a gaussian time-series simulation with a linear variogram and only two conditioning data -- one at each end of the simulation domain. Compared to unconditional simulations, the fluctuation variance was cut roughly in half at short distances |h|, was cut much more than half at distances close to the length "L" of the domain, and of course was reduced to zero at h=L.

Further evidence of the power of conditioning may be found in Figure 12, which shows strongly anisotropic directional semivariogram plots, even though the semivariogram model (formulated in the numerical example of Section 3.2) used in the unconditional simulation was isotropic. An isotropic model was used in that case because too few

data were available to infer an anisotropic model with much confidence, but conditioning (including the use of many fictitious data, discussed further in Section 3.7.5) eventually restored the lost anisotropy, at least at distances larger than the data spacings.

Microergodicity. Alfaro (1979, 1984) has pointed out some serious pitfalls that can occur in the reproduction of short-scale variogram structure in a simulation. Conditioning is of little help in the reproduction of variogram structure at distances less than the data spacings, so we must rely on the ability of the unconditional simulation to reproduce the structure faithfully. Alfaro introduces the concept of "microergodicity" in a variogram function, which is defined in terms of the relative fluctuation variance, $\mathrm{Var}[\gamma_{(v)}(h)]/[\gamma(h)]^2$. A variogram function is microergodic if:

$$\lim_{|h|\to 0} \frac{\mathrm{Var}[\gamma_{(v)}(h)]}{[\gamma(h)]^2} = 0$$

If this condition holds, it is possible to estimate (or reproduce in a simulation) the theoretical $\gamma(h)$ very accurately for small $|h|$ from a single realization of the process over the domain V. If the random function is a gaussian process and $\gamma(h)$ resembles a power model (i.e., $\gamma(h) = A|h|^{\nu}+\ldots$) close to the origin (applicable to almost any model), then the condition is satisfied for $0<\nu<2$, with the limit being approached more rapidly for smaller values of ν. For $\nu=2$ (corresponding to smooth functions Z(x)), the limit is a constant nonzero value.

Perhaps one can think of this more easily in the frequency domain: for $\nu<<2$, the process is dominated by high-frequency (local) variation whose average behavior, as characterized by the variogram, can be well

sampled within a relatively small spatial domain V. As ν approaches 2, the process becomes very continuous at short distances, and most of the variance is concentrated at low frequencies, requiring a much larger domain for adequate sampling.

Thus accurate reproduction of the short-scale behavior of a gaussian-model semivariogram in a simulation, or reproduction of any model with parabolic behavior near the origin, can be difficult with a single realization over a small spatial domain. Estimation of such a model from a local data set is similarly risky (strictly because of domain errors, not estimation errors). Models with linear behavior near the origin (most of the models used in geostatistical practice) are relatively easy to reproduce. Figure 44 illustrates the instability of short-scale semivariogram behavior for a smooth-looking phenomenon.

Alfaro (1979, 1984) points out a particularly alarming problem in the reproduction of variograms for nongaussian processes. Variograms of the commonly observed lognormal processes (i.e., $Z(x)=\exp(Y(x))$, $Y(x)$ being a gaussian process), and of most other nonlinearly transformed gaussian processes, are exceedingly difficult to reproduce within a small spatial domain, even for variograms with linear behavior near the origin. This is particularly true for unconditional simulations of lognormal processes, because lognormal processes exhibit a "proportional effect" (Journel and Huijbregts, 1978, p. 186), wherein the local semivariograms $\gamma_{(v)}(h)$ are proportional to the squares of the local means. Very large domains of estimation and simulation are required to obtain stable estimates and good reproduction of the mean and variogram of the random function.

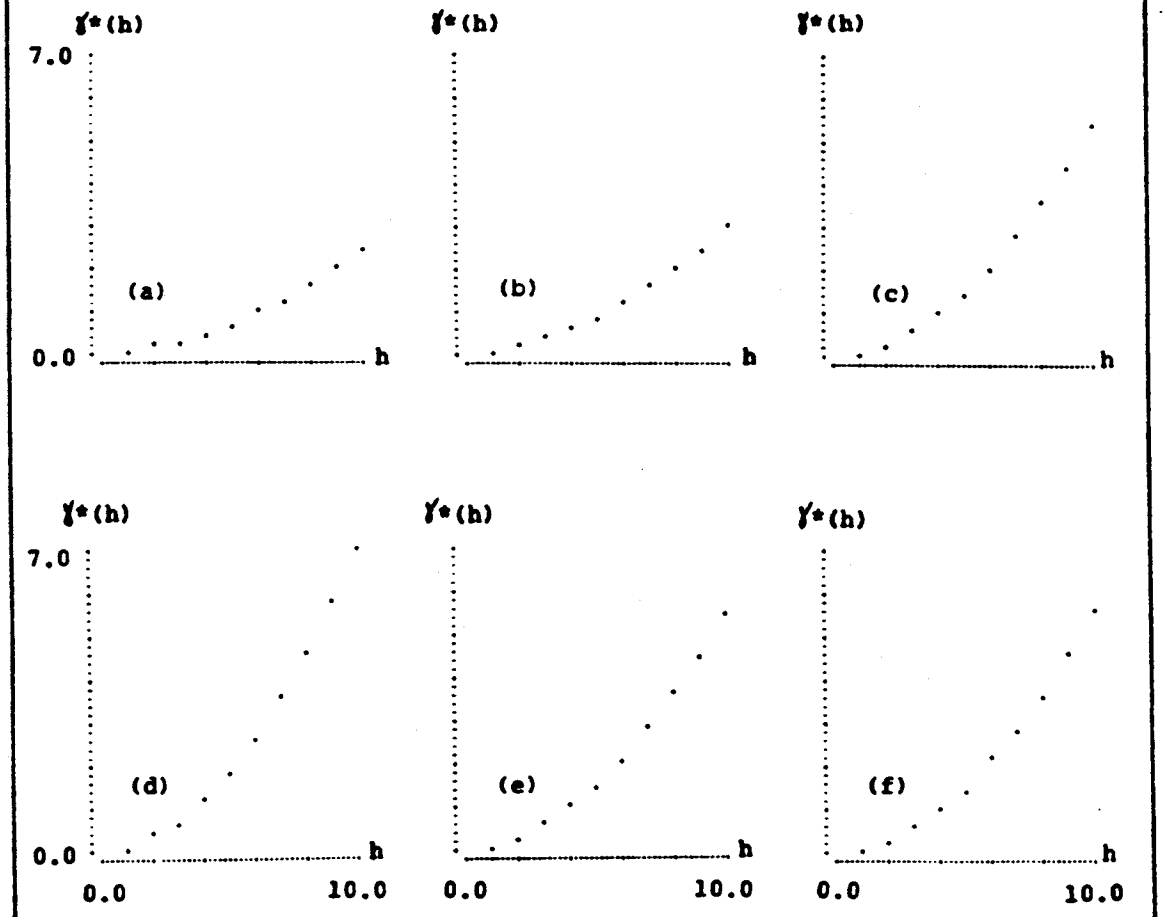


Figure 44: Instability of the short-scale semivariogram behavior of a smooth phenomenon. These sample semivariograms were calculated from six successive ((a) through (f)) disjoint strings of 300 data from a single time-series realization of a simple-transitive gaussian process with range 100, regularized by a moving window of width 50.

In practical simulation problems, we are not interested in the random function (which has no physical meaning anyway if we are studying the unique realization of some geological phenomenon), so the two caveats to keep in mind are that nongaussian processes in general, and lognormal processes in particular, require very good conditioning for adequate

reproduction of large-scale variogram structure, and that the short-scale structure of a nongaussian realization is apt to be poorly reproduced if the domains of estimation and simulation are not large or do not essentially coincide. If we are interested in a very small simulation domain, we might obtain a feeling for the magnitude of this problem by running several independent conditional simulations over the domain and observing the variability of the summary statistics among the different realizations. (Other situations in which repeated simulations are useful are described in Section 3.7.7.)

Now here is an effort-saving practical comment.[65] We have seen that good conditioning will do much to insure reproduction of local means and local variogram structures (at least for large $|h|$) for gaussian processes, even if the corresponding statistics of the unconditional simulation have run far afield. Therefore, why should we construct elaborate semivariogram models with complicated nested or anisotropic structures for our unconditional simulations, if the effects of those complications will be evident only at large distances $|h|$? If the simulation domain V_s is densely and evenly covered by conditioning data, then only the short-scale variogram structure and the microergodicity of the model are important considerations at the unconditional-simulation stage. Fussiness about variogram behavior at distances much longer than the typical data spacing is unjustified.

For example, the nested simple-transitive model used for the unconditional simulation in Figure 20(b) is really needlessly complicated and was used only because it was not convenient to make this

---

[65] Alas, one of the few effort-saving practical comments made in this dissertation.

remark earlier. Because data were available for that simulation on fifteen-minute intervals, compared to nested variogram ranges of 60 and 105 minutes, only the shape of the variogram up to $|h| \approx 30$ is really important.[66] A single simple-transitive structure with the same short-scale shape should have done just as well in practice, although the theory underlying conditional simulation assumes that the observed and unconditionally simulated models are identical.

So in the presence of good conditioning (and particularly in the absence of any big gaps in the data, where conditioning would have little effect on the simulation), a good rule of thumb is: model the semivariogram carefully at distances around and within the typical data spacing -- essentially up to the distance that would be included in a practical kriging neighborhood (Journel and Huijbregts, 1978, p. 345) -- then model it only approximately at greater distances. Let the conditioning do the rest.

A dramatic example of the application of this rule is shown in Figure 45, which shows sample semivariogram plots from four unconditional ((a) to (d)) and three conditional ((e) to (g)) time-series simulations. Every fifth simulated datum from the first unconditional simulation (semivariogram (a)) was used to condition the unconditional simulations represented by semivariograms (b) through (d), which differ vastly from one another at large distances $|h|$, but are all the same at short distances. The resulting conditional simulations,

---

[66] The shape of the variogram at larger distances might be important for kriging, but in this example it was not even important for kriging because the closely spaced time-series data screened out the influence of far-away data.

represented by semivariograms (e) through (g), are all virtually identical in structure to semivariogram (a) at all distances |h|.

In the presence of poor conditioning, the expenditure of great effort on modeling (minimization of MEE[γ(h)]) may still be open to question, as the domain error DE, and possibly the estimation error EE as well, may be so large that careful reproduction of a poorly known function by attempted minimization of MEE would not be worth the trouble.

By now it should be clear that, in most applications, the most important characteristic of the data that an unconditional simulation should reproduce is the short-scale variogram structure. If the unconditional simulation is later conditioned to the data, the larger-scale structure and the other statistical characteristics of the data will be imposed by conditioning, and perhaps by subsequent back transformations. Even if the simulation is not conditioned, or is conditioned by a sparse or unevenly distributed data set, reproduction of short-scale structure remains more important than reproduction of large-scale structure, because large fluctuation variances at large scales are natural and appropriate in unconditional simulations. The common practice of standardizing unconditional simulations to a sample variance or sill value is thus incorrect in most applications, because dispersion variances and sill values depend largely or entirely on large-scale variogram structure, which commonly is poorly known and poorly reproduced in unconditional simulations. Standardization of a simulated data set to reproduce the sample sill or sample variance of the data within a finite domain may have two unfortunate effects: (1) it may bias the relatively well reproduced short-scale structure of the unconditional simulation by altering the variogram's slope to

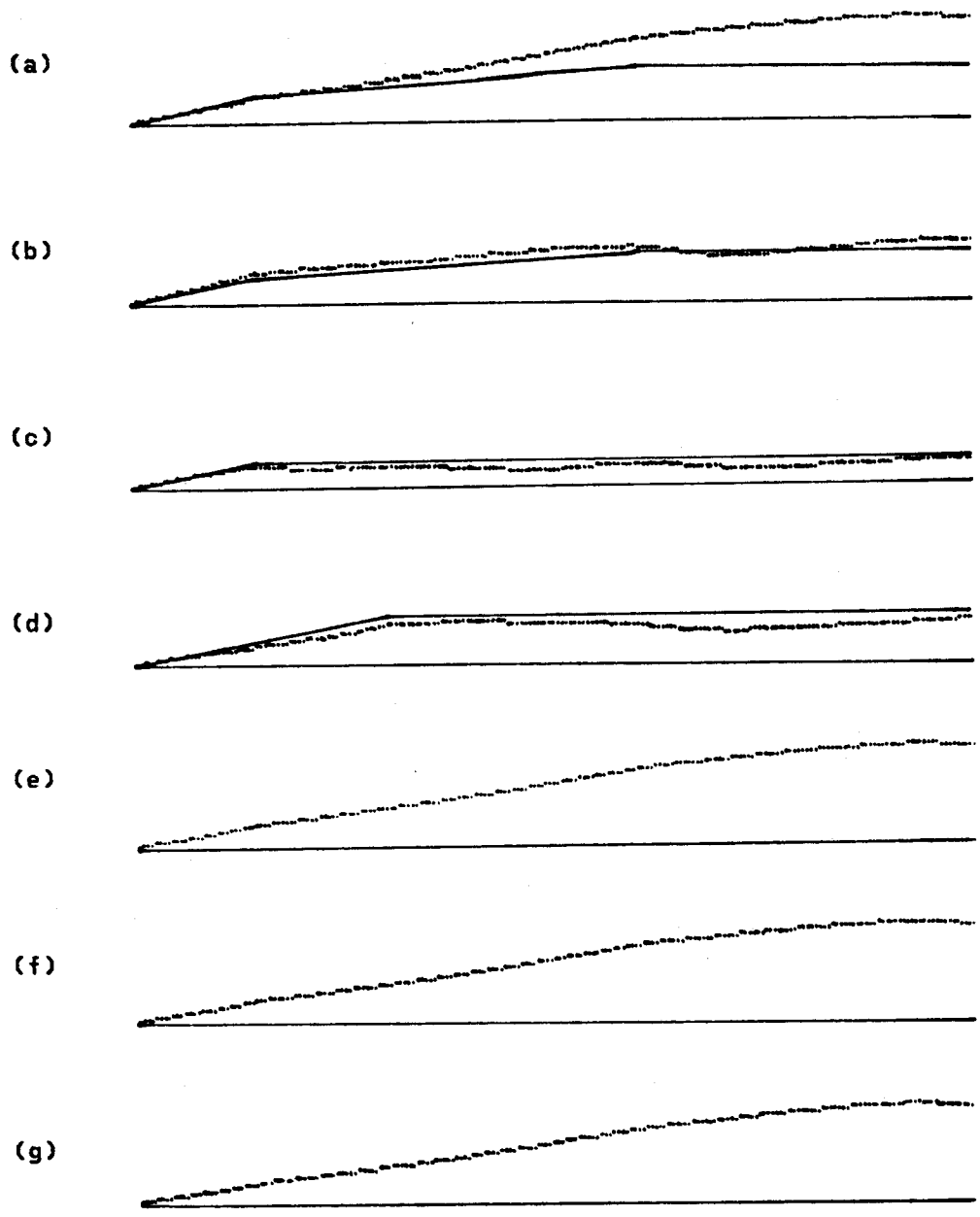


Figure 45: Comparison of the semivariograms of unconditional and conditional time-series simulations. The models used for the four unconditional simulations are represented by solid lines. Every fifth datum from simulation (a) was used to condition simulations (b) through (d), producing conditional simulations (e) through (g), respectively.

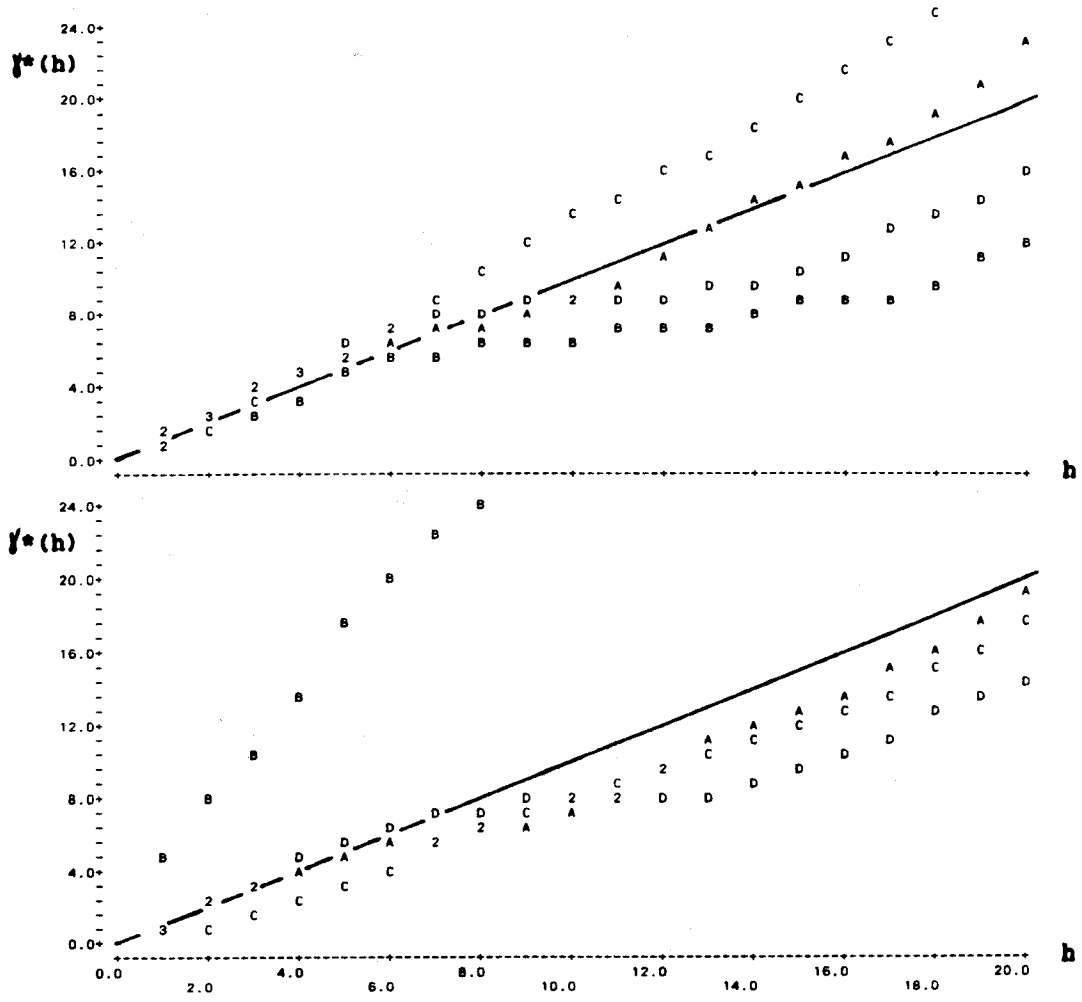


Figure 46: Illustration of the ill effects of standardization on the short-scale structure of sample semivariograms. Sample semivariograms of four unconditional time-series simulations of length 150 are shown in the upper diagram. The model (solid line) is simple-transitive with a range=50, sill=50, and no standardization has been performed. The lower diagram shows semivariograms for the same simulations after standardization of the sample variance to exactly 50. Notice that reproduction of the model semivariogram structure is adversely affected at short lags. Even if standardization had been performed to the proper dispersion variance $D^2(0/150)$ within the domain of 150 data instead of to the sill, $D^2(0/\infty)=50$, the individual plots still would have been more dispersed around the model at short lags after standardization.

compensate for the relatively poorly reproduced large-scale structure (Figure 46), and (2) it will destroy any value the unconditional simulation might have had as a simulated realization of the model random function. Independent realizations of a random function are not supposed to have the same sample sill or the same sample dispersion variance, unless the realizations are defined over practically infinite spatial domains. Forcing such a characteristic on them is artificial and in some applications could be misleading.

3.7.3 Estimation Errors (EE and MEE)

Estimation errors (EE) naturally tend to decrease with an increasing number of discrete data within the observation domain V_0, but they also depend on the amount of detail that is required in the estimate. For instance, in the case of irregularly spaced data, the estimation error $EE[\gamma(h)]=[\gamma^*(h)-\gamma(h)]$ for the semivariogram depends on the widths of the distance and direction classes over which the squared-difference summation is carried out. In general, as the class widths increase, the number of data pairs falling within each class increases, and the estimate within each class becomes more stable; i.e., if we add or subtract a few data the estimate will not fluctuate so wildly. However, if the class widths are made too large in order to increase the number of data pairs, this increased stability may have been traded for a loss of sensitivity in the estimate. For example, very wide distance classes may make the variogram's exact range hard to observe, as data pairs at distances well beyond and well within the range may be averaged together. Similarly, a sample variogram calculated from a

two-dimensional realization using wide azimuth (direction) classes may not reveal a subtle anisotropy in the true variogram. This tradeoff between sensitivity and stability (i.e., accuracy and precision, or bias and variance) of an estimate also appears in other contexts. For example, in the estimation of density functions, an increase in the widths of histogram classes may dampen spurious spikes and troughs in the histogram while smearing out the details of the true density; and in checks on stationarity by calculation of local means and variances within moving windows, an increase in the window size will decrease sampling variability unrelated to the structure but also smooth over real small-scale features.

A corresponding tradeoff occurs when constructing a model of the continuous phenomenon from a discrete estimate, involving the commission of a model estimation error, MEE. Many long and fruitless arguments have been waged over whether an apparent break in slope in a ragged sample semivariogram plot is real and deserving of incorporation into the model, or a sampling artifact that is best ignored. Unfortunately, the classical semivariogram estimator introduced on page 66 is very sensitive to extreme values and spatial clustering in the data, so large estimation errors and consequent modeling errors are easy to make. If problems of this kind are suspected, an investigation of "robust and resistant" variogram estimators and procedures for assessing the quality of semivariogram models may be worthwhile. Some of these are described in Part I of Verly et al. (1984).

Fortunately, it is evident from the results of Section 3.7.2 that accurate semivariogram modeling is not critical for a well conditioned

simulation, except at distances shorter than the radius of a practical kriging neighborhood. Unfortunately, the structure at very short distances usually is poorly known, especially at distances shorter than the typical data spacing. If one is fortunate enough to have a few "twin" drill holes, or a "cross" or "fence" of holes drilled at relatively close spacings, it may be possible to get a fairly good estimate of this short-scale structure, although even in the most favorable of real situations, the number of pairs available for semivariogram estimation near the origin is typically far less than the recommended practical threshold of thirty (Journel and Huijbregts, 1978, p. 194).[67] "Simple", and thus sparsely drilled, deposits of limestone, clay, or coal are particularly likely to be deficient in closely spaced data. The problem is more serious for nongaussian data, owing to the microergodicity problem summarized in Section 3.7.2.

The problem of modeling short-scale variogram structure is well known in geostatistics, as the short-scale structure (particularly the size of the nugget constant) has a large influence on kriging variances and the relative sizes of kriging weights. In simulation, the problem is even more serious: we need the short-scale structure not only for kriging but even more importantly for modeling the semivariogram to be reproduced in the unconditional (and, via the kriging errors, in the

---

[67] The variance of estimation of a semivariogram is inversely proportional to the number of data pairs used. For a one-dimensional gaussian process with a power-model semivariogram ($\gamma(h)=|h|^{\nu}$), the variance of estimation of the semivariogram is equal to (Journel and Huijbregts, 1978, p. 193):

$$\mathrm{Var}\{E[\gamma(h)]\} = E\{[\gamma^*_{(v)}(h)-\gamma_{(v)}(h)]^2\} = 4\gamma(h)\cdot D^2(0/V)/N(h)$$

where $\gamma_{(v)}$ is the true local semivariogram within the one-dimensional domain V, $\gamma^*_{(v)}(h)$ is its estimate, $D^2(0/V)$ is the dispersion variance of the (assumed point-support) data within V, and N(h) is the number of data pairs at lag h available to calculate $\gamma^*_{(v)}(h)$.

conditional) simulation. The validity of the final results of a simulation study -- e.g., the estimate of short-term variability of mined ores delivered to a plant -- commonly depends critically on the short-scale structure of the simulation, and this structure is not imposed by conditioning but by the unconditional simulation and the quality of the semivariogram model (as illustrated in Section 3.7.2). Hence it is vitally important in real-world simulation studies to have an adequate set of closely spaced data, particularly if the original data are not gaussian.

Closely spaced data are particularly important if the phenomenon is not very continuous at short distances (e.g., a phenomenon with a linear semivariogram $\gamma(h)=|h|^{\nu}$ with $\nu<<2$, or a variogram with a short range). Alfaro (1979) demonstrates that variogram estimation errors are <u>larger</u> in such cases, in contrast to fluctuation variance, which tends to be <u>smaller</u> in these same cases. Alfaro's "equilibrium principal" summarizes the relationship: what you lose in fluctuation variance, you gain in estimation variance.

Large data sets are also required for the data-directed transformations discussed in Section 3.5. The piecewise normal-scores transformation and its interpolated inverse transformation (Figure 29) are not satisfactory if there are insufficient data to provide a smooth-looking sample distribution. Furthermore, owing to the curse of dimensionality, the multivariate density estimates required to implement the multivariate stepwise-conditional gaussian transformation of Section 3.5.2.3 are unreliable without a large data set.

Obviously, massive estimation and modeling errors can occur if the effects of drifts, discontinuities (contacts, faults), multiple populations, and clustered or preferential drilling schemes are not recognized and accounted for before the modeling stage of a simulation study. Practical approaches to some of these problems are discussed in later sections of this chapter, and some are illustrated in the case study of Section 4.2. The practical advice on structural analysis offered in Section III.C of Journel and Huijbregts (1978) is also applicable.

3.7.4 Simulation Discretization Error (SDE)

Simulation discretization error comprises all of the interrelated errors that are committed when a continuous spatial phenomenon is simulated at a discrete set of grid points, using a simulation method involving discrete approximations of the covariance function or spectral density function.

The most straightforward of these component errors is the error involved in discretizing the simulation domain V_s into a regular grid of points x_i at which simulated realizations of the random variables $Z_s(x_i)$ will be generated. The total number N of grid points x_i, i = 1 to N, is a major factor in determining the cost of a simulation, so it should be chosen carefully. Once the boundaries of the simulation domain V_s have been chosen (with due regard to the effects of the domain errors DE and MDE), the choice of N amounts to a decision on the spacings of the simulation grid and, to a much lesser extent, the orientation of the grid. A number of factors should be considered when choosing grid spacings and orientation:

3.7.4.1 Choice of Grid Orientation

If the phenomenon to be simulated is anisotropic, an unconditional simulation is usually easier to perform if the grid is oriented parallel to the major axes of anisotropy. A geometric anisotropy on a square or cubic grid can then be simulated by performing an isotropic simulation (usually simpler to do) on a rectangular grid whose spacings are inversely proportional to the variogram ranges in the corresponding directions (Figure 47). This is the method incorporated into subroutines CS2D and TB3D, presented in Appendices A and B.[68] Zonal anisotropies are also easier to incorporate into a simulation if the axes of all additive zonal structures can be oriented parallel to the grid.

If the phenomenon is isotropic, if the data are located on a perfectly regular grid that is convenient for kriging, if the simulation domain V_s is large and rectangular in shape, or if mining blocks are to be laid out in some prespecified orientation, then it may be easier in the long run to orient the simulation grid parallel to the data grid, the rectilinear boundaries of the simulation domain, or the boundaries of the mining blocks. If the anisotropy is in conflict with one of these orientations, one may have a difficult choice to make; however, it may still be possible to satisfy the conflicting goals simultaneously. Sample variograms commonly do not yield sufficient detail to choose the axes of anisotropy precisely, so it may be possible to juggle the anisotropy a bit so that it corresponds to a 45-degree diagonal direction in the conflicting grid system (Figure 48). One should keep

---

[68] An oblique anisotropy can be used in CS2D by generating the window externally, as explained in Appendix A.

in mind that the distance between colinear grid points in this oblique direction will be longer than the spacings of the rectangular simulation grid, so in some applications it might be necessary to decrease the spacings to make the simulation detailed enough in all directions of interest. In a conditional simulation, a slight alteration of the directions of anisotropy in the unconditional simulation will not usually be detectable, because conditioning will force the anisotropy of the conditioning data onto the simulation, at least at large distances.

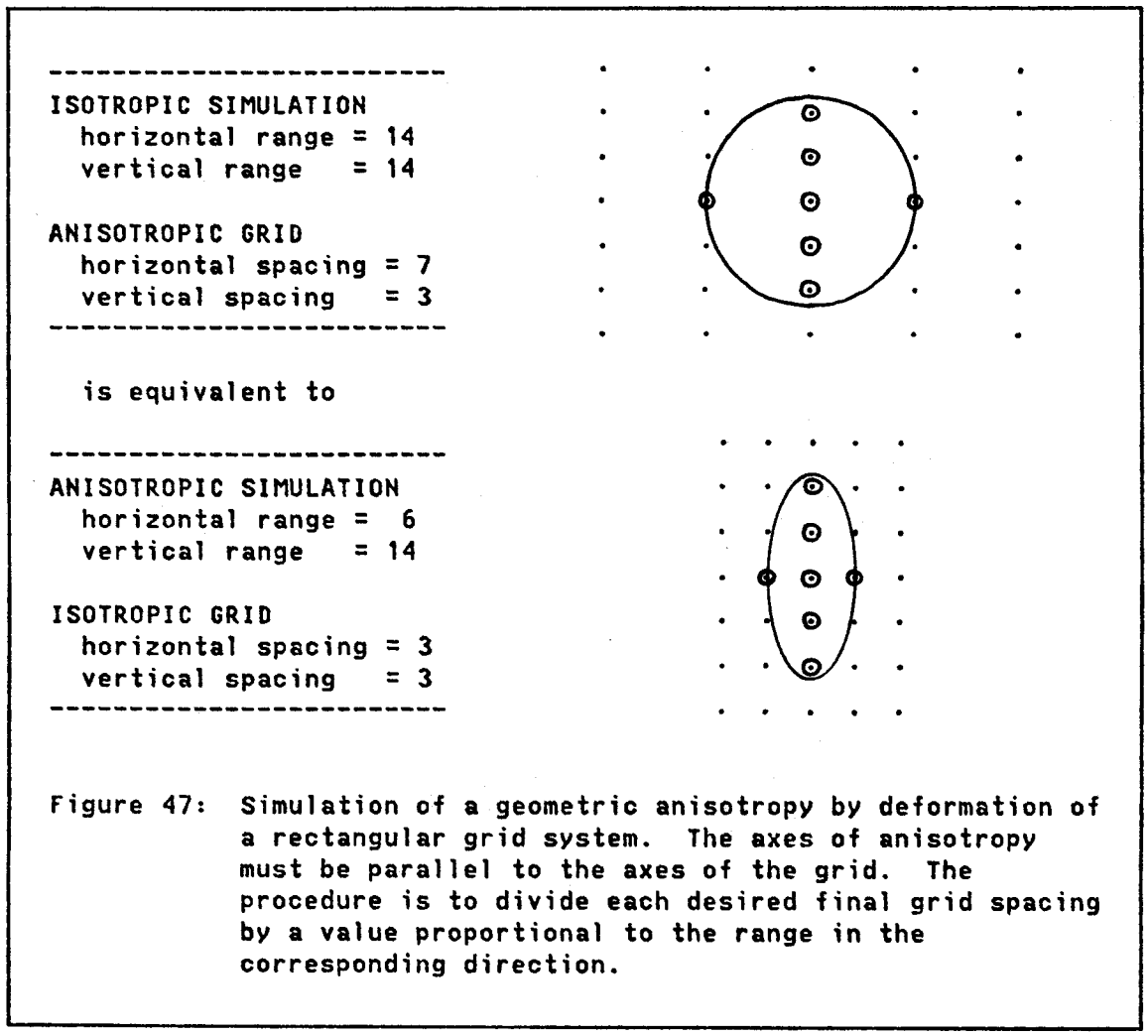


Figure 47: Simulation of a geometric anisotropy by deformation of a rectangular grid system. The axes of anisotropy must be parallel to the axes of the grid. The procedure is to divide each desired final grid spacing by a value proportional to the range in the corresponding direction.

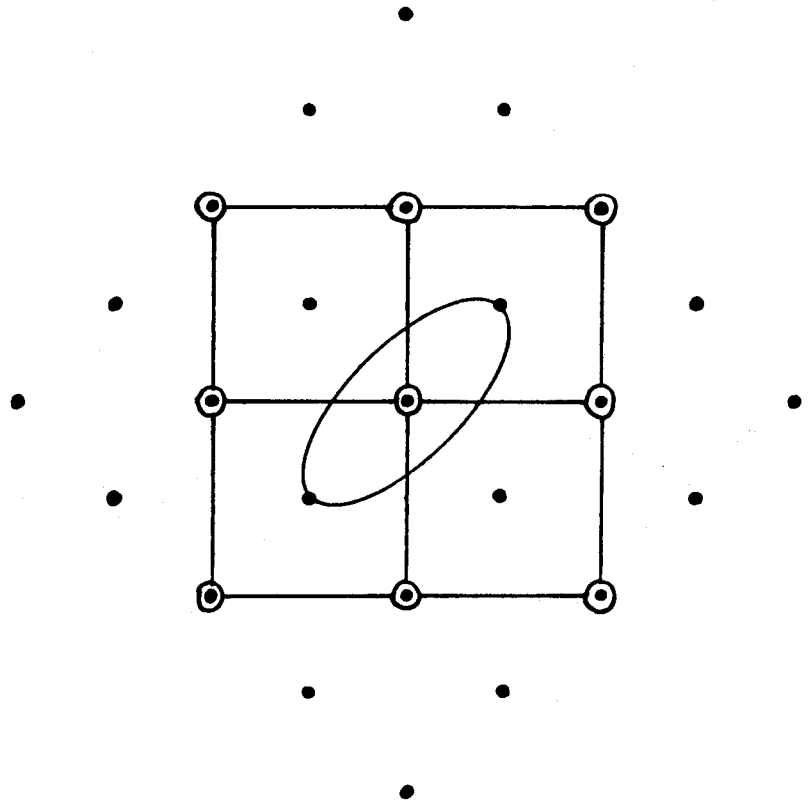


Figure 48: Orientation of a simulation grid system where the anisotropy is oblique to the preferred grid directions. Alternate points from the original 45-degree grid (all black dots above) are retained in the final simulation grid (circled dots).

3.7.4.2 Choice of Grid Spacings

At least four considerations are important when choosing grid spacings: (1) the form (particularly the range) of the variogram in the corresponding direction, (2) the spacing and accuracy of data locations, (3) the support of the data, and (4) the nature of the problem being investigated. These matters are discussed in detail in the following paragraphs:

(1) Form of the variogram. The spacing in each direction should be at least fine enough to allow reproduction of the important features of the variogram in that direction. In particular, the range of any nested structure will be difficult to see if the spacing is greater than about one fourth of the range. The parabolic behavior of the gaussian model near the origin may also be hard to observe if several spacings are not available within this short distance.[69] Furthermore, the all-important kriging errors that are generated during kriging of the unconditional simulation will be more realistically structured if the grid spacing is small with respect to the range of the variogram and the typical distance between the data points used for kriging.

In simulating an anisotropy, one should consider the reproduction of the structure in all directions, not just those directions that correspond to grid (or structural) axes. Figure 49 shows plots of several discretized elliptical windows for a moving-average simulation by subroutine CS2D; although most windows reproduce the range very well in the axial directions, notice that the discretization is effectively coarser in diagonal directions. Some diagonal ranges are thus poorly reproduced, and the elliptical anisotropy as a whole is correspondingly poorly reproduced.

---

[69] Variograms of regularized phenomena also exhibit parabolic behavior near the origin, owing to overlap of the nonpoint support at short distances. If the simulation is supposed to represent non-overlapping volumes of material centered at necessarily greater spacings, the simulated realization will not exhibit a clear parabolic behavior. A shorter spacing than the width of the support could be used in the simulation to check the accuracy of the simulated variogram, but then only simulated data at the centers of nonoverlapping blocks should be retained for subsequent studies. More comments on the subject of support appear in paragraph (3), below.

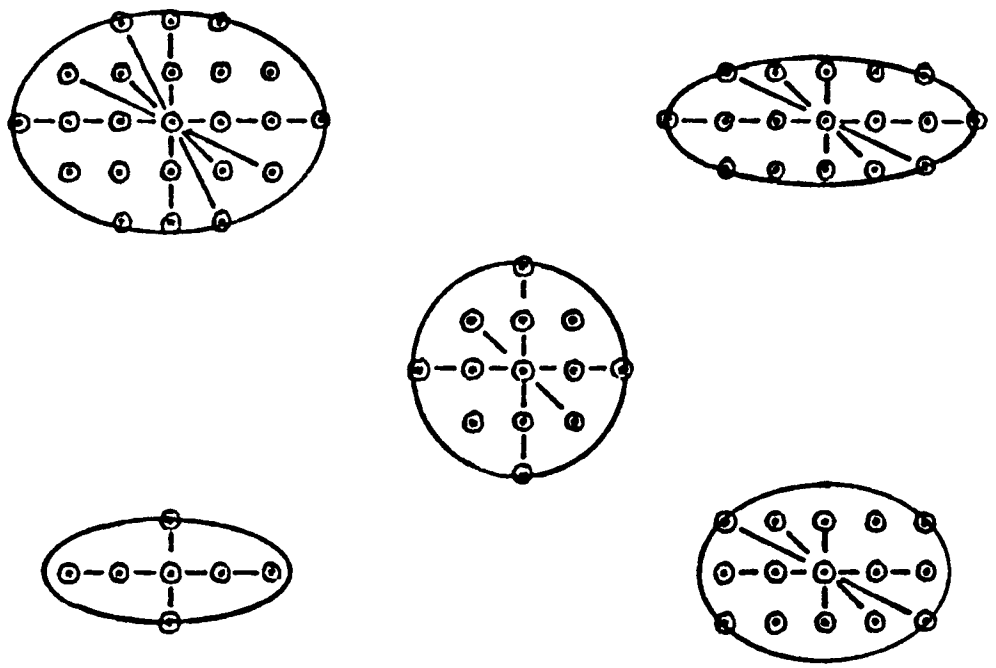


Figure 49: Discretized elliptical weight functions used to simulate realizations with anisotropic circular variograms. Compare the discretized ranges in various directions with the smooth geometric anisotropies in the model weight functions.

(2) Data locations. The theory of conditioning requires that kriging be performed using real and unconditionally simulated data at the same locations. Thus, if conditioning data are not on a regular grid, it is important to make the simulation grid fine enough so that a grid point in the unconditional simulation will fall acceptably close to each real data point. (One could also simulate additional data at the exact locations using a matrix approach from Section 3.3.1.1, but this would complicate the procedure.)

It is also important to remember that the domain of the unconditional simulation must be large enough to cover the whole set of conditioning data, even though the domain of the conditional simulation, V_s, may be a small subset of this region. The seemingly silly mistake of not making the unconditional domain big enough is easy to make.

(3) Data support. The support of the estimation and conditioning data should be considered when choosing grid spacings. If the data represent the average properties of constant sample volumes "v" of specified dimensions, rather than point-support measurements, then it is simplest to simulate only the regularized process represented by the data. Then each simulated value at a specific grid location represents the average content of a volume "v" centered at that grid point. If the grid spacings are equal to or greater than the corresponding dimensions of volume v, the simulation can proceed exactly like a simulation of point-support data. If the spacings are less than the corresponding dimensions of v, then more than one simulated value may lie inside a given volume with the dimensions of v. Then alternative approaches might be taken:[70]

(a) Use a model obtained directly from the nonpoint data to simulate the regularized process in the usual way, but retain only simulated data at locations corresponding to the centers of nonoverlapping blocks of support v. (A dense simulation of physically overlapping blocks is useless for most applications.) If the original data were on a regular grid, this procedure would be wasteful, as many simulated data would be discarded. If the data were irregularly located, then the procedure

---

[70] Paragraphs (a) to (d) that follow can be skipped by readers uninterested in the details of this narrow subject.

makes more sense, because the finer grid would allow the data locations to be reproduced more accurately. (This is the approach mentioned in the footnote of paragraph (1) and used for the unconditional simulation in Section 4.1.)

(b) Use the model obtained from the nonpoint data to perform the kriging, and attribute the known conditioning value inside each sampled volume v to all grid points that fall inside the volume. This will result in a smooth realization (especially near the locations of conditioning data) that will not really represent the short-scale variability of either a point-support process or a process regularized over smaller volumes v'<v, centered on the simulation grid points. Thus there is really no advantage to having so fine a grid, and usage of such a simulation to study the effects of short-scale variability could be very misleading. This approach is not recommended.

(c) Use a "deregularized" (deconvoluted) model (e.g., obtained using the procedure suggested by Journel and Huijbregts, 1978, pp. 90-91) to perform the unconditional simulation,[71] followed by kriging as above. This will still result in an overly smooth conditional realization in the vicinity of the conditioning data.

(d) Use a deregularized model as above, and substitute a set of "roughened" conditioning data on the deregularized (usually point) support for each real datum on support v. The roughened data should be

---

[71] In deriving the deregularized model, it is advisable to have some point-support data from which to estimate the behavior of the deregularized semivariogram, particularly near the origin. "To be rigorous, it is not possible to reach a greater degree of precision than that of the smallest support v of the data without introducing supplementary and unverifiable hypotheses." (Journel and Huijbregts, 1978, p. 231)

unconditionally simulated, using a structure consistent with the structure of the deregularized model at distances less than the dimensions of v (e.g., their variance should approximate the dispersion variance $D^2(v'/v)$ of the simulation grid support v' within the original data support v). Their values should be standardized so that the average of the simulated values within a conditioning volume v is exactly the real conditioning value if the data points represent volumes v'<v that sum exactly to the conditioning volume v.[72] If the simulated conditioning values represent only a grid of point data within v, then their average should be standardized to a random value consistent with the estimation variance of the known value of support v by the internal grid of point data. Although complicated (perhaps needlessly so, in practice), this approach will provide a fairly rigorous simulation of the deregularized process.

If a simulation is required only for block data on a larger support than the actual data, then the alternative change-of-support methods described by Journel and Huijbregts (1978, Section VII.A.5) can be used to perform a block simulation. The easiest and most flexible (though not necessarily cheapest) approach is simply to perform a point-support (or data-support) simulation and then average the points to obtain the block values, as described below and in Journel and Huijbregts (1978, pp. 511-513).

---

[72] If the average of the simulated grid data on small support v' is to equal the known conditioning value within the larger volume v in the final back-transformed conditional simulation, then the average of the transformed grid values will not generally equal the transformed conditioning value within v if the transformation is nonlinear. Thus the standardization of the conditioning values should be performed on back-transformed data.

In simulations of tabular deposits, such as most sedimentary units, not only the grid spacings but also the dimensionality of the simulation domain may be open to choice. If the entire thickness of the deposit is mined in one bench, there may be no point in performing a three-dimensional simulation. However, if the data are drill-core analyses and these analyses do not routinely intersect the entire thickness of the deposit, a three-dimensional simulation may still be necessary. An advantage of two-dimensional simulations of tabular deposits is that the common problem of variable core-sample lengths can simply be averaged out.

(4) Nature of the problem. The grid spacings of a point-support simulation should be fine enough to provide an adequate approximation of a continuous regionalized variable realized at all points within the continuous domain V_s. Adequacy in this sense depends on the purpose of the study. For instance, if we are simulating an ore deposit for a study of fluctuations in the grades of mined ores delivered to a plant, we should have in mind some minimum sampling increment (support or tonnage) of ore whose characteristics the simulation should respect faithfully. Obviously the simulation cannot practically be expected to reproduce the very short-scale behavior of a stream of delivered ore (rock by rock, for example). But if we are interested in the variability of lots of delivered ore as small as, say, 100 tons, then we would want the simulation to reproduce the variability of hundred-ton blocks of ore (or fifty-ton blocks, if two faces are mined simultaneously, etc.) faithfully. In other words, if "d" denotes the set of simulated point-support data located inside a block "v" of

specified dimensions, and Q_s is the simulated quantity in question, then we want the error $[Q_{s(d)}-Q_{s(v)}]$ to be tolerably small. In mine-planning and mineral-processing applications, the dispersion variance of blocks of ore within some region V of the deposit is commonly of great interest. Thus we want:

$$D^2(d/V) \approx D^2(v/V)$$

$$\bar{\gamma}(V,V)-\bar{\gamma}(d,d) \approx \bar{\gamma}(V,V)-\bar{\gamma}(v,v)$$

$$\bar{\gamma}(d,d) \approx \bar{\gamma}(v,v)$$

For simulations, a reasonable criterion of acceptability would be that the difference $[D^2(d/V)-D^2(v/V)]$ should be small in comparison with $D^2(v/V)$ -- the quantity we wish to reproduce. If the volumes v and V are rectangularly shaped, the quantities $\bar{\gamma}(v,v)$ and $\bar{\gamma}(V,V)$ can usually be calculated easily by means of the auxiliary function "F" or looked up in a chart (e.g., see Journel and Huijbregts, 1978, Section II.E). Then a simple computer program to calculate $\bar{\gamma}(d,d)$ among progressively denser grids of points within v can be used to determine the maximum grid spacings that will result in acceptably close values of $D^2(d/V)$ and $D^2(v/V)$. As a rough guide, Journel and Huijbregts (1978, p. 97) suggest that the following numbers of regularly spaced points provide adequate approximations to continuous $\bar{\gamma}(v,v)$ values for <u>most</u> practical applications:

For a linear v, 10 points;

For a square v, 6x6 points;

For a cubic v, 4x4x4 points.

Of course the proper discretization depends on the characteristics of $\gamma(h)$, so these suggested discretizations may be inappropriate for some

simulations. A preliminary computation of $\bar{\gamma}(d,d)$ for some alternative grid densities is recommended.

One should keep in mind that a simulation may be used for several purposes, and the spacing should be adequate for any such usage. For instance, the sizes, shapes, or orientations of mining blocks may be changed, or may be undecided at the simulation stage. For a coregionalization, different variables may have different structures. Usually the spacing allowing proper reproduction of the shortest-range structure should be adopted. If transformations are performed on the data, the ranges observed in the original and transformed processes might differ. The spacing should then reflect the shortest range found among both the original and transformed structures -- the original structures, because these structures should be reproduced in the final back-transformed simulation that is actually used, and the fully transformed ("gaussian-process") structures, because without reproduction of these structures the entire spatial distribution of the gaussian data cannot be reproduced, and thus the quality of the back-transformed simulation may be jeopardized.

3.7.4.3 Choices of Other Parameters

Improper choices of other parameters used at the unconditional-simulation stage can result in discretization errors of other kinds. Many of these choices are related to the choice made for grid spacings, so it is better to choose the grid spacings first and allow this prior choice to influence the choices of other parameters.

For some simple examples of the effects of grid spacings on the choice of unconditional-simulation parameters, consider first the choice of the number of Poisson points that must be generated in a simulation by a random-average technique (Section 3.3.1.3). To avoid generating a "mosaic" realization in some areas, the number of Poisson centers must greatly exceed the number of grid points in the simulation domain, although probably many fewer points would adequately reproduce the sample variogram. If we use a spectral method for the unconditional simulation, we should consider reproducing the spectral density function for all wavelengths greater than or equal to twice the minimum grid spacing (corresponding to the Nyquist frequency). In a turning-bands simulation, the spacing of simulated values of the line processes should be dense enough so that adjacent grid points will not receive contributions from exactly the same set of simulated line points. (Subroutine TB3D determines an appropriate spacing automatically.) In a two-dimensional turning-bands simulation, we also can choose the number of lines that are simulated. Mantoglou and Wilson (1981, Section 3.6) have carefully explored the effects of the number of lines, discretization along lines, and spectral discretization for the line processes, and have concluded that in practical two-dimensional turning-bands simulations, the finite number of lines is the <u>main</u> source of simulation discretization error. Nevertheless, they do not recommend more than eight evenly spaced lines for most practical simulations, and their reasoning leads us to a final important consideration:

Before choosing to endure the expense of simulations on a very fine grid, spectral discretizations with many fine frequency increments,

semivariogram models with complicated nested structures and anisotropies, two-dimensional turning-bands simulations on a great number of lines, and numerous other niceties, we should always carefully consider the quality of the data and of the model we have constructed from the data. It is not worthwhile to drive the simulation discretization error SDE (which we can usually control) to the vanishing point when some of the components of the total functional error TFE (most of which we cannot control without more data) are likely to be several times larger:

> There is a tradeoff between accuracy and cost. A more accurate model will be more expensive. In practical . . . applications, the imposed (theoretical) statistics are very often obtained from a limited set of data points in the field and a limited number of realizations (very often only one). So it is expected that the errors in the estimation of the covariance function, etc., are quite significant and the true covariance function is different from the one we fit with a model. [This is the total functional error, TFE[C(h)]; see Figure 37.] Consequently, in the simulation of the process, we may not always be justified in using a very exact (and thus more expensive) generation model, by increasing the number of lines, etc. How well should we preserve the model covariance function, when it does not represent exactly the reality? It would be wiser to choose the parameters of the model such that there is consistency between these input data estimation errors and the simulation model errors. [Mantoglou and Wilson, 1981, p. 96]

In practice, a numerical value for the variance of the total functional error is not generally obtainable (at least not using the sketchy ideas presented here), but the discussions in Sections 3.7.2 and 3.7.3 should provide an intuitive feeling for how much detail and expense are justifiable in order to minimize the simulation discretization error.

3.7.5 Conditioning Methods and Applications

Several of the advantages of "good conditioning" have already been emphasized in previous sections of this chapter. All of these advantages can be summarized in one general remark: conditioning imposes the local unmodeled idiosyncrasies of the conditioning data onto a simulation of a more broadly specified phenomenon. Conditioning is especially beneficial if the spatial domain of the simulation is small with respect to the variogram range. In such a case, different unconditional simulations tend to differ strongly from one another, particularly in their large-scale features, and thus may depart severely from the characteristics of the real data in the area of interest. (These problems are discussed at greater length in Section 3.7.2.) Conditioning is also helpful if the specification of the model is in doubt, or simply in error. Characteristics of the phenomenon that occur on a scale larger than the typical spacing of the conditioning data will be imposed on the simulation by conditioning, whether incorporated into the model or not.[73]

What constitutes "good conditioning"? To answer this question, we first have to decide how accurate our simulation is supposed to be; i.e., we must assign a maximum tolerable value to the total simulation error, $TSE=[Q^*_{s(vs)}-Q_{0(vs)}]$, for each characteristic Q that is of interest. Then we must ask whether this maximum value can be respected,

---

[73] For instance, in the case-study results summarized in Section 4.2.4, the short-scale structure of one of the variables, PCTDOL, was smoothed out by an inadequate despiking procedure. The semivariograms of the simulated data are parallel to those of the original data, as a result of conditioning. However, conditioning did not correct the bias in the short-scale (essentially nugget) structure of the simulation, so the semivariograms of the real and simulated data do not coincide.

assuming that we employ the best available procedures to create a conditional simulation using the available data. If the total simulation error cannot be kept under control even with the best procedures, then we probably do not have adequate conditioning: more data, carefully placed, are needed to reduce some or all components of the total simulation error.[74] Unfortunately, a quantitative assessment of the total simulation error, or of most of its components, is seldom possible; so the decision of whether the total error is under control and, if not, what must be done to reduce it, is a subjective one. Fortunately, most deficiencies should be obvious during preliminary inspection and structural analysis of the data. Gaps in the data coverage leading to large domain errors are usually obvious, and it is the domain errors DE and MDE that can be most readily reduced by an expansion in the conditioning data base.

Guidance in finding the best locations for additional conditioning data can be obtained from a map of local kriging variances. The true characteristics of areas with high kriging variances will tend to be less faithfully reproduced in a conditional simulation, and multiple independent conditional simulations (using independent unconditional realizations) will tend to differ from one another most strongly in these areas. Areas with high variances will usually be those with a low density of nearby data, or with data only on one side. To prevent a

---

[74] Additional data may be needed for purposes other than just conditioning. For instance, reduction of the estimation error for variogram behavior near the origin cannot be improved by conditioning, because the error occurs at too short a scale. However, additional data from twin holes or a cross or fence of holes will usually reduce this error throughout the simulation domain, although it will improve conditioning only very locally.

simulation from running wild at the edges of the simulation domain, it is always a good idea to place some conditioning data on or just outside the boundaries. Reproduction of "global" characteristics, such as the average grade or total quantity of metal within a very large domain, is not ordinarily an important objective of a simulation. Consequently, the rather complicated procedures that have been devised for locating additional data for reduction of global estimation variance are not usually pertinent in the location of data for conditioning.

In addition to its primary advantage of introducing unmodeled characteristics of the data into the simulation, conditioning affords an opportunity to introduce subjective information into the simulation, either by ordinary kriging using fictitious conditioning data, or indirectly by cokriging, "soft" kriging, or universal kriging with subjective shape functions (all discussed below). Observed regionalized variables that have not been simulated but are cross-correlated with simulated variables can also be used to improve the quality of conditioning by means of cokriging. In general, the decision of whether to introduce more data or whether to use a more elaborate kriging method for conditioning a simulation should be based on the same considerations that are important in devising a solution to an estimation problem. The remainder of this section provides some details on alternative conditioning methods.

Fictitious data. Fictitious (or "subjective", to use a more polite term) conditioning data for use in ordinary kriging should certainly be used with care. There should be a very good reason to believe that a fictitious value is very close to the correct value, particularly if its

location is inside the simulation domain. Otherwise the conditional simulation will be biased. Fictitious data are perhaps easier to justify if their locations are outside the region of direct interest. For example, if there are good geological reasons for believing that a regionalized variable has a certain average value or exhibits a certain trend in a border area of the simulation domain where there are no real data, then a fringe of fictitious conditioning data, designed to force such an average or trend onto the kriged estimates in this area, might be placed just outside the simulation domain (Figure 50) to "tie down" the conditional simulation in the border area. In the case of a trend, both the positions and the values of the fictitious data will influence the trend. A map of the kriged values is thus helpful in visualizing where to add the data, what their values should be, and how many such data are needed.

"Soft" kriging. If subjective conditioning data are placed within the simulation domain, it is probably better to consider these data as a separate variable to be used in cokriging. (This requires a subjective cross variogram as well.) A more elaborate extension of this approach would be to use "soft" kriging (Journel, 1984e), which allows us to use subjective (prior) probability distributions, or incompletely known distributions (as in the case of inequality constraints, described below), instead of deterministic fictitious data at several unsampled points in and around the simulation domain. A series of krigings or cokrigings using known and "soft" data can then be performed to obtain estimated probability distributions of the variable being simulated at all points in the simulation domain. To select a few fixed values for

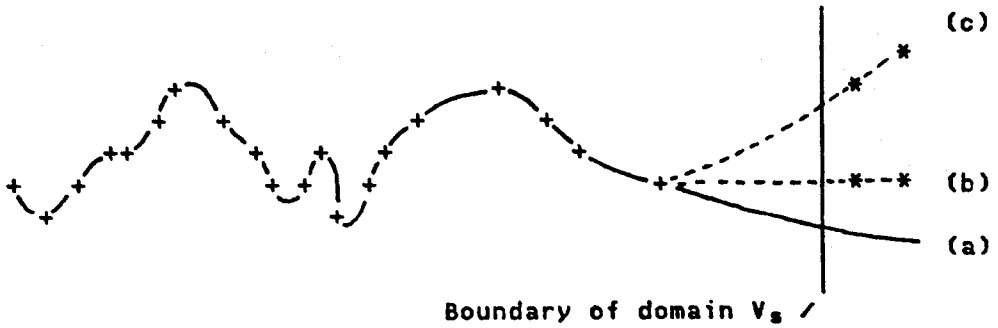


Figure 50: An illustration of the use of fictitious data in the conditioning of a time-series simulation. (a) Kriged values obtained using a limited data set that does not cover the whole simulation domain V_s may extrapolate undesirable local trends in the data into outlying regions where there are no data. (This is particularly likely if the phenomenon is very continuous.) (b) Fictitious data are used to remove an undesirable local trend in the kriged values. (c) Fictitious data are used to force a local trend onto the kriged values.

conditioning the simulation, we might extract the expected value from each predicted probability distribution at each new conditioning point, or draw a value at random from its probability distribution. This relatively complicated approach could be justifiable for the conditioning of a badly undersampled variable that cannot be effectively cokriged using existing real data, or for situations in which a great deal of important qualitative information is available. An application of soft kriging is described by Kulkarni (1984).

Kriging with inequality constraints. In two related papers, Dubrule and Kostov (1984) and Kostov and Dubrule (1984) present an approach for taking inequality data into account in an estimation. Their method is

useful in situations where we do not know the exact values of a regionalized variable $z(x)$ at some locations x_i, but we do know upper and/or lower bounds for these values (i.e., we know $z_{1i} \leq z(x_i)$, $z(x_i) \leq z_{2i}$, or both). The most common situation in which such constraints occur is in an estimation of the elevation of a stratigraphic horizon using drill-hole data. In some locations the holes may not penetrate all the way to the horizon of interest, but it is known that the horizon must lie below the bottom of the hole. In a simulation context, this situation could easily occur if we were simulating a contact between two populations as a regionalized variable (Section 3.8.2). To assure that the kriged estimates $z^*(x_i)$ satisfy the known constraints at all locations x_i where we have inequality data, we can reformulate the kriging as a quadratic-programming problem (Hillier and Lieberman, 1980, p. 751). Kostov and Dubrule (1984) provide examples of how this formulation can be applied to the mapping of subsurface elevations and bed thicknesses.

This problem can also be regarded as a special case of the "soft" kriging problem and is approached from that point of view by Journel (1984e).

Subjective trends and universal kriging. Omre and Holden (1984) have suggested an application of universal kriging (Journel and Huijbregts, 1978, p. 313) that allows subjective information on the forms of drifts and discontinuities to be taken into account in an estimation (or conditioning) problem. Instead of the polynomial drift functions usually assumed in universal kriging, Omre and Holden allow arbitrary "shape functions", including discontinuous functions (e.g., representing

faults) to be incorporated into the estimation. This approach is helpful if we have a good subjective notion of the overall shape that a realization should take within some area where we have few data, but we do not necessarily have any idea of what the values should be at any points in that area, as in the previous two approaches.

Cokriging. In practice, ordinary kriging usually is employed to condition simulations of coregionalization, simply because cokriging is more complicated and expensive. If all of the simulated variables have been observed at all of the conditioning points, and if the unconditional simulation of coregionalization has been properly modeled, then little has been lost by using this shortcut. In the special case of intrinsic coregionalization, nothing has been lost, as ordinary-kriging and cokriging estimates are then the same if the data configurations are the same for all variables (Journel and Huijbregts, 1978, p. 326). However, if all simulated variables have not been measured at all conditioning points, cokriging may be necessary to control the simulations of the undersampled variables.[75] Otherwise the cross correlations among the variables that are built into the unconditional simulation of coregionalization may be altered as different variables are conditioned by different configurations of data (Figure 51). Of course, if some variables are not simulated at all but are nevertheless associated in some way with the simulated variables, improved correspondence of the real and simulated phenomena might still be obtained by cokriging using all data. This would correspond to the example in Figure 51 if we were simulating the lower variable only.

---

[75] Another serious difficulty with undersampling is discussed at the end of Section 3.7.8.

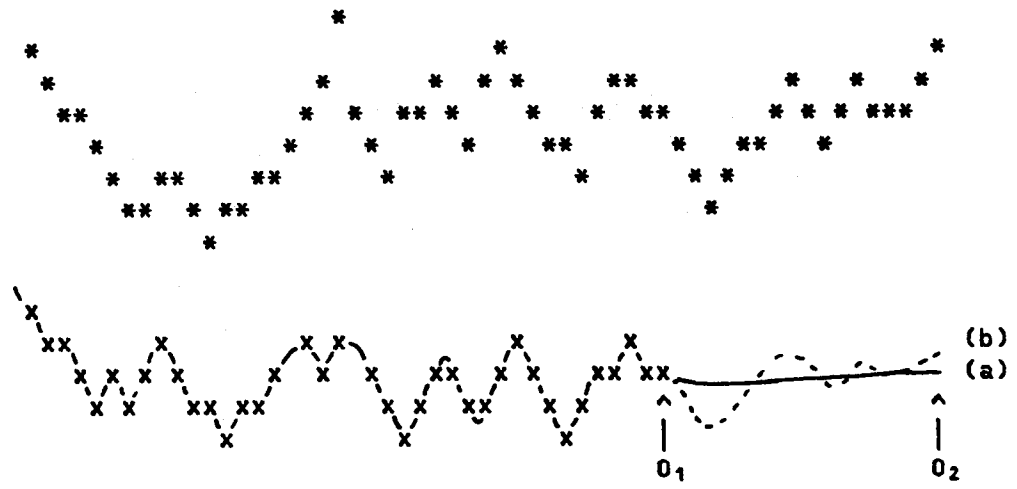


Figure 51: Estimation of coregionalized time series by ordinary kriging and by cokriging. In this example, the lower series has been undersampled with respect to the upper one, with which it is positively correlated. In case (a), the additional data of the upper variable have not been used in the kriging (solid line O_1O_2) of the lower variable. In a simulation, the cross correlations among the variables at distances longer than the data spacings would be largely removed by conditioning in this way, although very short-scale fluctuations, which show up only in the kriging errors, would still be cross-correlated. In case (b), the longer-range cross correlation is restored by cokriging (dashed line O_1O_2).

Are there any situations in which conditioning is not necessary? If we are concerned with a phenomenon of which multiple realizations are physically possible (e.g., the regionalization of rainfall from passing thunderstorms), then observed data from one realization may contain no information about another except for gross statistical characteristics (e.g., histograms and variograms that can be expected for rainfall from a storm of a certain magnitude). In such a case, a conditional

simulation using data collected from one realization would grossly underestimate the diversity of results that could be expected from other realizations. (However, if the terrain the thunderstorms are crossing is irregular, certain trends in rainfall might be applicable to most realizations.) In geological phenomena of which only one realization is possible, it is ordinarily much safer to perform a conditional simulation, unless the simulation domain is very, very large with respect to the variogram range and we do not intend ever to use any part of the simulation to represent any particular part of a real physical domain. In mining applications, the range is almost always large in comparison to the dimensions of the domain, and we usually have definite plans to start mining in one particular place and chew our way into several others, in something like an orderly manner. Therefore, for mining applications, a good rule of thumb is: don't even think about using an unconditional simulation for anything except input to a conditioning procedure. Furthermore, the "better" the conditioning (generally, the more evenly and densely spaced the conditioning data), the more realistic, useful, and trustworthy the results.

One common misconception about the applications of unconditional simulations in mining can now be readily cleared up. Suppose we simply wanted to find out what an "average" year's production from a certain mine would look like in the mill. Could we not use the results of an unconditional simulation as input to a mining simulation to produce a realization of an average year's production? No, we could not, because in a typical mine in which the range of the variogram is large in comparison with the spatial extent of the reserves, there simply is no

such thing as an "average" year. The phenomenon is not ergodic; therefore, it is entirely possible that no two years' output will ever look the same in the mill. This is something to keep in mind when designing mineral-processing facilities: the plant must be able not only to handle variations in the ore on the scale of minutes, hours, and weeks, but also to adapt itself to very long-range fluctuations possibly on the scale of decades and possibly involving very large changes in the "average" characteristics of the ore. Conditional simulations of entire ore deposits or entire properties should therefore prove valuable in planning both the short-range homogenization and proportioning capabilities of the plant and the long-range flexibility (particularly in proportioning) that will be required. In this sense, the output of a conditional simulation and associated mining model might be superior to actual historical quality data from an existing mill using raw materials from the same mine.

3.7.6 Selection of Methods for Unconditional Simulations

In Section 3.3, several methods were proposed for the unconditional simulation of realizations from a specified gaussian random function. Now that the most important errors that can creep into such a simulation have been described, it is possible to make a few recommendations for choosing among these methods. The major considerations seem to be the characteristics of the random function (particularly the semivariogram model), the tolerable simulation discretization error SDE (particularly as it relates to the choice of simulation grid spacings), the size and dimensionality of the simulation domain, the execution speeds of the

various algorithms and specific computer programs that are available, and of course the simple availability of the programs or the complexity of writing one.

Most of the advantages and disadvantages of each method already have been stated in Section 3.3. The following paragraphs recapitulate those remarks and add some new comments based upon the results of Sections 3.7.1-3.7.4.

Matrix methods. For very large simulations of geological phenomena, these procedures may be too slow on most computers and may demand too much array storage, although for some types of matrices (Davis, 1985a) storage requirements can be reduced considerably. These methods have the advantage that repeated simulations (Section 3.7.7) can be easily performed, and repeated conditional simulations, performed without kriging, are possible using the conditional mean and covariance matrices provided on page 109. Matrix methods might thus prove useful for generation of the repeated unconditional line simulations required for turning-bands simulations. A final advantage of matrix methods is the ease of simulating data at irregularly spaced locations (e.g., to coincide exactly with the locations of conditioning data).

Moving-average methods. The execution time of a moving-average simulation is roughly proportional to the number of grid points to be simulated times the number of grid points spanned by the weight function.[76] (In subroutine CS2D, this means the product of all dimensions of the arrays "Y" and "W".) If the window is large in

[76] Keep in mind, however, that the span of the weight function (range of the variogram model) can usually be reduced, with a consequent saving of computer time, if closely spaced conditioning data are available. This is discussed in Section 3.7.2 and illustrated in Figure 45.

comparison to the simulation domain, the time becomes even greater, because many random numbers must be generated at grid locations outside the domain. Hence very large two-dimensional simulations are expensive, and three-dimensional simulations may be virtually ruled out. However, if the weight function is constant-valued, as in the case of hyperspherical semivariogram models, it may be possible to save considerable execution time by means of an "updating" algorithm, as explained briefly in Section 3.3.1.2. The main advantages of moving-average methods are their simplicity and their accuracy. For weight functions with finite spans (corresponding to finite variogram ranges), no approximations are made except for discretization of the weight function and simulation domain. Thus the "striping" that can occur in a turning-bands realization (discussed below) cannot occur in a moving-average realization.

Random-average methods. Although Chiles (1984) was able to reproduce his model semivariogram functions satisfactorily using a relatively small number of Poisson centers (3000 to 10,000 centers for a two-dimensional grid of about 12,000 points), it is apparent that such realizations would yield some areas with "mosaic" structure (Section 3.3.1.3). This could be dangerous if the simulation is used to investigate the effects of very small-scale fluctuations in the realization. Conditioning does not cure such small-scale inaccuracies: a conditioned random-average simulation would contain "tiles" with smoothly varying internal values instead of the constant values turned out by the unconditional simulation. Back-transformed values would generally be smooth as well. Thus the random-average method seems to be

weakest precisely where an unconditional simulation method should show its greatest strength: in reproduction of the short-scale variability of the model random function. Its use is thus not recommended, except possibly in two cases: (1) simulations incorporating a large nugget component that would add some short-scale randomness, and (2) simulations in which reproduction of short-scale variability is of little importance. In other cases, moving-average or turning-bands simulations would be more reliable.

Turning-bands method. This method achieves its great speed by reducing a two- or three-dimensional simulation to a set of much faster one-dimensional simulations. Its superiority to the moving-average methods lies solely in its speed. Mantoglou and Wilson (1981, 1982) also demonstrate that it is faster than the spectral methods of Shinozuka and Jan (1972) and Mejia and Rodriguez-Iturbe (1974).

There is some loss in accuracy owing to the finite number of line simulations. This seems to be the source of the "striping" -- the appearance of nebulous cross-hachured zones of predominantly high or low values -- that can be observed in two-dimensional layers extracted from three-dimensional unconditional turning-bands simulations. This effect can be visualized in Figure 52, where zones of very high or low values occurring by chance on the lines may exert a strong influence on all values occurring within the associated turning bands. Striping is particularly easy to see if one extracts a plane of simulated data from a three-dimensional turning-bands simulation and performs a cutoff on the simulated values -- attributing an indicator "1" to all values above the cutoff, "0" to all values below. If the cutoff is placed at one of

the upper quantiles of the simulated distribution, a few vague linear zones containing mostly 1's will be visible on a map of the indicators. Each zone represents the contributions of a zone of very high values on one of the lines. Parallel zones are contributed by the same line. Wide zones may have been contributed by very wide anomalous zones on the lines, or possibly by a line that is nearly orthogonal to the observed plane. Fortunately, such striping is easily removed by conditioning, so it is only of concern (if then) in three-dimensional unconditional simulations. In two dimensions, striping can be obliterated by increasing the number of lines, thus diluting the effects of extraordinary zones along any one line.

Two-dimensional turning-bands simulations have been underutilized in the past because of the difficulty of deriving appropriate models for the associated line processes. The spectral method of Mantoglou and Wilson (1982) and the one-dimensional covariance solution of Brooker (1985) have cleared up this difficulty, so two-dimensional simulations should now become more popular.

Autoregressive methods. This approach is convenient for discrete time-series simulations of AR(p) processes conditioned only by a set of "p" initial data. Higher-dimensional autoregressions may not be realistic, are relatively difficult to model and to perform (although the new approach of Sharp and Aroian (1985) appears much simpler), and are not well suited to conditioning by kriging because of the peculiar forms of the covariance functions.

Frequency-domain approaches. Mantoglou and Wilson (1981, 1982) have amply demonstrated the advantages of the turning-bands method over the

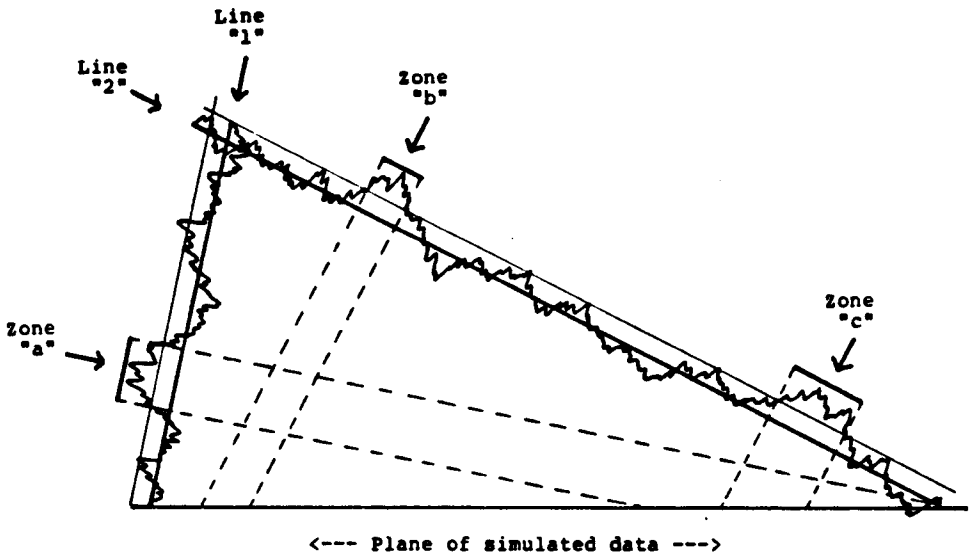


Figure 52: Origin of "striping" in an unconditional simulation produced by the turning-bands method. This is a side view along a plane of simulated data from an unconditional three-dimensional turning-bands realization. Two lines, numbered "1" and "2", are shown above the plane, along with graphical representations of the simulated line realizations. An "anomalous" zone "a" on line 1 contributes a slice of anomalous values to all points in space lying within the associated turning bands. The intersection of these bands with a plane extracted from the simulated domain creates a stripe of high values, not adequately diluted by contributions from the other fourteen lines. Anomalous zones "b" and "c" on line 2 produce parallel stripes on the plane, which will tend to be narrower than the stripes associated with line 1 because of the higher angle between the associated bands and the plane.

spectral methods of Shinozuka and Jan (1972) and Mejia and Rodriguez-Iturbe (1974). Borgman et al. (1984) claim great speed for their method, but direct comparisons with other methods have not been

made, and the method will not accept a great number of simulated layers in the third dimension. Spectral methods can be useful for turning-bands line simulations, as demonstrated by Mantoglou and Wilson (1982), although for some covariance models (e.g., the spherical model) the corresponding spectral formulations are mathematically complex. The finite-Fourier-transform program of Davis et al. (1981) may see increasing applications in turning-bands simulations, particularly now that Brooker (1985) has provided a means for deriving the required one-dimensional covariances for two-dimensional simulations.

Spectral approaches are more convenient for the simulation of very continuous phenomena, which have predominantly low-frequency variability. Covariance models with short-scale linear behavior correspond to long-tailed spectra, which can be reproduced only inaccurately or expensively by spectral simulation methods. A simulation discretization error can be committed by discretizing the spectrum into too few harmonics -- either by truncating the spectrum at too low a frequency or by spacing the harmonics too far apart. (However, the spectrum should be truncated below the Nyquist frequency.) Of course, nugget effects are not at all suited to spectral simulation.

3.7.7 The Advantages of Repeated Simulations

Because conditional simulations are not unique, it is important to determine under what conditions one simulation will be just as good for our purposes as any other, and under what conditions it would be preferable to compare the results of several independent simulations. Independent simulations, using the same model and methods but

independent sets of random numbers, will all look practically alike under either of two circumstances: (1) the conditioning is uniformly very dense in comparison to the scale on which the phenomenon is being investigated, or (2) the simulation (conditional or unconditional) covers a spatial domain that is very large in comparison to the variogram range. The second case requires ergodicity: the overall, but not necessarily local, characteristics of a single realization must approach the average characteristics of a large suite of independent realizations as the size of the simulation domain increases. In a way, the first circumstance is a special case of the second. In conditioning we impose the large-scale fluctuations of the conditioning data onto all conditional simulations. The only differences that remain are in the short-scale variability of the residuals. In a well conditioned simulation, the range of the residuals is much shorter than that of the original realization, and thus usually small in comparison to the size of the simulation domain. Notice that if we simulate within a large domain but subsequently make use of only a part of the domain, we give up this advantage.

For cases in which the above circumstances do not apply, including many cases in mining practice, different simulations may yield noticeably different results, so independent replications of the simulation should be considered. But how should we use these replications? The futility of merely <u>averaging</u> repeated simulations is explained at the end of Section 3.1.7: the average of a suite of conditional simulations is just an approximation of the kriging. (For unconditional simulations, the average approximates the stationary

mean.) Furthermore, the variance of the simulated values at a given point is just an estimate of the kriging variance (or of the sill, for unconditional simulations). In other words, mere statistical summaries of repeated simulations of regionalized variables tell us nothing that a kriging would not tell us more easily.

If we use each independent conditional simulation as input to some other kind of simulation (e.g., a mining simulator operating on simulated ore grades, a reservoir simulator operating on simulated porosities and permeabilities, or a drainage-basin simulator operating on simulated rainfalls), the statistics of the final output may be exceedingly hard to predict without actually plugging in several conditional simulations and watching what comes out. A good example in the field of hydrology is provided by Delhomme (1979), who discusses the influence of different conditional simulations of log-transmissivity in an aquifer on the spatial distribution of hydraulic head values. Different realizations of the head values are obtained by applying a numerical model of groundwater flow to different simulated log-transmissivity fields.

It is important to realize that this usage of repeated simulations reveals the influence of only one source of variability on the final output: variability among the independent sets of random numbers used to generate independent simulated realizations of the model random function. The same model is used in each case. Obviously, if the range or overall shape of the variogram, the size of the nugget constant, the form of the raw histogram (or scattergram at h=0), the reliability of the conditioning values, and numerous other parameters are in question,

then the total uncertainty in the final output may be much greater than that indicated by rerunning several simulations with fixed parameters.[77] It would be possible to design an experiment to test the sensitivity of the final output to all parameters in question, using an analysis-of-variance approach. However, such an experiment would be cumbersome to conduct and would consume vast chunks of computer time, so a more practical approach is to begin with good estimates of all parameters. Of course, good estimates are obtained from good data sets, and good data sets also promote good conditioning. As pointed out above, the better the conditioning, the better confidence we can place in a single simulation. Extremely unstable results for the final output of a study, as demonstrated through the use of repeated simulations, might thus be used as a good argument for gathering additional data.

If there are large gaps in the simulation domain with few or no data, the local kriging estimates (from both real and simulated data) inside the gaps will be very smooth. In such a case, the unconditional simulation imposes not only its small-scale features onto the conditional realization, but its larger-scale features as well. The sample semivariogram of an unconditional simulation is much more subject to fluctuation variance at these larger scales. Consequently, the characteristics of sparsely conditioned regions may vary appreciably from one realization to another. In some situations, there may be enough qualitative information about a sparsely conditioned area to allow one simulation out of several replicates to be selected as somehow

---

[77] This additional uncertainty corresponds essentially to the functional model error FME of Figure 37, whereas repeated simulations differing only in their random numbers would differ by amounts equal to the functional realization error FRE.

"more reasonable" than the others. This would be an alternative to the use of subjective conditioning methods, described in Section 3.7.5. Selection of particular realizations might also be made on the basis of variogram reproduction, which can be poor if the simulation domain is small, if the semivariogram model is not microergodic, or if the original data are not gaussian (Section 3.7.2).

3.7.8 Transformation Errors

Because earth-science data rarely appear to have been drawn from realizations of gaussian processes (Section 3.5.1), it is usually necessary to transform the data before a gaussian simulation can be performed. The form of the transformation usually is inferred from the available data, with the aid of the multigaussian hypothesis (which, although critical, cannot be fully checked), and perhaps with the aid of some knowledge of the physical processes involved. If the transformation does not correctly yield a sample from a gaussian realization, then a gaussian simulation (the only kind considered here) cannot correctly reproduce the entire spatial variability of the phenomenon. Furthermore, the lack of microergodicity in nongaussian realizations makes unconditional simulations particularly hazardous and the short-scale behavior of the realization questionable even in conditional simulations. This problem is discussed at greater length in Section 3.7.2, page 219.

In practice, it has often been observed (but not well discussed in the literature) that the variogram of a conditional simulation reproduces the variogram of the fully transformed "gaussian" data very

well, but discrepancies appear between the variogram of the back-transformed simulation and that of the raw data.[78] In other words $[\gamma^*_{s(vs)}(h)-\gamma^*_{o(vo)}(h)]$, which is the sum of the total model error and the simulation discretization error in $\gamma(h)$, is small for the gaussian data but large for the raw data. This suggests that the transformation has been specified incorrectly and thus the transformed data are not really multigaussian -- not a surprising development, considering the serious inference problems that are discussed in Section 3.5.1 and in the subsequent descriptions of individual transformations in Section 3.5. This can be regarded as a type of model estimation error, but one that is committed at a very early stage in a simulation project and not discovered until the simulation has been completed. (With reference to Figure 34, the error is committed prior to creation of "transformed data set number 2" and usually not discovered until the very last step, labeled "check results".) Because so much work may be wasted before the poor quality of the result is revealed, it is important to insure that transformations are modeled correctly in the beginning.

Transformations are usually constructed by trial and error. We search for a transformation of the raw data that produces a transformed data set with the characteristics of a gaussian realization (essentially, a gaussian distribution at h=0, and gaussian h scattergrams). We hope that the inverse of this transformation, applied to a simulated gaussian realization modeled after the

---

[78] The faulty reproduction of the "PCTDOL" semivariogram discussed in Section 4.2.4 is a fine example. In that case, the despiking technique (a type of transformation) applied to the original data caused the problem.

transformed data, will then reproduce the characteristics observed in the raw data. An error occurs if the forward transformation does not really produce a sample from a gaussian realization. There are two general ways to mitigate this problem: (1) perform thorough checks on the transformed data, confirming either that nongaussian characteristics have been eradicated or that some other transformation is needed, or (2) collect many more data. A large data set is important because both the checks and the estimation procedures necessary to derive some of the transformations are sensitive to the curse of dimensionality (Section 3.5.1). Furthermore, as the density of conditioning data increases, the conditional simulation will conform to the characteristics of the original realization more closely. This makes the simulation more robust to all kinds of modeling errors, including transformation errors. The following paragraphs briefly summarize the major pitfalls involved in the various transformations proposed in Section 3.5.

The fundamental danger in the usage of the gaussian transformations in Sections 3.5.2.2 and 3.5.2.3 is the danger that the multigaussian hypothesis -- the hypothesis that normality at h=0 implies normality at h≠0 -- is invalid. It may be the case that no transformation exists that would convert the observed realization to a gaussian realization, or perhaps a transformation does exist but is too complicated to be inferred from the available data. If the best procedures are used to construct a transformation at h=0, but checks on the distribution at h≠0 reveal that the spatial distribution still is not gaussian, there appears (up to now) to be little else that can be done in most cases,

except to hope that conditioning will make the simulation robust to the remaining discrepancies. Transformations of individual h scattergrams are not feasible, as the same data appear in many scattergrams. Simultaneous iterative transformations of many h scattergrams are conceivable but hardly seem practical. In some cases, transformations of the spatial domain of the data (and simulation), rather than of the data values, might improve the normality of h scattergrams, but no theory on how to derive such transformations seems to be yet available. (Unfolding" of deformed deposits prior to geostatistical calculations is a commonly used transformation of this type, although used for a different purpose.)

Some messy phenomena commonly observed in nature may not be transformable into gaussian realizations by any means. Examples of apparently hopeless cases include almost all natural landscapes, such as those contoured in Figure 53. The local kinks and other peculiarities (which might be viewed as "systematic nonstationarities") in the shapes of these contours seem to require highly parameterized models that could not be produced merely by transforming the distributions of the surface elevations to a gaussian distribution. Methods that would reproduce multivariate properties, such as multidimensional h-scattergrams among $\{z(x), z(x+h_1), z(x+h_2), \ldots\}$, would come closer to reproducing these phenomena, but such methods are unavailable. The "fractal" approach of Mandelbrot (1977, 1982), which is more directly concerned with modeling the |shapes| of natural phenomena, may be helpful in some of these cases, although again only a few parameters are being modeled. Simulation of the physical processes that formed the phenomena may

provide another approach (e.g., Harbaugh and Bonham-Carter, 1970), but there seems to be no way of conditioning most simulations of this type.

Certainly it is advisable to make the multigaussian hypothesis as justifiable as possible by insuring that the distribution at h=0 really is normal. For instance, if we have a multivariate data set, it could be dangerous to depend only on independent transformations of the marginal distributions (Section 3.5.2.2) if checks on multinormality at h=0 reveal serious departures from gaussian behavior in the transformed data. The most comprehensive multivariate transformation at h=0, the stepwise-conditional approach (Section 3.5.2.3), unfortunately requires a good estimate of the true multivariate distribution at h=0 and thus is thinkable only when there are many data, or estimation procedures requiring fewer data, or a thorough knowledge of the natural phenomenon that allows some characteristics of the distribution (e.g., constant-sum or mineralogical constraints) to be specified without data. In practice, less complicated transformations such as the one proposed by Chiles (1984, and Section 3.5.2.3) probably will suffice in many cases.

It is important always to remove any known constraints (Section 3.5.2.4) on the raw data prior to further transformations, as the constraints cannot be accurately respected by a gaussian simulation and probably cannot be accurately modeled by a smooth data-directed transformation or density-estimation technique. Constraints can be removed by the "graphical" transformation procedure (Section 3.5.2.2) for univariate data.

All multivariate transformations that involve combinations (linear or otherwise) of several coregionalized variables carry a risk that small

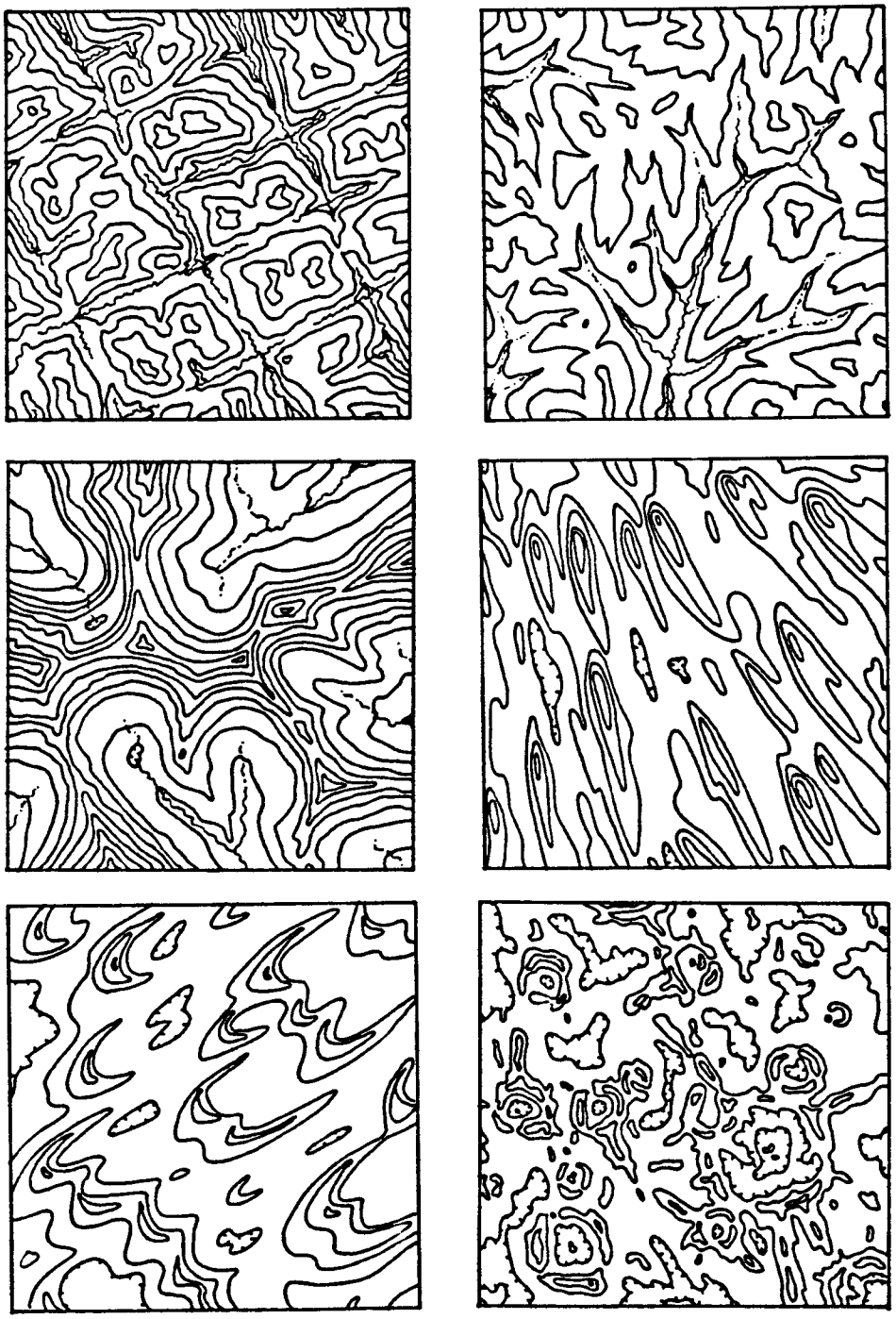


Figure 53: Some examples of landscape topographies that would be difficult to transform into stationary gaussian surfaces. Clockwise from upper left: rectangular drainage, dendritic drainage, drumlins, impact craters, barchan dunes, and alpine glaciation.

but potentially important characteristics of the original variables may become so diluted in the transformed variables that they will not be recognized in the transformed data. For instance, the principal-components transformation for simplifying the covariance matrix (Section 3.5.3) and the successive-remaining-space transformation for elimination of the constant-sum constraint (Section 3.5.2.4) yield some transformed variables that include contributions from most or all of the original variables. It is possible then that the influence of some individual components of a complicated nested variogram structure in one of the original variables may not be observed or modeled in the transformed data. Actually this may not be serious, as conditioning should impose all but very short-scale structures onto the final simulation. Moreover, there may not be anything uniquely valuable about the original data: we might argue just as forcefully that models constructed from the original data would be deficient if they did not embody structures that appeared in any arbitrary combination of the data. The characteristics that should be checked most carefully upon completion of a conditional simulation are the characteristics of variables that will be important when the simulation is actually used. Combinations of the original variables, rather than the raw variables themselves, may be of primary importance in applications. For instance, in cement quality-control applications, checks on the reproduction of the distribution of Bogue C_3S (Table 1) may be of much greater importance than reproduction of the individual characteristics of CaO, SiO_2, Al_2O_3, Fe_2O_3, and SO_3, if C_3S is actively used in quality control and its component oxides are not.

Most of the multivariate transformations described in Section 3.5 involve all k coregionalized variables, either simultaneously or in a stepwise fashion. Thus serious problems may occur if some of the k data are missing at some locations. If stepwise transformations are being used (e.g., the stepwise-conditional gaussian transformation of Section 3.5.2.3 or the successive-remaining-space transformation of Section 3.5.2.4), and the missing observations always occur in the same variables, then these variables should be transformed last. Cokriging can then be used to estimate missing transformed conditioning data from the other available data. If simultaneous transformations are being used (e.g., the principal-components transformation of Section 3.5.3), the missing data might be simulated by some approximate method prior to the transformation, or just replaced by estimated values if there are only a few missing data. The same approximations would be necessary in the case of sporadic missing values that occur in different variables at different locations. Some errors clearly would result in these simulations, but if there are few missing data, they probably would not be serious. Simultaneous transformations probably should be avoided if there are many missing data.

3.8 MODELING MULTIPLE POPULATIONS

3.8.1 The Problem of Multiple Populations

Geostatistical methods were invented to describe realizations of second-order-stationary random functions, although stationarity of the mean can be relaxed if the mean is a smooth drift function. However, geostatistical methods break down entirely if the mean, or any other characteristic of the phenomenon of interest, varies discontinuously in

space, i.e., if the phenomenon is actually a mixture of different realizations from one or more random functions. Then, if we are to use geostatistics successfully, we must isolate these different realizations and treat each as a separate population for which model inference, estimation, and simulation are performed separately (although not necessarily independently). This section describes ways for partitioning the simulation domain into subdomains containing different populations. The simulation domain is assumed here to be a three-dimensional rectangular prism; the procedures to be used in lower dimensions are analogous. The contacts between the different populations may be sharp or gradational, and the geometry of the contacts may be essentially deterministic (e.g., flat or regularly folded bedding planes) or partly random (e.g., the irregular boundary of a reef deposit, or an erosion surface).

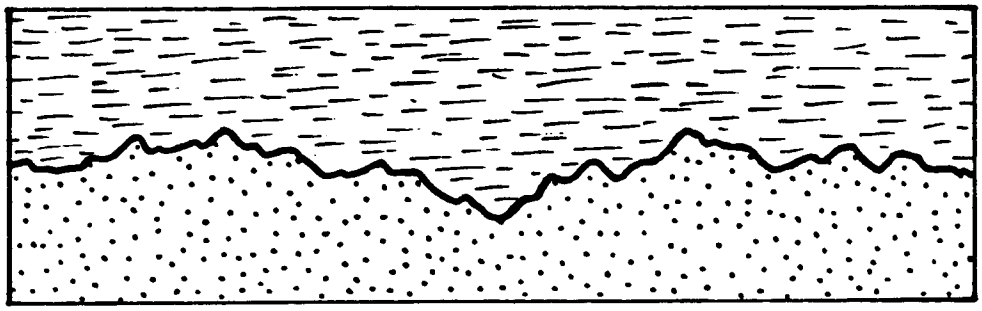
If the contact between two populations is sharp, well sampled, and essentially smooth at the scale of investigation, then it possibly can be modeled deterministically by a smooth surface-fitting technique or by the geometrical methods used in structural geology (e.g., Ragan, 1973).[79] If the contact is sharp but so irregular that deterministic methods are neither reliable nor realistic, then geostatistical means might be used to estimate the contact (in the context of a geostatistical estimation problem), or to simulate it (in a simulation problem). Then two situations arise: (1) the "elevation" (in some

---

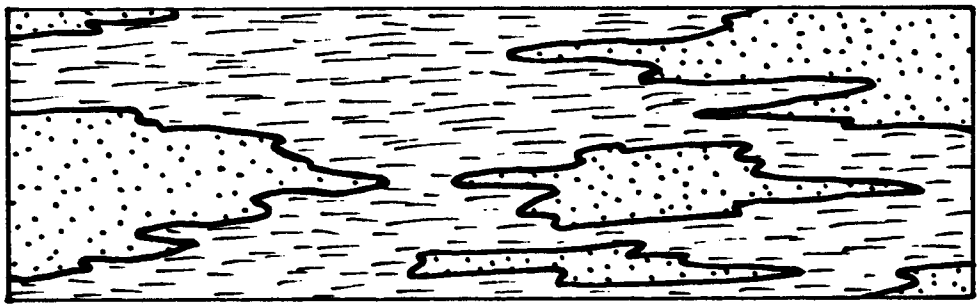
[79] If the rocks have been deformed internally by folding or faulting (i.e., if not only the boundaries between populations but also the internal spatial domains of the populations have been deformed), an "unfolding" technique (e.g., Dagbert et al., 1984) should be applied to rectify the coordinate system before geostatistical work begins.

coordinate system) of the contact surface in three-dimensional space can be regarded as a regionalized variable in two-dimensional space (Figure 54(a)), or (2) the contact is too convoluted for the first approach and thus must be modeled in three-dimensional space as the boundary between excursions of a regionalized variable above or below a cutoff value (Figure 54(b)). The second situation arises only if the contact cannot be modeled as a function of two spatial coordinates -- i.e., if the contact can cross a given coordinate location more than once. Sections 3.8.2 and 3.8.3 describe methods for simulating contacts in these two situations.

If the contact is gradational and the difference between the two populations is primarily a difference in mean, then it might be possible to combine the two phenomena (possibly after a transformation, if more than the mean differs), regarding the transition between them as a local drift. If the differences extend to the second- or higher-order properties of the two phenomena, then a drift model usually is not appropriate. In this case, one can define a sharp contact somewhere within the transition zone and then use conditioning to force the two realizations to converge at this contact. The sharp contact can be imposed a priori so that it is geometrically simple, or it can be estimated or simulated by one of the two procedures described below and conditioned by the available data. If it is estimated or simulated, some criterion (such as a cutoff value) must be selected to decide which available data are in which population. Techniques falling under the general category of "classification", including varieties of cluster analysis and discriminant analysis (Hand, 1981), can be used to allocate



(a) A regionalized variable in R^2



(b) Contour of a regionalized variable in R^3

Figure 54: Geologic contacts that can be modeled as a regionalized variable or as a contour of a regionalized variable. These are cross-sectional views of contacts separating three-dimensional domains.

multivariate data to inferred subpopulations; then the resulting classification of the data can be used to define the contacts. Classification techniques appropriate for analyzing geological data are summarized in Section 3.10.1.

3.8.2 Contacts as Regionalized Variables

If a geologic contact in a three-dimensional domain can be viewed as a function of two "horizontal" spatial coordinates, then the "elevation" of the contact in the third dimension can be viewed as a regionalized variable and can be simulated within a two-dimensional domain. Known contact elevations identified from field or drilling data can be used to estimate the characteristics of the regionalized variable and to condition the simulation. The simulated elevations divide the three-dimensional domain into irregularly shaped upper and lower parts. This operation can be performed successively to divide the three-dimensional domain into several parts, possibly to simulate a sedimentary sequence. Care might have to be taken to insure that successively simulated contacts do not cross (e.g., if some beds "pinch out"). This can be done through dense conditioning or through the use of a transformation. A likely choice of transformations is the "successive-remaining-space" transformation, described at the end of Section 3.5.2.4, which transforms each successive elevation (e.g., from bottom to top) into the proportion of the vertical dimension in the simulation domain that has not been already isolated by the lower contacts. An undertaking similar to this is described by Dowd (1978, 1984).

In some cases, the position of a contact may be statistically correlated with the values observed in the continuous regionalized variables (generally rock analyses) observed on either side of it. This can be taken into account by conditioning in the case of an indicator simulation (described in the next section) but not straightforwardly in an elevation simulation, as the contact elevation is simulated in two

dimensions, and the relevant data are recorded in three dimensions. The third dimension is precisely the one in which the correlation is likely to be observed (as in a gradational contact).

3.8.3 Contacts Defined by Cutoffs and Indicators

3.8.3.1 Overview of Conditional Indicator Simulation

If a contact is too convoluted to be represented as a regionalized variable in two dimensions, we may be able to use the technique of "conditional indicator simulation", which is related to the indicator approach to geostatistical estimation problems described by Journel (1983). To apply the technique in its original form (Isaaks, 1984a, and Journel and Isaaks, 1985), we must begin with two intermingled populations characterized by two sets of continuous data (e.g., chemical analyses). There must also be some direction in the (possibly transformed) variable space in which the two populations do not overlap; i.e., we must be able to distinguish between the populations simply by applying a cutoff to the continuous data. If the data are multivariate, we might apply the cutoff to some combination of the measured variables, such as a linear discriminant function (Hand, 1981, Chapter 4). Continuous data values above the cutoff are assigned indicator values of "1"; values below the cutoff are assigned indicator values of "0". An example is illustrated in Figure 55(a) for a one-dimensional two-population realization. The associated indicator realization is depicted in Figure 55(b). These two realizations are known only at a few discrete sample points, which are subsequently used in model inference and conditioning of the indicator simulation. Conditioning of

an indicator simulation using the original continuous data is discussed in Section 3.8.3.3 and model inference in Section 3.8.3.2. More general methods for obtaining the gaussian conditioning data that do not depend on the original data (or on continuous data at all) are described in Section 3.8.3.4.

The conditional indicator simulation should satisfy at least the same basic criteria that we impose on a conditional simulation of a gaussian process: it must equal the known indicator values at the data points, and it must have the same variogram as the known indicator data. (There is no reason why variograms and other statistics cannot be defined for indicator data.) The conditional indicator simulation is performed by first conditionally simulating a realization from a gaussian process, and then applying some cutoff to these simulated continuous values to generate the simulated indicator values. The locus of points within the simulation domain at which the gaussian realization crosses the cutoff value (i.e., a particular "contour" of the simulated data values) thus becomes the simulated contact between the two continuous populations defined in the original data. Notice that the location of the contact is not simulated precisely but is known to lie between adjacent grid points with differing indicator values.

The big questions are: (1) what variogram should the simulated gaussian realization have, (2) what continuous values should be used to condition it, and (3) what cutoff should be applied to those values? Some partial answers can be given immediately. (1) The variogram of the gaussian realization should be specified in such a way that the indicators generated by applying the cutoff will have the proper

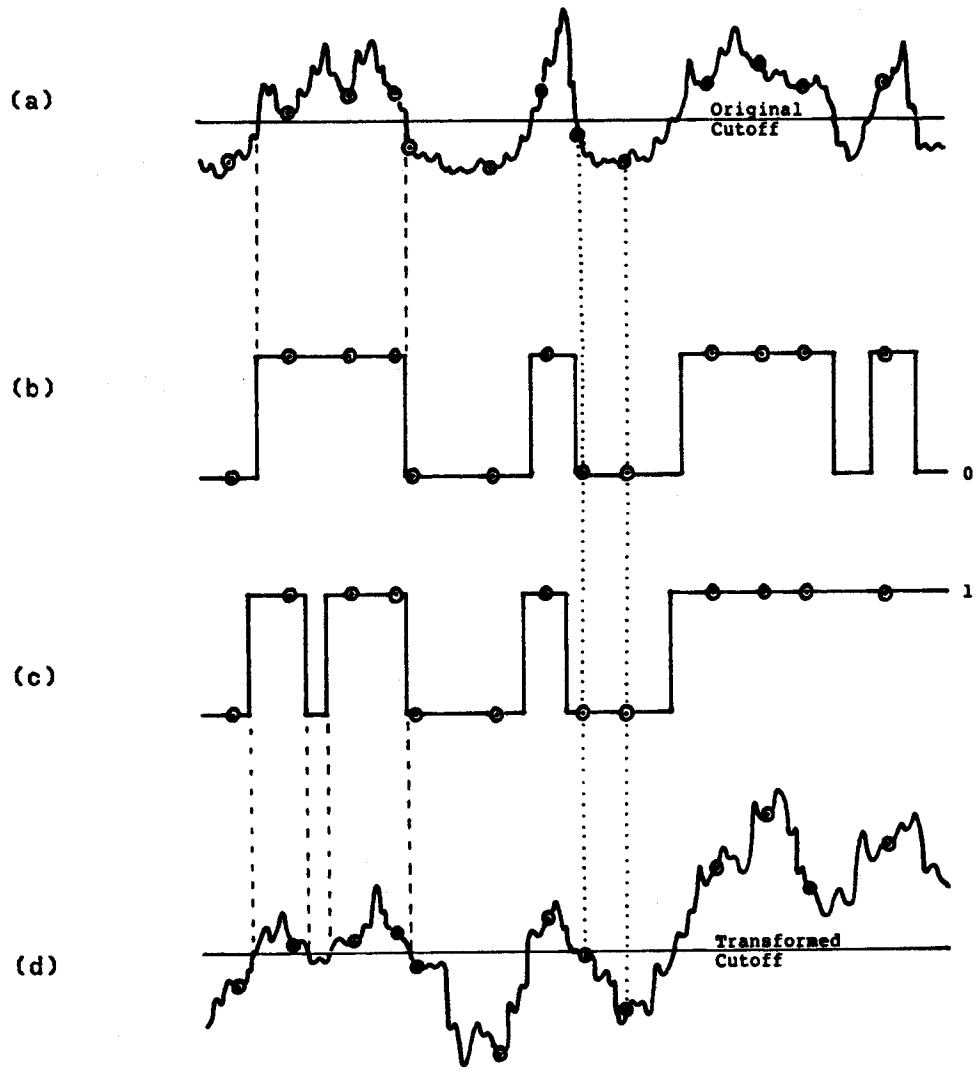
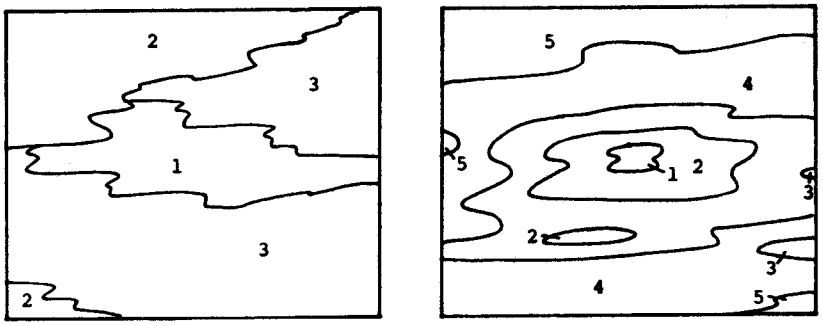


Figure 55: A one-dimensional indicator simulation conditioned by continuous data from two populations. (a) The original mixture of two populations, distinguished by a cutoff value. The true values are known only at a few points (circled). (b) The true indicators, known only at a few points. (c) Conditionally simulated indicators, obtained by applying a cutoff to the realization in (d). (d) Gaussian realization conditioned by transformed data sampled from (a).

indicator variogram. (2) The conditioning values should be continuous data (drawn from a single gaussian realization, to be rigorous) that are above the selected cutoff where the known indicator values are "1" and below the cutoff where they are "0". (3) The cutoff should be selected such that the proportion of gaussian values above the cutoff is roughly the same as the proportion of "1" indicator values in the data. If the data do not cover the simulation domain evenly, a declusterized histogram of the data (Journel, 1983) might be used.

To simulate contacts among several populations this way (Figure 56(a)), we would probably start by simulating the contact separating one population (perhaps a well conditioned one) from all others. Then we would simulate the contact separating another population from all others, possibly incorporating information from the previous simulation into the conditioning data. (This could produce a huge set of conditioning data, so some judicious sampling of the simulated data might be in order.) The contact separating the last population from all others would be fully determined by previous simulations. Notice that simply performing several cutoffs on a single gaussian simulation (Figure 56(b)) could be realistic only if contacts never intersect, as in a layered sequence. Then a simulation by the regionalized-variable approach (Section 3.8.2) usually would be feasible.

All of the varieties of contact simulation described here rely on the use of gaussian simulations, or transformations of them, to represent a complicated boundary between two or more geological populations. In nature, there is ordinarily no reason to suppose that geologic contacts



(a) (b)

Figure 56: Conditional indicator simulations separating more than two populations. (a) The numbered populations can be separated sequentially, e.g. by isolating population "1" first, followed by a separation of "2" and "3". (b) A rare situation in which multiple cutoffs of a single gaussian realization might be used to separate several populations simultaneously.

behave like transformations of gaussian spatial phenomena, so we have here the same serious limitations that were depicted previously in Figure 53. Some contact morphologies just may not be realistically "simulatable" with the techniques to be offered here (e.g., see Figure 57). We must rely on dense conditioning in these cases, or await the development of methods not tied so closely to stationary gaussian phenomena. A general method capable of reproducing complicated shapes like those in Figure 57 would have to be able to reproduce not only the bivariate properties of a phenomenon (e.g., the variogram, which measures squared differences between Z(x) and Z(x+h)), but its multivariate properties as well. Only the univariate and second-order

bivariate properties (mean and covariance function) can be modeled and reproduced by means of a gaussian simulation.

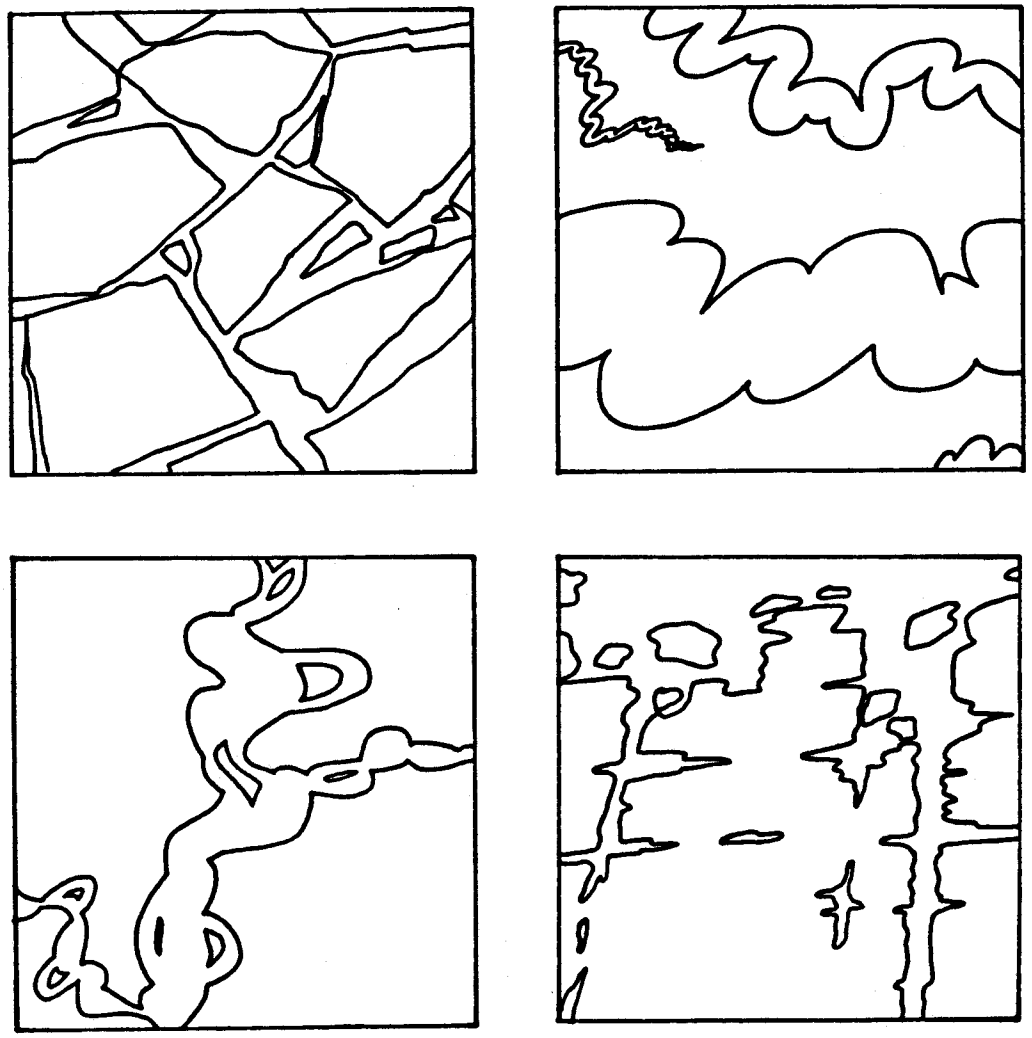


Figure 57: Some examples of geologic contacts that would be difficult to model as contours of stationary gaussian realizations. Clockwise from upper left: breccia, ptygmatic veins, karst surface (cross section), channel fillings (map view).

3.8.3.2 Variograms of Indicator and Gaussian Processes

Regardless of the type of data used for conditioning, a common approach is adopted for modeling the semivariogram function that is used for the unconditional gaussian simulation. If the unconditionally simulated standardized gaussian realization $y_s(x)$ has model distribution function $G(y)$, then the corresponding indicator $i(x)$, equal to "1" when $y_s(x)$ exceeds the cutoff y_c and "0" otherwise, has the following expected moments, derived from the characteristics of Bernoulli (indicator) random variables:

Mean: $E\{I(x)\} = 1-G(y_c) \equiv m_i$

Variance: $Var\{I(x)\} = G(y_c)[1-G(y_c)] = m_i(1-m_i) = C_i(0)$

Semivariogram: $\gamma_i(h) = \frac{1}{2}E\{[I(x+h)-I(x)]^2\}$

Covariance: $C_i(h) = E\{I(x+h)I(x)\} - m^2_i$

Noncentered covariance: $E\{I(x+h)I(x)\} = C_i(h) + m^2_i$

$= C_i(0) - \gamma_i(h) + m^2_i$

$= m_i - \gamma_i(h)$

$= L(y_c, y_c, \rho_{ys}(h))$

where $L(y_c, y_c, \rho_{ys}(h)) \equiv Prob\{ Y_s(x+h) \geq y_c \text{ and } Y_s(x) \geq y_c\}$. The function L is described and graphed in Abramowitz and Stegun (1972, pp. 936-939). Because $Y_s(x)$ is standard normal, its correlogram function, $\rho_{ys}(h)$, is just:

$$\rho_{ys}(h) = E\{Y_s(x)Y_s(x+h)\} = 1 - \gamma_{ys}(h)$$

Estimates of all of these indicator moments can be obtained from the available (possibly declusterized) indicator data, which are either defined from a cutoff on continuous data or based on other types of information, as explained in the next two sections. These statistics

are used to select parameters for the unconditional gaussian simulation. From the sample mean m^*_i and semivariogram $\gamma^*_i(h)$, we can obtain the noncentered covariance:

$$L^*(y_c, y_c, \rho_{ys}(h)) = m^*_i - \gamma^*_i(h)$$

The function L^* can be inverted by reference to Abramowitz and Stegun (1972) to obtain a value for $\rho^*_{ys}(h)$. Then a discrete estimate of the appropriate gaussian-process semivariogram $\gamma_{ys}(h)$ can be obtained by:

$$\gamma^*_{ys}(h) = 1 - \rho^*_{ys}(h)$$

We simply fit a continuous model $\gamma_{ys}(h)$ to this semivariogram, perform the unconditional gaussian simulation, and condition it (using $\gamma_{ys}(h)$) either to the gaussian-transformed continuous data (using the method of Isaaks, 1984a, and Journel and Isaaks, 1985, presented in Section 3.8.3.3) or to data generated directly from the indicator data set (using one of the methods from Section 3.8.3.4). The fitting of $\gamma_{ys}(h)$ at small $|h|$ should be done with care. If the populations have well defined contacts, a nugget effect probably should be avoided so the resulting population boundaries do not look overly fuzzy. If the contact should look very smooth, we should fit a model with parabolic behavior near the origin.

3.8.3.3 Conditioning With Continuous Data

Clearly one (not necessarily the best) choice for gaussian conditioning data would be the normal scores of the original data (or of the discriminant function used to define the cutoff, if the original data are multivariate), using both populations combined. The corresponding choice for the cutoff would then be the normal score of the cutoff that

was applied to the original data (or to the discriminant function). But what would we assume in this case? We would assume that the positions of the contacts and the continuous data values that occur in the two populations are statistically dependent. We would further assume that the pooled gaussian-transformed conditioning data "look like" data from a single realization of a gaussian process that crosses the cutoffs in the appropriate places. In such a realization, continuous values above cutoff should tend to decrease toward the contact and the continuous values below cutoff should tend to increase toward the contact, creating a gradational contact between the gaussian-transformed populations.

If the transformed continuous data look exactly like a gaussian realization, then the two populations appearing in the original data must have been derived from a single parent gaussian realization by the application of an invertible graphical transformation with a shape such as the one in Figure 58. (In the case of multivariate data, Figure 58 might represent a transformation of the discriminant function.) Now if this were strictly true, we could simply perform a normal-scores transformation on the original data (or on a transformation of the multivariate data), invoke the multigaussian hypothesis of Section 3.5.1, conditionally simulate a new gaussian realization using the variogram of the gaussian data to model the unconditional simulation and using the transformed data for conditioning, apply the inverse transformation, and end up with a simulated two-population data set having essentially the same entire spatial distribution as the observed one. If we applied a cutoff to this simulation, the resulting indicators necessarily would have the same variogram that we would find

for the indicators of the original data, because the entire spatial distributions of the gaussian-transformed data and of the gaussian simulation would be the same. The indicators would also be conditioned properly. Thus both requirements placed on the conditional indicator simulation would be satisfied.

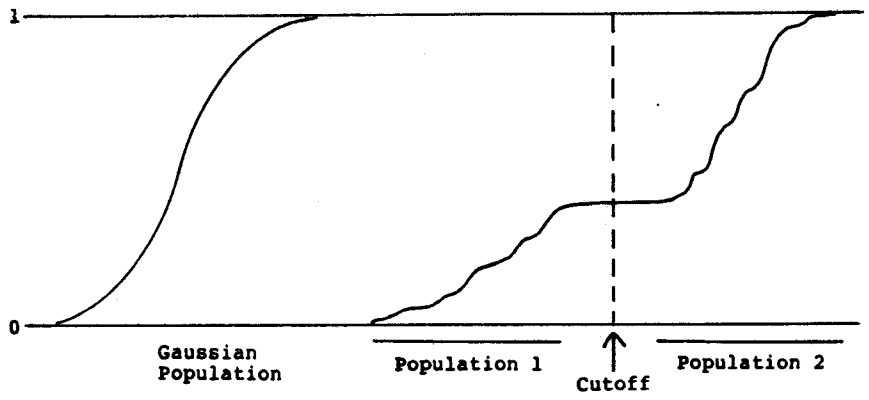


Figure 58: Transformation of a gaussian distribution into a bimodal ("two-population") distribution.

The supposition that the two populations were derived from one parent gaussian realization by a transformation is not generally geologically reasonable, although specific instances (perhaps including selective mineralization of a "gaussian" country rock) might be found. Nevertheless, the original formulation of conditional indicator simulation proposed by Isaaks (1984a) and Journel and Isaaks (1984) relies on the same conditioning procedure as the one above, differing

only in the derivation of the semivariogram model for the unconditional gaussian simulation. (This derivation is described in Section 3.8.3.2.) The normal scores of the original data were used for conditioning in their case study largely for convenience; however, the single-parent hypothesis for their case study (a sandstone uranium deposit) might be considered geologically feasible, and the ore grades in the two populations do tend to converge near the contacts, as would be expected in a phenomenon monotonically transformed from a gaussian realization. These characteristics, along with a good set of conditioning data, made their indicator simulation look very realistic.

If we use gaussian-transformed continuous data for conditioning, we will tend to force the larger-scale variogram structure of these data onto the variogram of the conditional gaussian simulation. This is one of the primary "advantages" of conditioning in most contexts, but it may not be such an advantage here. If the continuous data really represent (or at least resemble[80]) a transformation of a single parent gaussian realization into two artificial "populations" as in Figure 58, then the variogram of the gaussian-transformed data and the variogram used in the unconditional simulation (Section 3.8.3.2) should be the same. Otherwise they may not be very similar at all -- particularly if the two populations represent essentially unrelated geologic phenomena, such as igneous intrusions in a much older country rock. In that case, we would

---

[80] The data should at least be gradational at the contacts -- i.e., values above the cutoff should tend to decrease toward the contacts, and values below the cutoff should tend to increase toward the contacts. Satisfactory checks on binormality of the h scattergrams of the data would add further comfort. Another check is simply the equivalence of the semivariogram of the gaussian-transformed data and the model semivariogram $\gamma_{ys}(h)$ of the unconditional simulation, derived by the methods of Section 3.8.3.2.

be computing a variogram over a mixed population of apples and oranges, with unpredictable results. In this type of geological situation, a cutoff applied to the conditional simulation $y_{sc}(x)$ would yield an indicator simulation that would be conditional to the real indicator data. However, reproduction of the indicator variogram would depend on how closely the kriged values resembled a kriging of a single gaussian realization. An example has been cooked up in Figure 59, which is analogous to Figure 55(a), except that the contacts between the two populations have no tendency to be gradational -- in fact, just the opposite. (It may be comforting to realize that geological examples of gradational behavior near contacts are legion, whereas real-world examples of the "reverse-gradational" behavior in Figure 59 are imaginable but hardly common.) The behavior of the continuous conditional simulation $y_{sc}(x)$ obtained by kriging with the gaussian transforms of the data in Figure 59 would probably be rather different from that obtained using the data from Figure 55, and thus the indicator simulations might look rather different, although the sample indicator variograms in both situations are the same. Notice that the fundamental deficiency of the data set in Figure 59 is that it does not look like a monotonic transformation of a single gaussian realization, but more like a transformation of two or more realizations pieced together. This problem will appear also in the next section, where we attempt to generate gaussian data from indicators.

Figure 59: A mixture of two one-dimensional processes with nongradational behavior near their contacts. Compare this figure with Figure 55(a).

3.8.3.4 Conditioning With Indicator Data

In many practical situations, there is no reason to suppose that the original continuous data are appropriate for use in conditioning the indicator simulation. The use of these data severely restricts the possible outcomes that the conditional indicator simulation can have, and if indicator simulations are being performed to get an idea of the uncertainty that exists in the locations of the contacts, this type of conditioning may lead to both a bias in the assemblage of outcomes and an unrealistically narrow diversity of outcomes, as explained below. Moreover, there are many situations in which we may define populations (and thus indicator data) in ways other than by application of a cutoff to a continuous data set pooled from both populations.

For instance, a geologist may have logged the occurrence of granite versus wall rock in a set of drill cores that were subsequently analyzed for some metal. In our simulation of the metal deposit, we probably would want to treat the granite and wall-rock mineralizations

separately. If the granite intrusives are numerous and complex in shape (e.g., pegmatite bodies), an indicator simulation of granite versus wall rock might be considered, but there are only indicator data -- "1" for granite, "0" for wall rock -- available for conditioning. (The metal contents of the two populations might actually overlap, so the cutoff approach would be useless.) How do we generate gaussian data to condition the unconditional gaussian simulation?

Suppose we have discretized the simulation domain into a grid of points and we assign appropriate indicator values i(x) to the points x that roughly represent the locations of the logged drill holes. We end up with N indicator data, M of which are granite (indicator "1"). The immediate objective is to simulate an unconditional gaussian realization $y_s(x)$ with a specified semivariogram (modeled from the sample indicator semivariogram), and then condition it to a set of gaussian data. The logical cutoff y_c to perform on these data is the [M/(N+1)]th quantile of the gaussian distribution G(y), assuming that the domain has been sampled in an unclustered manner. We must generate (i.e., simulate) a set of continuous conditioning data $\{y(x_j), j = 1 \text{ to } N\}$ that appear to have been drawn from a gaussian realization y(x) with the same mean and variogram as the unconditional simulation. These data must also satisfy the conditions that:

$$y(x_j) > y_c \quad \text{if} \quad i(x_j) = 1$$

$$y(x_j) \leq y_c \quad \text{if} \quad i(x_j) = 0$$

Here we do not assume any continuous values or rankings for the continuous conditioning data -- only conformance to the inequalities imposed by the indicator values. These <u>inequality</u> conditions are <u>much</u>

less restrictive than the exact conditioning values imposed on the conditional gaussian simulations described in earlier sections. As the conditioning is less restrictive, the diversity of possible outcomes is greater.

Now how can we generate these data? We must transform indicators into continuous data -- a "one-to-many" transformation that can only be accomplished by some sort of simulation. The next few paragraphs[81] list several plausible approaches (numbered for convenience):

(1) The brute-force way would be by acceptance-rejection, in which we use some method to generate repeated unconditional multinormal simulations at the N data points (e.g., one of the matrix methods in Section 3.3.1.1) and accept the first simulation that just happens to verify all of the inequalities imposed by the indicator data. In the case of a large number N of indicator data, we could find ourselves generating and checking unconditional simulations for a very long time, so this is not generally an efficient approach.

(2) A better (but still very heavy) approach is a stepwise method similar to the nested conditional matrix-simulation method of Section 3.3.1.1. We start by generating a value $y(x_1)$ from the conditional distribution of

$$Y(x_1)|i(x_1),i(x_2),\ldots,i(x_n)$$

then generate $y(x_2)$ from the conditional distribution of

$$Y(x_2)|y(x_1),i(x_2),i(x_3),\ldots,i(x_n)$$

and so on through $y(x_n)$:

---

[81] Some readers may prefer to skip the rest of this section, which proposes several alternative methods that have not been tested in practical situations.

$$Y(x_n)|y(x_1),y(x_2),\ldots,y(x_{n-1}),i(x_n)$$

In practice, each $y(x_j)$ would be conditioned only on the $i(x)$ and previously generated $y(x)$ values that are no farther away from $y(x_j)$ than the ranges of the associated variograms.

This may seem confusing, so consider the trivial case of only two indicator data, $i(x_1)=1$, and $i(x_2)=0$. We must generate $y(x_1)$ and $y(x_2)$ from a doubly truncated bivariate standard normal distribution with means [0,0], variances [1,1] and covariance (equal to the correlation coefficient in this case) $\rho(|x_1-x_2|) = 1-\gamma(|x_1-x_2|)$, as shown in Figure 60. We know that $y(x_1)\rangle y_c$, because $i(x_1)=1$, and that $y(x_2)\langle y_c$, because $i(x_2)=0$. These constraints define a region within the domain of the unconditional bivariate normal density $g(y_1,y_2)$ -- the lower-right part in this case -- from which we are permitted to draw values of $y(x_1)$ and $y(x_2)$. To generate $y(x_1)$, we must somehow integrate the density $g(y_1,y_2)$ within this region over all allowed values of y_2, and then progressively over all values of y_1 above y_c, to produce a conditional marginal distribution function:

$$F(y(x_1)|i(x_1),i(x_2)) = \frac{\int_{y_c}^{y(x_1)} \int_{-\infty}^{y_c} g(y_1,y_2)\, dy_2\, dy_1}{\int_{y_c}^{\infty} \int_{-\infty}^{y_c} g(y_1,y_2)\, dy_2\, dy_1}$$

Then we can just generate a random number u_1 distributed uniformly between 0 and 1, and obtain an appropriate value for $y(x_1)$ from:

$$y(x_1) = F^{-1}_{y(x1)/i(x1),i(x2)}(u_1)$$

This reduces the domain of $g(y_1,y_2)$ from which we can generate a value for $y(x_2)$ to the heavy vertical half-line shown within the lower-right

part of Figure 60. Knowing the value of $y(x_1)$, we generate $y(x_2)$ from a truncated version of the conditional density $g(y_2|y_1=y(x_1))$ using the distribution

$$F(y(x_2)|y(x_1),i(x_2)) = \frac{\int_{-\infty}^{y(x_2)} g(y_2|y_1=y(x_1))\, dy_2}{\int_{-\infty}^{y_c} g(y_2|y_1=y(x_1))\, dy_2}$$

so that $y(x_2)$ is obtained from

$$y(x_2) = F^{-1}_{y(x2)/y(x1),i(x2)}(u_2)$$

where u_2 is distributed uniformly on [0,1], independently of u_1.

It is not obvious how this "better" approach, already complicated enough with only two data, could be programmed to generate several thousand gaussian conditioning values -- a realistic number for real-world applications. Thus it appears that we may have to resort to some kind of approximation. A few possibilities are described below.

(3) Suppose that we ignore the conditioning effect of <u>all</u> known indicator values (or of those no farther away than the range of the indicator variogram) on the simulation of each $y(x_j)$. Instead we condition each simulated $y(x_j)$ only on the indicator $i(x_j)$ and on the previously generated $y(x)$ values lying within the range of the $y(x)$ variogram. Then we could use the stepwise-conditional approach of Section 3.3.1.1, augmented by only one acceptance-rejection step to account for the single indicator. We start by generating a first trial value of $y(x_1)$ from an unconstrained standard normal distribution. If the value happens to land on the correct side of the cutoff, we accept it. Otherwise, we keep generating until we get what we want. Thus we have effectively generated $y(x_1)$ from the distribution of:

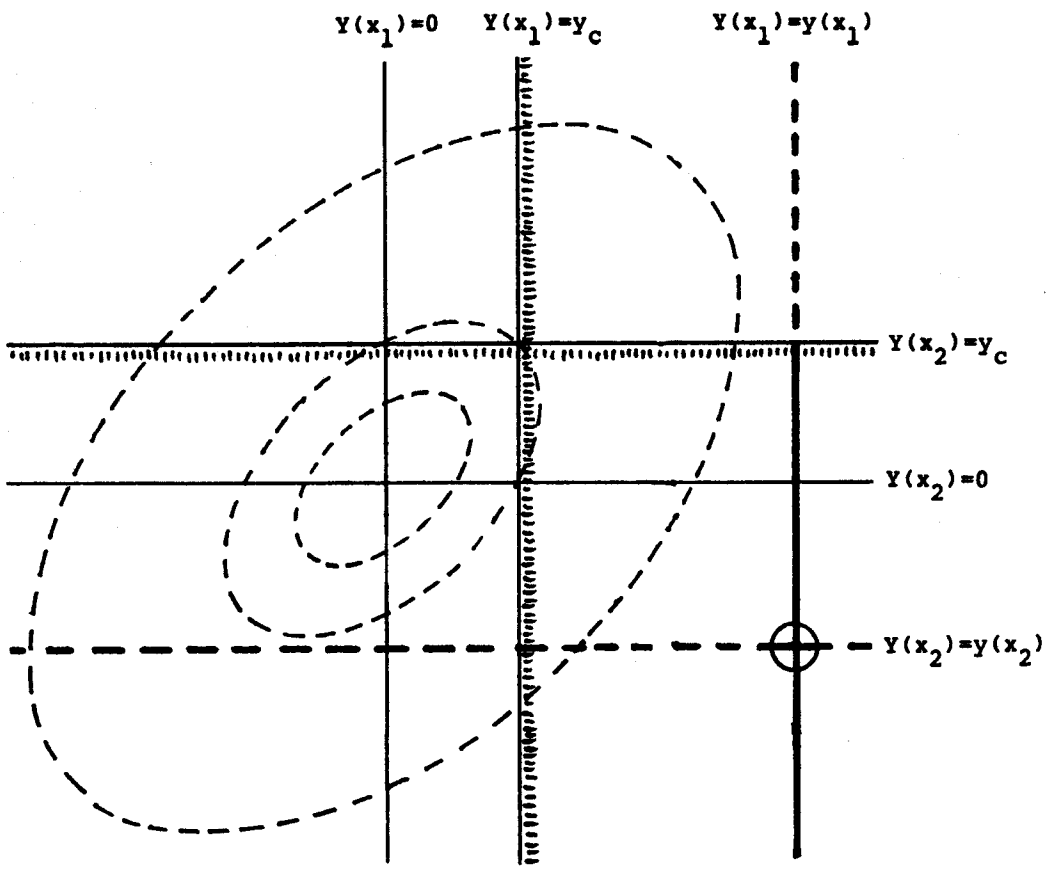


Figure 60: Generation of two bivariate gaussian data from a distribution conditioned by indicator data. We know that $i(x_1)=1$, so $y(x_1)>y_c$. We know that $i(x_2)=0$, so $y(x_2)\leq y_c$. Thus we must generate two values $y(x_1)$ and $y(x_2)$ randomly from the lower-right (doubly shaded) part of the unconstrained bivariate density $g(y_1,y_2)$.

$$Y(x_1)|i(x_1)$$

(Compare this with the distribution for $Y(x_1)$ on page 289.) Then we generate $y(x_2)$ from the conditional distribution of

$$Y(x_2)|y(x_1),i(x_2)$$

using the same acceptance-rejection procedure to account for the indicator $i(x_2)$. We continue until we get an acceptable value for $y(x_n)$ from the distribution of:

$$Y(x_n)|y(x_1),y(x_2),\ldots,y(x_{n-1}),i(x_n)$$

Viewed geostatistically, each successive $y(x_j)$ would be sampled by acceptance-rejection from a gaussian distribution with mean equal to the simple-kriging estimate (Journel and Huijbregts, 1978, p. 561 and p. 566) of $y(x_j)$ given the previously generated $y(x)$ values, and variance equal to the simple-kriging variance. Although this method is much simpler than method (2), the matrix calculations required to obtain the last several data are still very heavy. (However, if conditioning is limited to previously generated $y(x)$ values within the range of the y variogram, the matrices may not be so large in practice.)

The resulting set of conditioning data $\{y(x_j), j = 1 \text{ to } N\}$ would all satisfy their individual indicator constraints, but they would have been drawn from the wrong conditional distributions. In the bivariate case illustrated by Figure 60, this procedure amounts to drawing $y(x_1)$ from the marginal density obtained by integrating the whole right (shaded) side of the density $g(y_1,y_2)$ over y_2, instead of just the portion of the density lying in the lower-right (doubly shaded) sector. Once $y(x_1)$ is drawn in this biased way, $y(x_2)$ is obtained in the usual manner from its conditional distribution given $y(x_1)$ and $i(x_2)$. Although this method is not strictly correct, the values generated by the method might, when viewed as a group, look reasonable enough to be used as gaussian conditioning data. The earlier values, generated with less conditioning, would tend to be less realistic than the later ones. For example, early values generated near contacts between "0" and "1"

indicators might not tend to be close to the cutoff, as we would expect in a single gaussian realization; and values generated within clusters of "1" or "0" indicators might not tend to be "high" or "low", respectively (Figure 61). (This problem is similar to the one illustrated in Figure 59.) Therefore, the order in which the $y(x_j)$'s are generated probably should be randomized. To correct likely deficiencies in the earlier $y(x_j)$'s, we might perform a second pass over some of these values, obtaining new simulated values by conditioning on the later values obtained in the first pass.

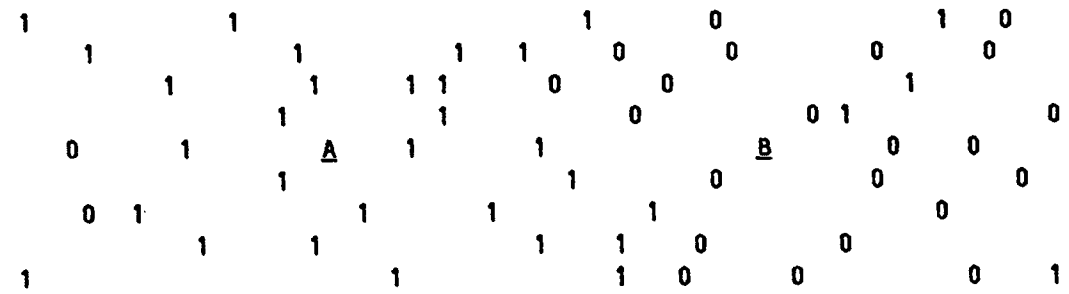


Figure 61: Map of a set of hypothetical indicator data to be used for conditioning a gaussian simulation. Gaussian conditioning values generated at "A" should tend to be higher than those generated at "B", even if both A and B have the same indicator value. The influence of the surrounding indicator values is explicitly considered in methods (1), (2), and (5) of this section, but not in methods (3) and (4).

A faster but even less satisfactory approach would be to generate an unconditional simulation at all conditioning points x_j and simultaneously accept all $y(x_j)$ values satisfying their own indicator

values. Then we could generate a new set of $y(x_j)$ values to replace the rejected values, conditional on the values accepted in the first pass, again accepting those satisfying their own indicators. All data could be simulated relatively quickly this way, but probably big local chunks of inappropriate-looking unconditional data would all be accepted together in the first pass.

Yet another alternative would be to replace the successive simple-kriging estimates of each $y(x_j)$ with a cokriging, taking into account all surrounding indicator data. Method (5), below, achieves a similar effect in a somewhat simpler way.

(4) Isaaks (1984b) has described another approach that generates conditioning data with a univariate gaussian distribution and the proper indicator variogram. However, like method (3), this method generates data that are not sampled from a single gaussian realization and thus will not generally have the correct variogram for conditioning the gaussian simulation. We begin by generating <u>two independent</u> unconditional gaussian realizations using the semivariogram model derived from the indicator semivariogram. The first realization, $y_{s1}(x)$, is generated at all points in the simulation domain and is used as the unconditional gaussian simulation. The second realization, $y_{s2}(x)$, needs to be generated only at conditioning points, but in practice it is probably easier to generate it everywhere in the domain. Each member of the data set $\{y_{s2}(x_j),\ j = 1 \text{ to } N\}$ is then transformed into one of two ranks, k_{0j} or k_{1j}, depending on the value of the associated indicator $i(x_j)$, and finally transformed into

$$y_{s3}(x_j) = G^{-1}\{k_{0j}G(y_c)/[(N+1)(1-m_i)]\} \quad \text{if } i(x_j) = 0$$

$$= G^{-1}\{G(y_c)+k_{1j}[1-G(y_c)]/[(N+1)(m_i)]\} \quad \text{if } i(x_j) = 1$$

where k_{0j} = rank of $y_{s2}(x_j)$ among all j for which $i(x_j)=0$

k_{1j} = rank of $y_{s2}(x_j)$ among all j for which $i(x_j)=1$

m_i = arithmetic average of $\{i(x_j), j = 1 \text{ to } N\}$

y_c = $G^{-1}[1-m_i]$

N = total number of indicator data

and G( ) = standard normal distribution function.

Each transformed value $y_{s3}(x_j)$ then represents a quantile of one of the complementary truncated normal distributions in Figure 62.

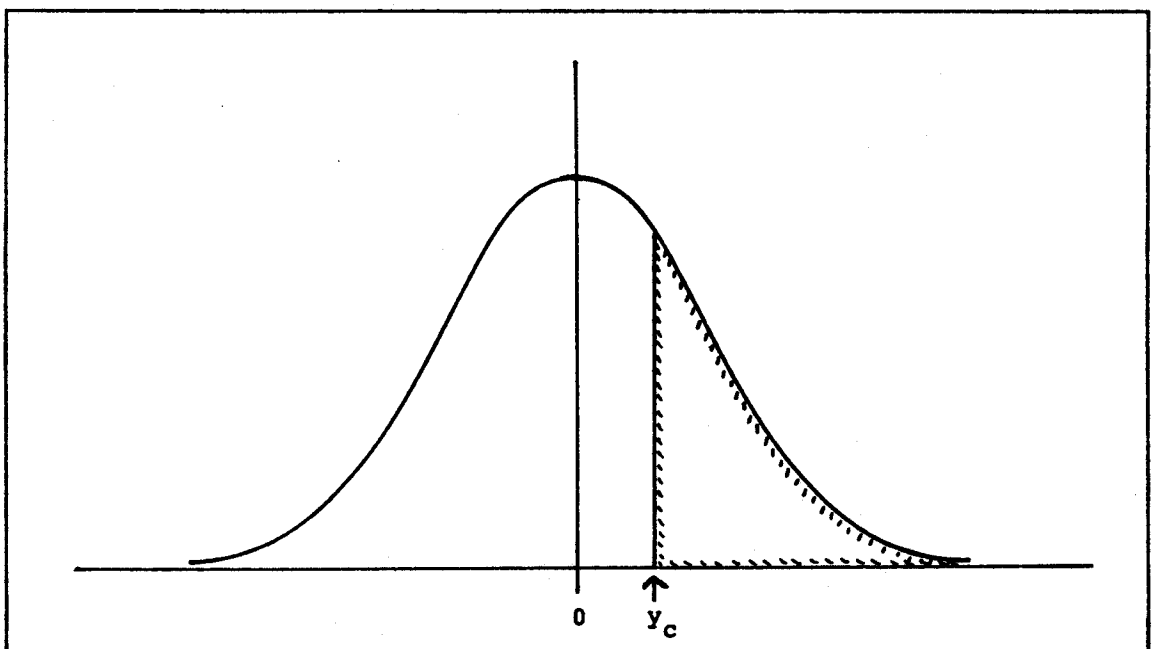


Figure 62: Truncated normal densities used in the transformation of Isaaks (1984b). The value y_c separates the two truncated densities. Simulated values $y_{s2}(x_j)$ are transformed into values from the left density if $i(x_j)=0$, from the right density if $i(x_j)=1$.

The resulting data set $\{y_{s3}(x_j), j = 1 \text{ to } N\}$ is then used to condition the unconditional gaussian simulation, $y_{s1}(x)$, using the

semivariogram model attributed to $y_{s1}(x)$ in the kriging. The transformation above, explained in detail by Isaaks (1984b), is designed to impose univariate normality on the whole set of N data, while assuring that data associated with indicators $i(x_j)=1$ exceed the cutoff y_c and other data are below the cutoff. Although the indicator variogram is strictly reproduced, this method does not assure reproduction of the semivariogram model for $y_{s1}(x)$, as it does not in general produce a realization of a single gaussian process. Because the starting data $\{y_{s2}(x_j)\}$ are taken from a completely unconditional simulation, the values of the data $\{y_{s3}(x_j)\}$ may not tend to rise and fall toward the contacts in a natural manner as would be expected in a single gaussian realization. This method might therefore produce an even less realistic data set for conditioning than the previous ones. However, it is computationally much easier than the other methods, avoiding the heavy matrix manipulations required for the stepwise-conditional simulations.

(5) A final approach, which might be developed into several related variations, attempts to strike a compromise between the rigorous stepwise-conditional approach (2) and the abbreviated stepwise approach (3). We again generate the $y(x_j)$ values by a successive conditional method, at first using only transformed indicator data for conditioning, but gradually replacing these with previously generated $y(x)$ data. We begin by (somehow) transforming the indicators into a univariate gaussian data set $\{w(x_j), j = 1 \text{ to } N\}$, having a standard normal distribution $G(w)$, a tendency to grade toward the contacts like a well behaved single gaussian realization, and a consistency with the

indicator data, so that $w(x_j)>y_c$ if $i(x_j)=1$, and $w(x_j)\leq y_c$ if $i(x_j)=0$. The transformation might be done in several ways, allowing several variations to the approach.

For instance, we might replace each indicator value $i(x_j)$ with a continuous variable $i^*(x_j)$ kriged from the other indicator values using a model indicator semivariogram estimated from the indicator data. Then we could rank the kriged values into two categories (depending on the true indicators $i(x_j)$) and apply the transformation of Isaaks (1984b), which is provided in the description of method (4), to obtain the new data set $\{w(x_j)\}$ with gaussian distribution $G(w)$. The data $w(x_j)$ would then be consistent with the indicator data and usually close to the cutoff value in areas where both "0" and "1" indicators occur. However, because the $i^*(x_j)$'s are generated by a kriging, the $w(x_j)$ values probably would be too smooth to reproduce the desired semivariogram model for the gaussian data, so another step probably would be required to get the conditioning data.[82]

We could perform a stepwise-conditional generation of the conditioning data $y(x_j)$, taking the locations x_j in a random order. The first value, $y(x_1)$, is generated from the conditional distribution of

$$Y(x_1)|i(x_1),\ w(x_2),w(x_3),\ldots,w(x_n)$$

where the $w(x)$ data are treated as though they were previously generated $y(x)$ data, and $i(x_1)$ is satisfied by acceptance-rejection. Then, successively, we generate $y(x_2)$ from

$$Y(x_2)|y(x_1),i(x_2),\ w(x_3),w(x_4),\ldots,w(x_n)$$

---

[82] An example of the ill effects that could arise by not proceeding to the next step is contained in the discussion of a smooth "despiking" method in Section 4.2.4.

and so on to $y(x_n)$ generated from:

$$Y(x_n)|y(x_1),y(x_2),\ldots,y(x_{n-1}),i(x_n)$$

For a variation on this approach, we could replace the kriging of $i^*(x_j)$ with an average of the $i(x_j)$ data within a moving window (a simpler approach, maybe just as good, and probably not so smooth).[83] We could also add some noise to the $i^*(x_j)$ before transforming them to the $w(x_j)$ to roughen them up, or smooth the $i^*(x_j)$ with a moving window before transforming them to the $w(x_j)$, to tone them down. This way we might achieve a semivariogram model for the $w(x_j)$'s that would look very much like the desired model, in which case the arduous stepwise-conditional method might be skipped.

Notice that, of the five approaches above, only methods (1) and (2) are rigorous, but they do not seem practical for indicator data sets of realistic size. If the number of conditioning data to be generated is very large, the approximations proposed in methods (3) and (5) may be impractical as well, because of the matrix sizes involved in the stepwise-conditional generation method. (The matrix size is not such a problem if only a few previously generated data lie within the range of each new value to be generated.) Because many practical situations can be visualized in which a realistic simulation of complicated contacts among different populations would be considered crucial, the general field of indicator simulation has to be considered ripe for research. Better methods for obtaining gaussian conditioning data are needed.

---

[83] Essentially the same approach, allowing nonuniform moving-average weight functions as well as uniform "windows", has been suggested by Journel (1984d).

Methods not dependent on gaussian simulations, and thus able to reproduce more highly parameterized shapes, are needed even more.[84]

3.8.4 Conditioning Continuous Data Within Populations

In theory, all regionalized variables have infinite spatial domains, but in nature this is never the case. Furthermore, natural regionalized phenomena are seldom independent of their natural spatial boundaries. For instance, unless a geologic contact is a razor-sharp fault, the rocks on either side of it usually exhibit some sort of systematic behavior with proximity to the contact. Usually this behavior cannot be strictly reproduced in a conditional simulation simply by truncating the simulation at the contact, unless a very good set of conditioning data is available near the contact.

Several types of behavior near geologic contacts can be imagined. In a completely gradational contact, the statistical and geological characteristics of one phenomenon grade smoothly into those of another. This situation, common in sedimentary facies contacts, is illustrated in Figure 63(a). Igneous intrusive contacts more commonly exhibit systematic behavior that depends only on distance from the contact: the igneous rock may become finer-grained and more contaminated by assimilated wall rock near the contact, and the wall rock may become

---

[84] Davis (1985b) has proposed a matrix method for indicator simulation that is not based upon gaussian simulations, but it is not specifically designed to reproduce the multivariate properties of the data set, and it is subject to the usual size limitations of matrix simulations. Also, methods to generate unconditional simulations of two- and three-dimensional Markov random-set processes with a finite number of states are described at length by Lin and Harbaugh (1984). The properties of the process in two dimensions were derived by Switzer (1965). The method produces mosaic-like realizations with linear or planar boundaries between tiles.

more recrystallized and metasomatized near the igneous body. The compositions of the two rocks may not appear gradational, although they may grade toward each other, as in Figure 63(b). (In this case, the contact may appear completely gradational after a transformation; cf. Figures 55(a) and (d).) If the contact is an erosion surface, any of the situations in Figure 63(b), (c), or (d) might occur.

Contact effects are not the only ways in which the compositions of two adjacent rock types can be cross-correlated. For example, if a shaly limestone is locally dolomitized after deposition, the contacts between the limestone and dolomite facies may be sharp and nongradational, but the amount of shale in the two facies may appear as a continuous regionalization, independent of the dolomitization. Then the calcium contents of the limestone and dolomite would also be cross-correlated, owing to a constant-sum constraint involving the shale, but the calcium regionalizations in the two populations would not be continuous across the contacts.

If there are no geological reasons for suspecting a certain kind of cross correlation or contact behavior, the data should nevertheless be carefully explored to see if such effects may be present. Then three questions arise:

(1) Are the continuous data statistically independent of their positions with respect to the contacts? Such a dependency may appear as a local drift near the contact, which might be modeled as a regression -- viewing the regionalized variable as a response variable with distance-from-contact as a predictor. (This would not be a linear regression, as the effect of the contact could be expected to die out

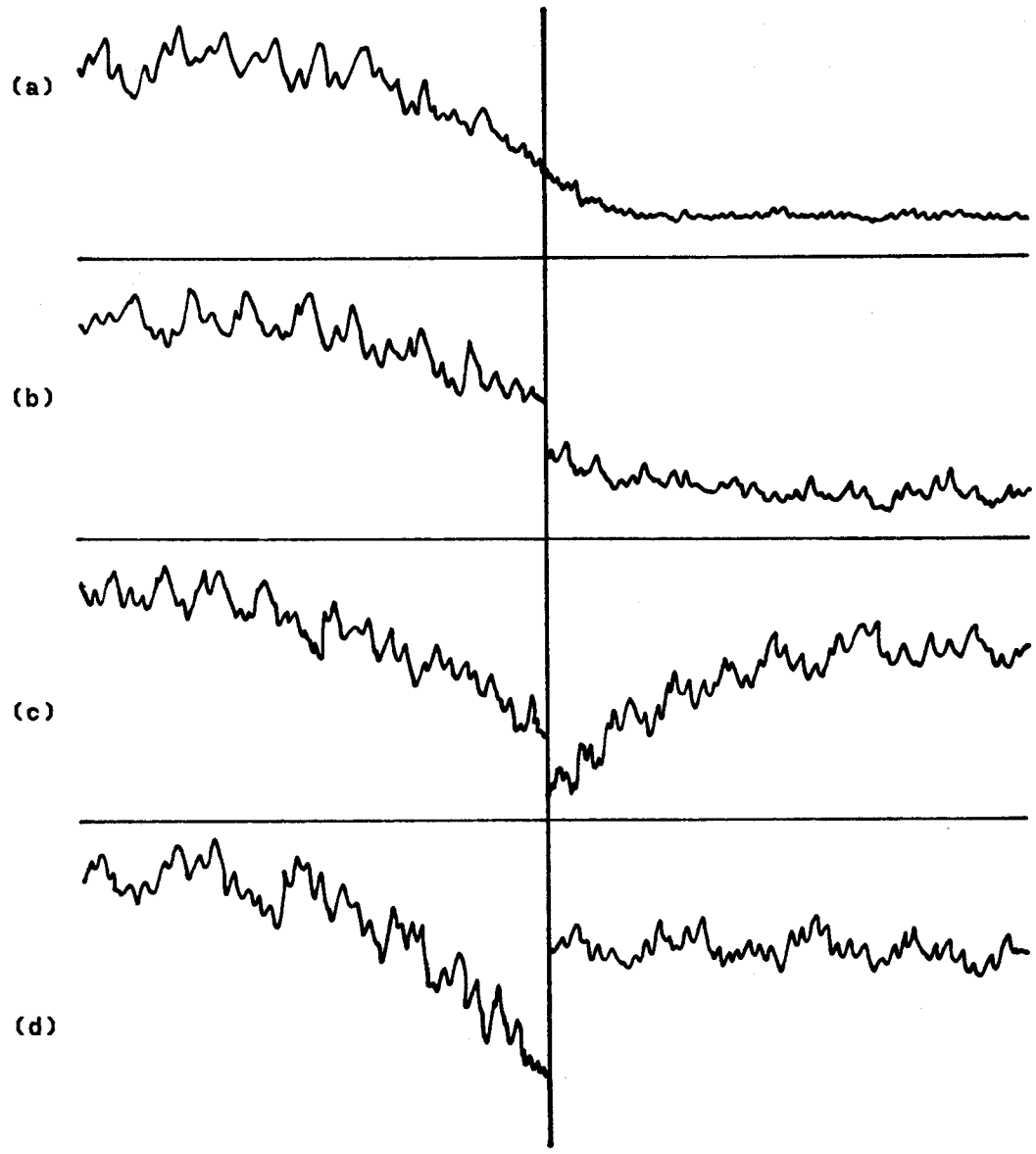


Figure 63: Possible trends in regionalized data near population boundaries. In case (a), two populations with very different statistical properties grade into each other across a contact. In cases (b) and (c), the contact is not gradational, but the two populations nevertheless appear to be locally statistically dependent. In case (d), the two populations are independent, although one shows a relationship to the contact.

gradually away from the contact.) Another possible model would be a coregionalization between the regionalized variable and a rock-type indicator variable. An analysis of the data should reveal which model is more appropriate.

(2) Are the continuous data in each population independent of the data in the other populations? Although data in different populations may appear to be cross-correlated near the contacts, this might be explained entirely by a "distance" or "indicator" effect of type (1). However, populations separated by completely gradational contacts, or populations consisting of "mineralized" and "unmineralized" rock (e.g., the locally dolomitized limestone described above) may be more readily modeled as spatially cross-correlated variables.

(3) In the case of a coregionalization, do all variables in a population behave in the same way with regard to questions (1) and (2)? Although it is possible that not all variables will show obvious dependencies, it is likely that any dependencies that do exist will all be of the same type: either dependence on contact proximity, or spatial cross correlation among populations.

A preliminary analysis of the data to answer the three questions above, as well as a consideration of the geological origins of the populations in question, should provide some clues to how the observed effects could be incorporated into a simulation. Once the population boundaries have been fixed (possibly after a careful consideration of these questions, particularly if indicator simulation is used), simulation of the continuous phenomena inside them can utilize both the available continuous data and the position of the contact (or the

associated indicator data) for conditioning. Furthermore, if one population is simulated before another, the simulated data from one population can be used to condition the simulation of another with which it is correlated, although simultaneous simulations via a combined linear model might be preferable. Some alternative procedures for different situations are described below.

Before we can determine whether a natural regionalized phenomenon is dependent in some way on the proximity of a contact, we must know the contact's location. If the contact is well behaved -- i.e., if it is deterministic or at least predictable within a small estimation error -- then we can easily find the shortest ("vertical") distance between each data point and the contact. A scattergram of values of the regionalized data versus their distances from the contact should reveal any trend that may exist in the local mean. Then the simplest way to model the effect is to fit a smooth curve to the trend (generally a curve whose slope decreases in absolute value with increasing distance from the contact, like a spherical semivariogram curve) and regard the residuals as samples from a stationary regionalized variable. Subsequent semivariogram modeling and conditioning should be performed using these residuals. (The residuals should be inspected carefully, as a trend may exist in the local variances of the data as well as in their means. A further transformation to stabilize the variance may be needed.) In a sedimentary deposit with a pervasive vertical drift that extends through the whole thickness of the deposit, the local drift near the contact may be modeled together with the overall drift. (Methods for simulating vertical drifts are discussed in Section 3.10.4.)

If the contacts appear to be convoluted, or if their true locations are very poorly known, any effect on a regionalized variable that results from proximity to contacts will have to be modeled using only the known indicator data. Some type of contact effect involving the boundary between two populations should be suspected if there is any structure in the sample noncentered cross covariance,

$$NCC^*_{zi}(h) = [1/N(h)] \sum_{j=1}^{N(h)} [z(x_j)i(x_j+h)]$$

between the continuous data $z(x_j)$ within <u>one</u> population and the indicator data $i(x_j)$ of <u>both</u> populations. (Notation is analogous to that in Section 3.1.3.) Notice that the sample cross semivariogram,

$$\gamma^*_{zi}(h) = [1/(2N(h))] \sum_{j=1}^{N(h)} \{[z(x_j)-z(x_j+h)][i(x_j)-i(x_j+h)]\}$$

would be meaningless in such a case, because the first difference $[z(x)-z(x+h)]$, for $z(x)$ defined in a given population, exists only where the second difference $[i(x)-i(x+h)]$ is zero. The centered covariance,

$$C^*_{zi}(h) = NCC^*_{zi}(h) - m^*_z m^*_i$$

is of questionable value as well, because the mean m^*_z of the $z(x)$ data applies to only part of the domain, and the mean m^*_i of the indicator data is probably nonstationary (certainly nonstationary if the domain is small). Fortunately, a noncentered covariance can be used in a cokriging system, so the procedure for modeling a contact effect using indicator data is to simulate a stationary $z_s(x)$ within the domain of interest and use cokriging with the known indicator data (or even better, with previously simulated indicator data known everywhere) to condition the simulation. Although more difficult to visualize than the

local-drift model of the previous paragraph, an indicator approach to contact effects might be more realistic in complicated situations where rocks from one population are surrounded by several masses of material from another population. In such a case, a reliance on the distance to the nearest contact (if known) as a predictor would tend to ignore the likely effects of other nearby contacts. Cokriging with indicator data would allow the proximity of all nearby materials to be taken into account.

To check whether the continuous data of two populations are cross-correlated, we must again calculate sample noncentered cross covariances instead of variograms, because two variables $z_1(x)$ and $z_2(x)$ defined on disjoint spatial domains have no cross variogram. If the cross covariances show some structure, two approaches are possible: (1) simulation of each variable independently, using cokriging to condition each simulation, or preferably (2) simulation of both variables simultaneously as a coregionalization, <u>again</u> using cokriging for conditioning. Cokriging is necessary in both cases, because the data coverage for each variable contains large gaps corresponding to the disjoint domain of the other variable. Hence data near the contacts would not be properly conditioned using only ordinary kriging. If the variables are simulated independently, conditioning might be done sequentially, using the conditionally simulated data from one population to condition the other.

If two populations are believed to be completely gradational, and the difference between them is primarily a difference in the mean, the contact region can be simulated in either of two ways: (1) as a local

drift, or (2) as a single stationary phenomenon, conditioned by data from both populations. The first approach probably requires fewer conditioning data. The easiest way to use method (1) is to estimate the drift from local averages, by universal kriging, or (suboptimally[85]) by trend-surface techniques (Davis, 1973, pp. 322-358). If residuals from the drift are stationary, the drift can simply be subtracted out; then the simulation can proceed using the residuals instead of the original data.[86] If the variogram of the residuals is not stationary, the next paragraph may offer a way out. To use method (2) we should break the data set into two stationary populations, leaving out the transition zone, and estimate the variogram separately within each population. The method should work well if the inferred semivariogram models are the same, or close enough to be pooled together with little loss of accuracy. We then perform a single unconditional simulation over the combined domain of both populations using the pooled semivariogram model. The transitional behavior is imposed by conditioning.

If there is a difference in the variograms of the data in two gradational populations, an approximate solution, which models the transition zone as a weighted average of the two regionalizations, may still be available. An example is provided in Figure 64. First we should remove any trend in the mean of the two populations by method (1) above. Then we calculate semivariogram models for three disjoint domains: the "left" population (in Figure 64), the transition zone, and

---

[85] The trend-surface fitting technique assumes that residuals from the fitted surface are uncorrelated. This is not usually the case.

[86] However, the first case study of Chapter 4 (particularly Section 4.1.3.5) illustrates a difficulty that can develop when using normalizing transformations in the presence of drifts.

the "right" population. If the semivariogram within the transition zone (which presumably is not stationary) appears like a mixture of the left and right semivariograms, we can proceed with the simulation.

First we simulate unconditional realizations from the semivariogram models of each population. Each simulation domain should cover the domain of the corresponding population, plus the domain of the transition zone, so the two realizations overlap. Now we find a linear combination of the left and right semivariograms that looks reasonably like the calculated sample semivariogram of the transition zone. This model should be roughly appropriate for the central part of the transition zone (the "contact", if we had to pick a sharp one). We then construct a single unconditional simulation of the whole phenomenon as follows. At the center of the transition zone, we average the two unconditional simulations according to a linear model derived from the coefficients of the combined semivariograms, determined above. At the contact between the stationary left population and the transition zone, we apply a coefficient of "1" to the simulation of the left population, "0" to the simulation of the right population. At the contact between the right population and the transition zone, these coefficients are reversed. At all other points within the transition zone, linear combinations interpolated between the central combination and the two endpoints are used, as shown in Figure 64. Finally, the simulation is conditioned using ordinary kriging within the two "pure" populations, using the respective semivariograms of those populations, and including a few bordering data from the transition zone. Within the transition zone, kriging can be performed using the nested model derived earlier,

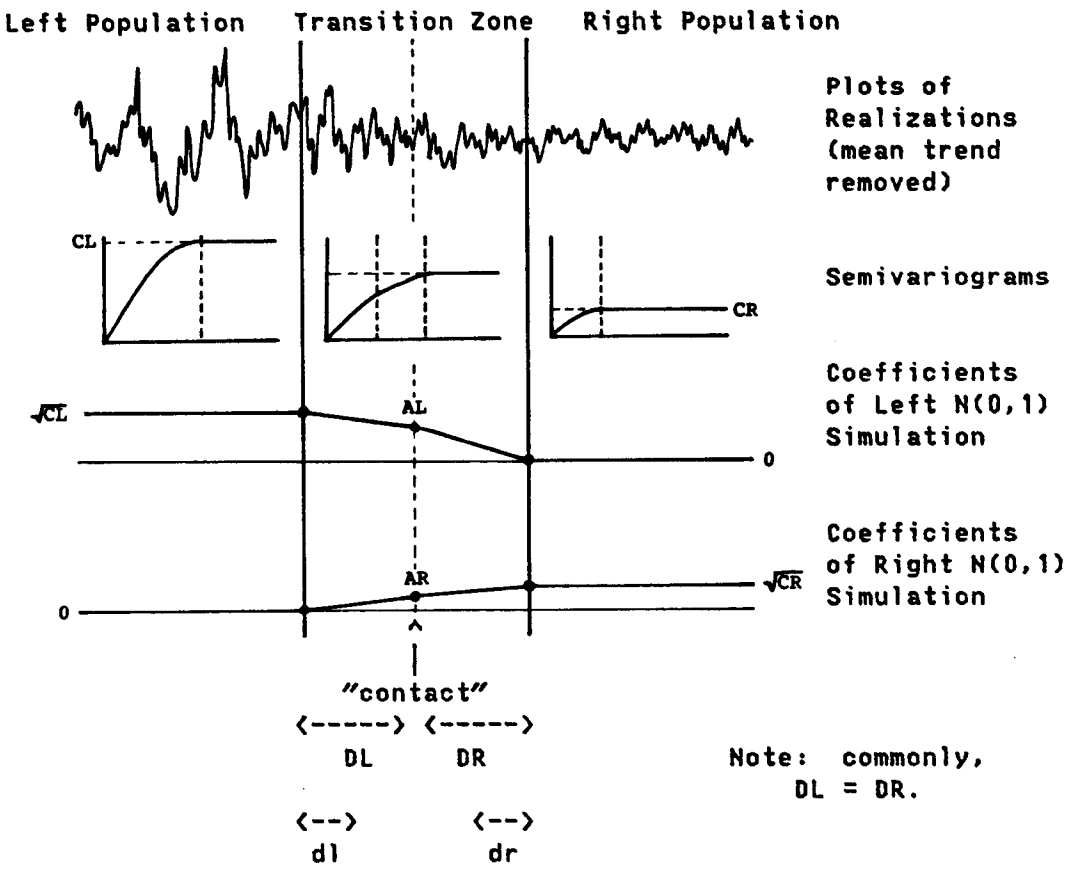


Figure 64: Unconditional simulation of a transition zone between two populations with different variograms. The left and right unconditional simulations (standardized to sills of 1) are combined within the transition zone according to these formulas:

At "contact": coefficient of left simulation = AL
coefficient of right simulation = AR

Left of contact: coefficient of left simulation
$= [dl/DL]AL + [(DL-dl)/DL]\sqrt{(CL)}$

coefficient of right simulation
$= [dl/DL]AR$

Right of contact: coefficient of left simulation
$= [dr/DR]AL$

coefficient of right simulation
$= [dr/DR]AR + [(DR-dr)/DR]\sqrt{(CR)}$

or (with more difficulty) using a model that progressively changes across the zone.

The method can be simplified at the unconditional-simulation stage by making a direct interpolation between the simulations of the two populations instead of working through the model derived from the transition zone. This may be necessary if there are too few data in the transition zone to provide a good sample variogram. Otherwise, the use of the central point should provide a better simulation of the transitional phenomenon. This method could be extended to unconditional simulations of phenomena that have only locally stationary ("quasi-stationary") variograms. We would first break up the simulation domain into quasi-stationary regions and perform an unconditional simulation within each region, overlapping large areas of adjacent regions. We would then simulate the transitional areas between regions in the manner described above.

3.9 APPLICATIONS OF CONDITIONAL SIMULATION IN MINING AND MINERAL PROCESSING

> We can be thankful that the pure beauty of mathematics is part of our life experience, but we must remember that our proper place is on the mainland of "doing." [S. C. Florman, The Existential Pleasures of Engineering, Chapter 11]

3.9.1 Mining Applications

Since its development in the early 1970's, conditional simulation has been applied most frequently to the design of mining systems. The general approach has already been described in the introductory paragraphs of Section 3.1.7. Briefly, the idea is to try out various

simulated mining procedures on a simulated mineral deposit; the procedures that perform favorably are then considered for installation in the actual mine. The closer the simulated mineral deposit and simulated mining systems come to the actual situation that would be encountered in the real mine, the more confidence we can place in the simulated results. If either the conditional simulation of the deposit or the mining simulation is unrealistic, the results will be of little value. Interestingly, the recent published literature on applied conditional simulation in mining seems to dwell more on the development of realistic mining models than on the creation of more realistic conditional simulations of mineral deposits. The emphasis in this dissertation is just the opposite: mining simulations are covered only briefly in this section, and only very simple simulations are attempted in the case study of Section 4.2.

Many extended examples of conditional simulations applied to mining problems have appeared in the literature. Very brief guides to the early literature are presented by Journel(1980, pp. 98-108) and by Journel and Huijbregts (1978, pp. 545-554); a more lengthy description of an important early application by Deraisme (1977), similar in some respects to the study presented in Section 4.2, is contained in each of those reviews. Several more recent applications are described by Deraisme and Dumay (1981), Deraisme and Marbeau (1983), and Deraisme and de Fouquet (1984). Mining-simulation programs commonly are interactive, utilizing graphical displays to enable a human operator to control the progress of the simulated mining operation. Of course, human interventions render a mining simulation irreproducible and in a sense

less easily comparable with simulations of other alternatives, because the effects of human and preprogrammed factors are confounded. However, such simulations are usually more realistic, inasmuch as real mining operations are never fully automated. Furthermore, some of the important factors that must be considered during the development of a real mine (drainage, working room, noise, traffic flow, etc.) may not be easily reduced to a set of programmable instructions.

Mining simulations may be designed to answer a variety of practical questions about the interaction between mining procedures and the spatial distribution of ore grades within a mineral deposit. A common problem, addressed in the case study of Section 4.2, is the selection of a mining procedure that will somehow minimize the variability of ore grades delivered to a processing plant, while satisfying constraints on product quality. If the compositional variability of the delivered ore can be cheaply reduced by a simple change in mining procedures, it may be possible to simplify or reduce the capacity of the plant's homogenization systems, described in Section 2.2.2 and 2.2.3. Characteristics of the mining procedure that might affect variability include: the dimensions and orientations of mining blocks; the number of blocks being mined and blended together; the possible use of blast-hole or core-drill information to select the blocks to be blended; the possibility of splitting some heterogeneous blocks into ore and waste subblocks; and the possibility of routing mined ore to different stockpiles having different chemical qualities and subsequently proportioning the outputs of these stockpiles to achieve an overall quality aim.

Conditional simulations can also be used to study the effects of the spatial distributions and sizes of high- and low-grade areas on the total recoverable reserves of a deposit that is being mined selectively; to study the effects of dilution at contacts between rock types (Section 3.10.2) by conditionally simulating the contact and the rock compositions on either side of it (Section 3.8), and then simulating a mining operation recovering ore from the vicinity of the contact; and to study the effectiveness of various estimation procedures on the prediction of local ore grades and of recoverable reserves.

3.9.2 Mineral-Processing Applications

> Any resemblance between prospect data and actual plant feed is coincidental.
> The design average is that statistical point through which a value most frequently passes when going from one extreme to the other.
> Remember, if it can, it will. If it cannot, it might.
> [selections from "Raulerson's Rules for the Design of Mineral Beneficiation Plants", by John Raulerson, Jr., as quoted in Mining Engineering, March 1985, p. 207]

3.9.2.1 Design of Homogenization Facilities

> The most important and most difficult feature to specify is the size of the stockpile required to achieve the required reduction in material variability between successive stockpiles. If the design is insufficiently precise then the stockpiles must be made extra large for safety, causing excessive cost. [Parnaby et al., 1973, p. 325]

The output of a mine (or of a mining simulation operating on a simulated mineral deposit) is a time series of (possibly multivariate) ore grades and ore tonnages. Stockpile homogenization systems (Section 2.2.2.1) are designed to filter out some of the variability in this time series. To a close approximation (explained below), the effect of a stockpile

prehomogenizer can be concisely expressed, in geostatistical terms, as a regularization (Section 3.1.6) of the time series. (If this is not apparent, a quick review of Figures 2 and 3 should make it so.) If the mine-output time series were stationary and observable over a long period of time, the effectiveness of any homogenizer could be calculated straightforwardly from the variogram of the time series, using the formulas for dispersion variance and regularization provided in Sections 3.1.5 and 3.1.6. However, because actual observed or simulated mine-output time series commonly do not exhibit stationarity and ergodicity, it is usually safer, and perhaps easier as well, to study the effects of various homogenization systems experimentally, by calculating simple moving averages of the time series data using several different lengths (spans) for the uniform moving-average weight function.

If both tonnage and grade time series are available, the moving average might be calculated over the time series of accumulations (tonnage x grade), or the time series might be converted to a "tonnage series", in which grade is expressed as a function of the cumulative tonnage of ore delivered to the stockpile. The tonnage-series representation is the correct choice if the stacker is designed to stack a constant tonnage per unit length along the axis of the stockpile (a good idea anyway). This is the only approach available in a simulation study, unless one is prepared to simulate a time series of both tonnages and grades delivered to the stockpile. A detailed simulated time series of tonnages would involve some modeling of the hauling and crushing systems, or reliance on historical tonnage data.[87] This level of detail

---

[87] The important word in this sentence is "detailed". In the case study of Section 4.2.5, a time series of tonnages <u>is</u> simulated without the

is not justified in most simulation studies of homogenization systems, where the main parameter of interest is the size of the pile. The rate at which the pile is built is of little or no relevance in determining pile sizes.

The principal objective of a homogenization system is to reduce the variability $D^2(Q/L)$ of a raw material to be used in the plant. The support "Q" is a "critical sampling quantity" (Schofield, 1980, p. 78) that is actually determined by the sampling system and quality-control objectives of the plant, but in this discussion it is more convenient to identify Q with the amount of material contained in one "slice" recovered from the face of the stockpile by a reclaiming device. Some reclaimers, such as bucket wheels, move cyclically back and forth across the face of the pile, so that a slice can be well defined as the material reclaimed from the pile during a single cycle. Some more elaborate reclaimers attempt to recover material from the whole face simultaneously and continuously, so that the definition of a slice may become arbitrary. The support "L" may be taken to be "∞" if the plant operates continuously for a long period of time, or it may be some finite tonnage of material if the plant works on a "batch" basis. (For example, in a cement plant, "L" may represent the total amount of raw material consumed by the plant during a "run" of a particular type of clinker.) Whatever the case, once the quantities Q and L are defined, the stockpile's efficiency can be measured in terms of the variance ratio $D^2(Q_o/L)/D^2(Q_i/L)$, where Q_i represents a continuous segment of the

---

need of any information on the hauling and crushing systems, but it is on a pile-to-pile scale. In that simulation, each pile of ore in the time series contains output from the same number of mining blocks, and the recoverable reserves of the blocks are not constant.

stockpile input stream containing a tonnage Q of material, and Q_o represents a similar segment of the output stream from the stockpile. Each variance can be decomposed into "slice" and "pile" terms, i.e.,

$$D^2(Q/L) = D^2(Q/P) + D^2(P/L)$$

where "P" is the amount of material contained in a full stockpile. A stockpile homogenizer of a given capacity is able to reduce only $D^2(Q_o/P)$, the slice-to-slice output variability from an individual pile. $D^2(P_o/L)$, the pile-to-pile output variability, remains equal to $D^2(P_i/L)$. However, $D^2(P_i/L) = D^2(P_o/L)$ can be decreased by increasing the size of the pile or by somehow decreasing the large-scale variability of the input time series, e.g., by selective mining, or by blending mined materials from several faces into each pile. (All of these methods for reducing $D^2(P/L)$ are simulated in Section 4.2.5.)

The effect of a stockpile homogenizer can be represented exactly as a regularization if $D^2(Q/P)$ can be reduced to zero,[88] i.e., if the material within the stockpile has been <u>completely</u> homogenized. In theory this could be accomplished by building up the pile from a practically infinite number of very thin layers. In practice this variance never vanishes and does not usually decrease noticeably beyond about 100 layers (Gy, 1981), because the varying compositions of the individual particles contained in the slices will always cause different slices to differ in composition, no matter how many and how thin the layers. This variability at the particle scale corresponds to Gy's (1982) "constitution and distribution heterogeneities" and comprises a

---

[88] In the unlikely case that simulated data are available on a smaller support than Q, we would have to reduce $D^2(data/P)$ to zero to have a rigorously defined regularization; however, the component $D^2(data/Q)$ of this variance is probably of no practical importance.

large part of the nugget effect observed in the input variogram of a real (not simulated) stockpile-input time series. It can be reduced only by increasing Q or decreasing the particle size within Q; otherwise this nugget constant persists in the output variogram whereas all other variation contained within $D^2(Q_o/P)$ can be effectively eliminated.[89] A reduction in particle size is accomplished in the mill, but only after the individual slices have already been recovered, so it does not affect $D^2(Q_o/P)$. However, if the mill has a large enough capacity to mix several slices of material together, or if the plant has a powder homogenization silo (Section 2.2.3), the output from the total stockpile-mill-silo system may have very nearly a zero $D^2(Q_o/P)$. In any case $D^2(Q_o/P)$ is very small in comparison to $D^2(Q_i/P)$ for a many-layered pile, so the view of homogenization as a regularization involves very little approximation in most plants.

In a simulation study, we can normally assume that the stockpile will be built of a sufficient number of layers to make $D^2(Q_o/P)$ small in comparison to $D^2(P/L)$ and insignificant after grinding and powder homogenization. Then the only influences on $D^2(Q_o/L)$ that must be considered are those involving $D^2(P/L)$: (1) the <u>size</u> of the pile, which in a simulation is represented by the span of the moving-average weight function applied to the input time series, and (2) the mining methods, the important features of which presumably are embodied in the mining simulation (Section 3.9.1). A simulation study of alternative stockpile homogenization systems is thus reduced to a mere smoothing of the outputs of various mining simulations using moving-average smoothers of

---

[89] A convincing plant-scale experimental study of these effects, conducted in a French cement plant, is described by Gy (1981).

various spans. The resulting sample values of $D^2(Q_o/L)$ are then compared with the value needed for acceptable plant performance.[90] The smallest (cheapest) pile that seems assured of producing an acceptably smooth output is the one chosen for installation. Batch and continuous piles (Figures 2 and 3) with the same active capacity[91] are represented by the same weight function. The only difference is that in a continuous system, the weight function slides over the time series continuously (point to point in a discrete simulation) as in Figure 3(c), whereas in a batch system it jumps from one pile-sized increment to another, attributing the average within the input increment to all slices within the corresponding output increment, as in Figure 2(b). Thus we have a straightforward procedure for specifying the size required for a homogenization stockpile; the "most important and most difficult" problem posed in the quotation at the beginning of this section can be easily solved if we have reliable mine-output time-series simulations.

We cannot specify the sizes of mills and powder homogenization facilities in this way. These facilities should be sized to handle variability on a particulate scale and variability that may be created by segregation in the stockpile reclaiming system or in the mill itself. These sources of variability might be modeled separately in a

---

[90] Criteria other than dispersion variance may be selected. In the case study of Section 4.2, the mean absolute difference between successive pile compositions (easily obtainable only by simulation) is used as a criterion (Section 4.2.5.3).

[91] The circular stockpile shown in Figure 3 has the same "active capacity" as a linear pile of the same size only if the stacker is stacking over the whole length of the pile. If part of the pile is just dead storage waiting to be reclaimed, the efficiency of the pile is reduced.

comprehensive plant engineering study, but no useful information at this scale is available in the simulated time-series data obtained by the methods described here. Of course, direct measurements of short-scale variability can be obtained after the plant is built and in production (or might be obtained from a similar plant built previously that uses similar materials) and certainly should be obtained if a powder homogenization system is to be designed for installation in an existing plant. Methods for collecting data on short-scale variability, primarily for the design of sampling and control systems but applicable to short-scale homogenization systems as well, are described by Gy (1982, Chapters 6 and 14). An example of the use of those methods to illustrate the effects of segregation by a bucket-wheel reclaimer is provided by Gy (1981).

3.9.2.2 Design of Proportioning Facilities

If time-series realizations (real or simulated) are available for all (or at least all "variable") raw materials to be used by a plant, valuable information can be gained on the degree of flexibility that should be built into the plant's proportioning facilities. The solutions to deterministic long-term proportioning problems (the "mix-design" problems of Section 2.3.2) provide only an impression of the "typical" amount of each raw material that will be required by the plant. If chemical quality specifications are tight and some of the raw materials are variable in composition, the actual day-to-day or even year-to-year raw-material proportions that the plant will require may depart severely from the deterministic solution.

To assess how serious these departures will be, we can apply the quality-control proportioning algorithm (or maybe several alternative algorithms) to be used in the plant to the time-series mine-output data and observe experimentally how much the mix proportions will vary from time to time. In this application the time-series data must first be expressed as "tonnage series" (described in the preceding section), because the rate at which each material is consumed will vary according to the instructions provided by the proportioning algorithm. The most important results to be collected from such a study probably are the ranges in mix proportions calculated for each raw material (Figure 65). The maximum proportions suggest how much storage and feeder capacity should be installed for each material. The minimum proportions suggest whether facilities should be available to measure out very small quantities of some materials, and whether some sources of raw materials may not be used at all from time to time.

Mining and proportioning simulations may be carried out simultaneously, incorporating feedback and feedforward control loops (Section 2.2.2.3) between the proportioning algorithm and the mine(s), in the manner depicted in Figure 6. Simulated "blast-hole data" extracted from the conditional simulations of the various raw materials, in combination with real core-drill data, might be used to obtain simulated estimates (possibly by kriging) for feedforward control, and the current and past simulated tonnage-series data encountered by the proportioning algorithm could be used for feedback control. "Control" might amount to nothing more than changes in mix proportions in the plant, as depicted in Figure 6, but it might also incorporate an

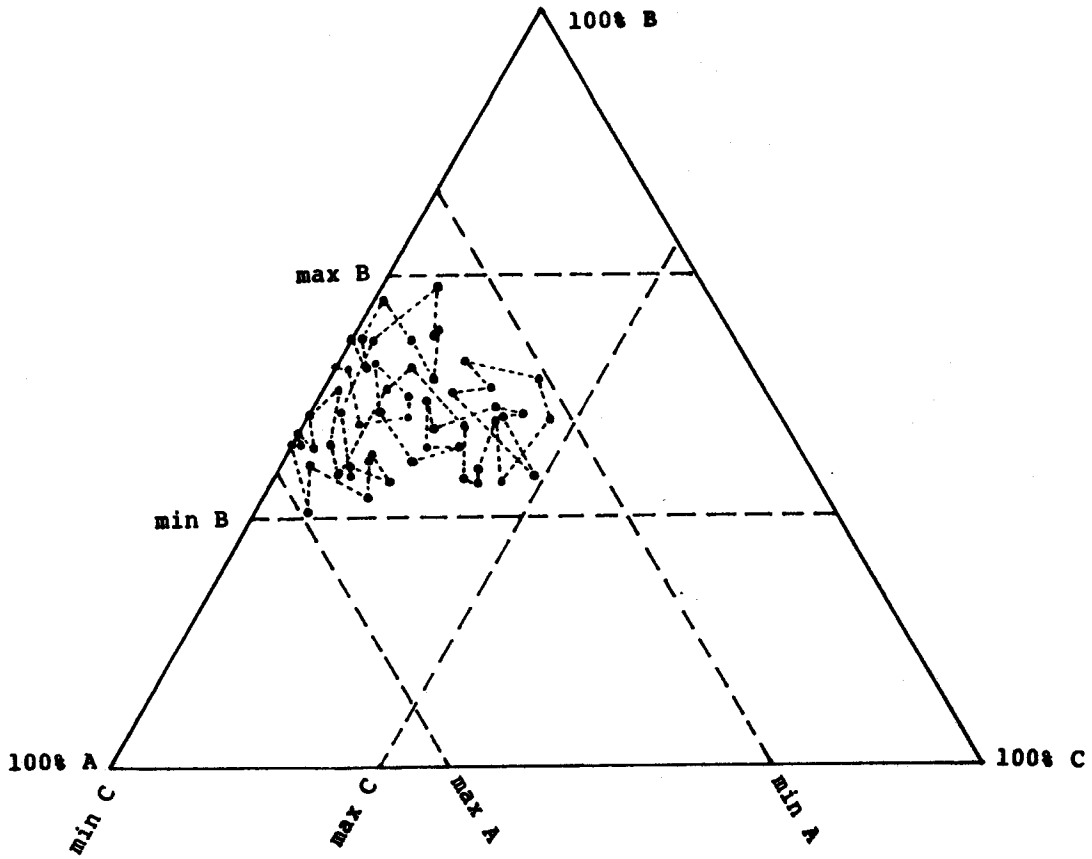


Figure 65: The usage of simulated tonnage-series data to assess variability in mix proportions. In this imaginary example, three raw materials (A, B, and C) are being proportioned to obtain successive equal-sized batches of mix with a constant target mix composition. As the raw materials vary in composition, the proportions also change (broken dotted line). Proportioning and storage facilities must be flexible enough to accommodate the long-run extremes in mix proportions represented by the dashed lines and the short-run changes in proportions indicated by the dotted line.

engineer's decision to change working faces in one of the simulated mines. An integrated mining and proportioning simulation of several years' mining operations might be performed interactively by a mining engineer and a quality-control supervisor in a few days' time.

No proportioning simulations are included in the two case studies of Chapter 4, as only data sets from single raw materials were available. (However, in both case studies these single raw materials are so variable that proportioning the outputs from different parts of the same mine would be a thinkable option.) Data sets of sufficient size and quality for a geostatistical study currently are a rarity in the cement industry, even for single raw materials. Drilling data adequate for reliable conditional simulations of all materials being used by a plant may not exist anywhere at this time. Nevertheless, Deraisme and Dumay (1981, Annexe B) have performed a multi-material proportioning study using data from a French cement plant. (This study is briefly reported in English by Deraisme and de Fouquet, 1984.) In that study, drilling data adequate for a conditional simulation were available for only one of the four materials being used. The other materials were assumed to be constant in composition. The four materials were being stacked onto a common blending stockpile, with an attempt being made to meet a target calcium concentration in the pile while respecting a maximum magnesium constraint and a target pile size. The high magnesium content of the principal (simulated) limestone commonly required the addition of a large amount of expensive purchased high-calcium limestone. Four mining simulations were performed on the principal limestone deposit: one representing actual mine practice, and three others using kriging to estimate block contents before mining, and using three different orientations for the mining blocks. The first simulation faithfully reproduced plant experience, whereas the improved control afforded by the kriged estimates enabled the simulated proportioning operation to decrease its average consumption of the expensive limestone.

3.10 SOME SPECIAL PROBLEMS

3.10.1 Identifying Multiple Populations

Section 3.8 suggests some methods for conditionally simulating the locations and compositions of multiple rock types occurring within the simulation domain. These methods presuppose that the rock types have already been distinguished in the conditioning data. For example, in the case study of Section 4.2, the limestone and dolostone populations were visually identified in the drill cores by a field geologist before the sampling and analysis of the cores were performed. But in many data sets no classification has been made *a priori*, and indeed a meaningful classification not based on the chemical data is not always possible. Then the chemical analyses should be examined to see whether the existence of multiple populations is indicated. If only one or two variables are of interest, the examination can be done graphically. Histograms (one variable) or scattergrams (two variables) of the data may reveal whether natural groupings of data values exist, if they can be distinguished without reference to their locations.

Spatial groupings of data values may also be present even though no clusters appear on histograms or scattergrams, if the compositions of the spatial groups extensively overlap. These groups may be revealed on contour maps or cross sections as zones with different contour levels or patterns. (In the case of multivariate data, some function of the data, such as a principal component, might be mapped.) It is good practice to plot maps, cross sections, histograms, and a variety of bivariate scattergrams to examine the data for unsuspected clusters or trends.

Discernment of multiple populations can be difficult in a high-dimensional multivariate data set. Trivariate scattergrams can be rotated and examined from different angles on computer graphics terminals, and split-screen displays of several multicolored trivariate scattergrams have proven useful in the examination of data sets with more than three dimensions (McDonald, 1982). In addition to graphical methods, many numerical methods for identifying groups within multivariate data sets, collectively called "cluster analysis", have appeared during the past twenty years. These are well described by Everitt (1980), Gordon (1981), and Hand (1981, Chapter 7). The class memberships assigned by any of these exploratory methods must be examined carefully, especially in the case of spatially distributed data. A useful approach is to plot maps and cross sections denoting the locations of samples assigned to different clusters with different symbols or colors. These plots can then be examined to see whether the spatial distribution of clusters makes geological sense.

Location or contiguity of data can be built into a classification system so that samples will tend to be lumped into spatially compact groups. Some methods, called "constrained classification", that achieve this are reviewed by Gordon (1981, p. 61). Methods particularly applicable to geological data sets have been presented by Webster and Burrough (1972)[92], who are concerned with soil classification and mapping, and by Sinding-Larsen (1975), who is concerned with subdividing a region into geochemically homogeneous areas identified by soil

---

[92] This paper is summarized by Marriott (1974, pp. 85-89).

samples. Some related techniques used for classifying remote-sensing data are described by Switzer (1983).

3.10.2 Dilution of Ore Grades During Mining

"Dilution" is a mining term meaning the contamination of mined ore by waste or off-grade materials. Typically dilution occurs when mining activities accidentally wander across an ore-waste contact and a small amount of waste material is incorporated into the recovered ore, "diluting" its grade. The effects of this kind of dilution on the time (or tonnage) series of ore grades delivered to a processing facility can be studied by simulation, provided that the ore grades in question have been simulated in both ore and waste populations (unless the grades are effectively zero in one of the populations), and provided that the locations of ore and waste within the simulation domain have also been simulated. Methods for performing simulations of multiple populations are discussed in Section 3.8. Simulated mining procedures can be applied to a multiple-population conditional simulation to observe what happens if ore recovery is not perfectly confined within the boundaries of the ore body. Some examples are provided by the mining simulations of Section 4.2.5.1, in which the limestone ore is contaminated by dolostone in amounts varying with the degree of selectivity of the mining procedure.

In surface mines, dilution by incompletely stripped overburden may be an important source of contamination of the ore. For instance, clay caught in small solution cavities near the surface of a high-grade limestone deposit may contribute most of the variability observed in the

delivered limestone grades. The effects of this kind of dilution are more difficult to reproduce in a simulation, as the amount of overburden recovered depends not only on the nature of the bedrock surface but also on the effectiveness of the stripping operation, which is difficult to model. In some cases a comparison of historical delivered ore grades with simulated grades of clean ore mined from the same area may permit estimation of the statistical properties of overburden contamination experienced in the past. For a new mine, one would simply have to make "reasonable assumptions" or rely on experience from similar operations.

In limestones and other rocks subject to the development of solution features, overburden may contaminate the ore not only at the bedrock surface but also within cavities and along joint surfaces throughout the deposit. Assessment of the extent of this kind of contamination cannot be made directly using only core-drilling records, as the unconsolidated clays and other materials filling the cavities usually are washed out of the rock during drilling. Serious underestimation of the amount and variability of such contamination is probably the most common reason for disappointment in the quality of newly opened limestone deposits. (Commonly the responsible geologist's reputation becomes as soiled as the raw materials.) In some cases, a better impression of the average composition of a clay-contaminated deposit can be obtained from chip samples taken from a percussion drill than from core-drill samples.

If the rock is riddled with fairly wide cavities that produce gaps in drill-core recovery, an upper limit on the amount of clay in the rock can be obtained during core drilling by carefully recording the percentage of core recovery at regular intervals down each hole.

Missing core footage may correspond to cavities (open or clay-filled), very soft or broken-up rock, or accidental grinding up of the core during drilling. Usually the driller knows how much of the missing footage is attributable to cavities. Core recovery (or percentage of cavities) can be treated as a regionalized variable and estimated or simulated with the techniques described previously in this chapter. If core-recovery data are compiled from several successive drilling programs, or possibly even from more than one driller or rig, the data from different sources should be compared carefully, as different equipment or personnel may have different degrees of success in recovering core from "problem" materials. (For example, core recovery data were available from both of the drilling programs contributing to the data set used in the case study of Section 4.2, but the recovery tended to be better in the second program. Recovery data were not used in the case study.) As the support (core length) for which recovery is recorded decreases, the core recovery tends to look more like an indicator variable (much like the percentage of dolostone simulated in Section 4.2).

Normally some hypothesis will have to be made regarding the amount of cavity space that is occupied by clay. Furthermore, samples of clay from the cavities may be few or nonexistent, so a further hypothesis on clay composition and variability is usually needed as well. (Subjective judgements of clay qualities and compositions might be incorporated via "soft kriging" (Journel, 1984e) or a similar approach.) However, even with these uncertainties a geostatistical view of clay contamination is still likely to yield a better ultimate impression of the effects of

contamination on the quality of the delivered ore than the simple application of totally subjective dilution factors to all results.

Clay contamination may be confined to sharply defined tabular "seams" occupying solution-widened joints or fracture zones, with wide intervening volumes of uncontaminated rock. In this situation a regionalized-variable model of the contamination may not seem realistic. Then direct simulation of the locations, orientations, and thicknesses of the seams may be preferable to simulation via core recovery. A straightforward, though simplistic, way to do this is to draw the dips, azimuths, trace lengths, thicknesses, and possibly depths of the seams randomly from their estimated (possibly joint) probability distributions (subjective, or from data), fixing the location of the center of each seam in space by simulation of a spatial point process, such as a Poisson process (e.g., using methods described by Diggle, 1983, Chapter 4).[93] "Conditioning" to a few known seams can be accomplished with little bias just by entering the known properties of the known seams directly into the simulated data set, drawing any unknown properties from their estimated distributions in the usual way. Seams can be kept out of areas where they are known not to occur simply by acceptance-rejection. This simple Monte Carlo method will not reproduce any spatial correlations in seam properties that might be evident from the data or from field observations. Methods for simulating spatially correlated fracture properties by spectral means are reported by Miller and Borgman (1985). Their method does not, as stated, admit the

---

[93] The method described by Lin and Harbaugh (1984) for generating random planes in three dimensions might also be used. This method is described briefly at the end of Section 3.3.1.3.

simulation of cross correlations, but extensions of the method to do this are under development. Presumably space-domain simulations of these properties could be performed as well.

3.10.3 Mixing and Segregation During Mining and Processing

Mixing and segregation of broken ore must be modeled, if at all, at the mining-simulation stage. Normally these effects occur on such a small scale in comparison to the scale of the simulation (small even compared to the density of discretization of the simulation domain) that they would not be observable in the simulation output anyway. For instance, several methods might be employed to recover a pile of broken ore blasted down from a given mining block. Each method would mix or segregate the ore from different parts of the pile, or from different size fractions of the pile, to some degree. But unless the mining block in question is large indeed or the simulation is performed on a very small support (dense discretization), any attempt to model the grades of the broken ore on this scale would push the conditional simulation well beyond its intended accuracy. Discretization errors committed by representing the block as a finite set of discrete points would likely exceed the errors committed by ignoring the mixing and segregation involved in mining out the block.

The scale at which the time-series output of a mining simulation can be useful is similarly limited by the effective support of the simulated time-series data. Even if the actual mineral deposit could be neatly mined out in a sequence of small parcels of ore corresponding to

individual simulated data[94], each datum would probably still represent several tons of ore. By the time each parcel of ore had passed through the mining, crushing, and conveying systems that would carry it to the plant, it would probably have been mixed with several other parcels. Thus the support at which simulated time-series data could be used for plant studies could certainly be no smaller than the discretization support of the conditional simulation, or this support stretched out according to the residence-time distributions of the mining, crushing, and handling systems. This means that questions about the very short-scale behavior of the actual time series in the plant cannot be answered using the output of a mining simulation applied to a conditionally simulated mineral deposit. This is just asking too much of the simulation. Questions that <u>can</u> be answered are those that involve supports at the scale of mining blocks, with compositions simulated as averages of several individual simulated data (unless a block-support simulation has been done). Thus the compositions of blending piles composed of ore from one or more mining blocks can be simulated nicely, but the minute-to-minute variability of ore streams that must be measured to calculate the accuracy of sampling systems (as in Gy, 1982, Chapter 14) is not accessible.

---

[94] And further assuming that support corrections had been made to make each simulated value representative of its entire discretization cell instead of the usual point or core support.

3.10.4 Drifts in Sedimentary Deposits

Sedimentary deposits commonly exhibit marked drifts in composition in the direction perpendicular to bedding (described below as "vertical"). If the vertical thickness of the deposit has been sampled in suitably short increments so that a long vertical sequence of samples is available from each drill hole, it may be possible to estimate the form and parameters of a drift function to describe this vertical trend within each hole. Different shapes of the drift in different holes usually can be represented by a single model for all holes with different parameters in different holes. These parameters may then be regarded as regionalized variables in the horizontal plane. Methods for estimating grades in such deposits are described by Rendu and David (1979) and by Royle et al. (1982).

A realistic simulation of such a deposit can be constructed from simulations of the drift parameters, plus a simulation of residuals representing the local departures of individual data from the drift. The drift simulation is done in two dimensions and the residual simulation in three. This type of simulation is particularly useful if the spacing between holes is wide and the drift fluctuates widely between adjacent holes. If the spacing is close enough that neighboring holes generally exhibit very similar drifts (as in the case study of Section 4.2), then the drift functions need not be modeled, as the tight conditioning will force the simulations to conform to the drifts embodied in the conditioning data.

An unconditional simulation incorporating a model of vertical drift was performed by Alfaro and Huijbregts (1974) and is also described by

Journel and Huijbregts (1978, p. 530). The simulation domain was discretized into 550x110 cells in its horizontal dimensions (x,y), with each cell represented by twenty simulated values in the vertical dimension (z). The regionalization of grades g(x,y,z) was considered to be the sum of two independent terms:

$$g(x,y,z) = t(x,y,z) + r(x,y,z)$$

where t(x,y,z) is the trend and r(x,y,z) the residual. The vertical trend was given a fixed cubic form:

$$t(x,y,z) = \alpha \left[a\, z^{3\beta} + b\, z^{2\beta} + c\, z^{\beta} \right]$$

where a, b, and c are constants and α and β are regionalized variables $\alpha(x,y)$ and $\beta(x,y)$ in the two-dimensional horizontal space. In this representation the parameter $\alpha(x,y)$ controls the amplitude of the vertical drift function at horizontal location (x,y), and $\beta(x,y)$ controls the function's shape -- specifically the vertical (z) locations of the local extrema of the trend function at (x,y). The residuals r(x,y,z) were modeled as the sum of three nested structures -- a nugget, an isotropic horizontal structure with a long range, and an isotropic three-dimensional structure with a much shorter range. In this way the residuals were correlated over long distances horizontally but only over very short distances vertically, as one would expect in a sedimentary deposit.

Chapter IV

CASE STUDIES

This chapter describes two case studies involving simulations of limestone deposits. Both deposits are being actively quarried as cement raw materials. The first case study is relatively short and is intended primarily to illustrate the use of principal components in multivariate simulations. The second case study is much longer, involving reproduction of constraints on the simulated values, simulations of multiple populations and their contacts, unfolding of deformed strata, and some applications of simulations to mine planning and the design of homogenization systems.

4.1 A SIMULATION OF COREGIONALIZATION OF LIMESTONE COMPOSITIONS USING PRINCIPAL COMPONENTS

4.1.1 Description of the Data

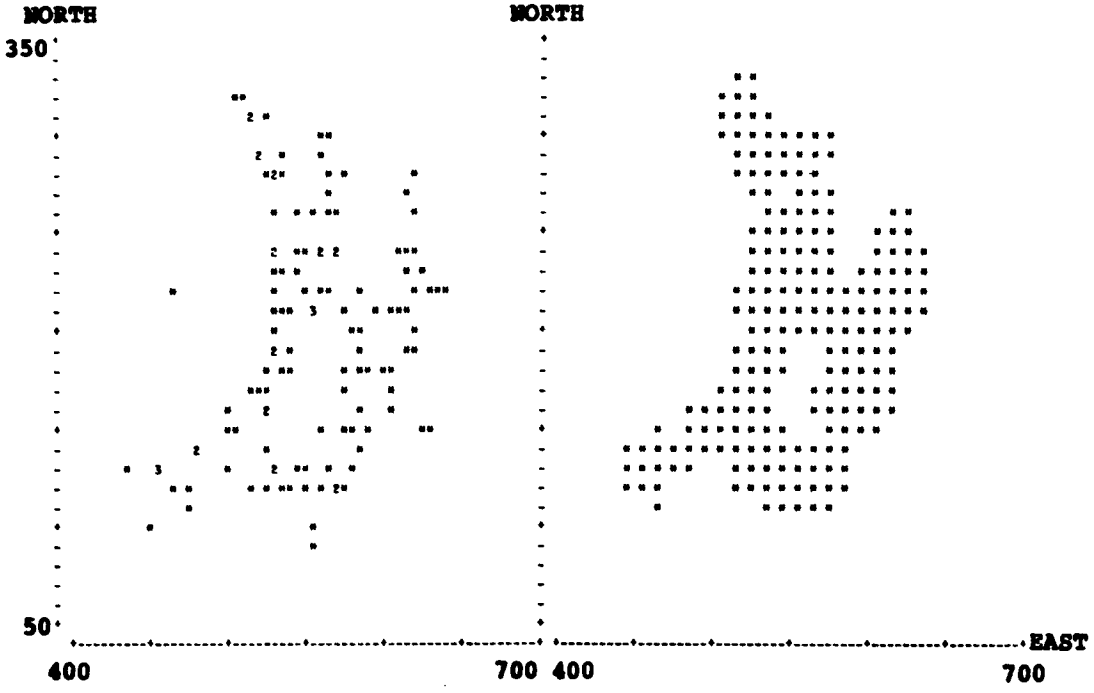
This brief case study is not directed toward a specific practical problem but merely illustrates some of the methods that can be used to simulate a coregionalization of several variables. The data consist of a set of limestone analyses from a quarry near Plymouth, England, provided by Schofield (1980, Chapter 11, pp. 196-212, and Appendix 11c, pp. 304-310). These analyses are estimates of the overall compositions of mining blocks, each obtained from a composite of the blast-hole samples collected within a block. Only 129 blocks located in the upper level of the quarry have been selected for this two-dimensional study.

The support of the data is not constant: the number of blast holes per block, the tonnage of stone per block, and the orientations of the roughly rectangular blocks are all highly variable. Not surprisingly, the larger blocks (in terms of either holes or tonnage) are slightly less variable in composition than the smaller blocks. However, this effect is small enough to be neglected for the purpose of this study, so all data are treated equally regardless of support.

The continuous two-dimensional spatial domain of the data, V_0 (see Section 3.7.1), is well defined in this case, as the block support of the data fills the entire domain (although there are some unsampled interior blocks). The general shape of V_0 can be seen clearly in the plot of block centers in Figure 66(a).[95] Within this domain there is almost no formal estimation error EE (Section 3.7.3), because the domain is nearly exhaustively sampled on block support. The most common general objective of conditional simulation is to manufacture additional data for a more exhaustive coverage of the domain. In the case of exhaustive block data this objective has already been realized, so for the present case study a different objective is proposed: to replace the irregularly spaced (owing to variable support) block data with a new set of regularly spaced, constant-support block data representing square blocks of roughly the same size as the average size of the real blocks. The centers of the 206 simulated blocks are plotted in Figure 66(b). Applications of such a simulation can be imagined; for instance, mine-

[95] The isolated point in the west-central part of Figure 66(a) may represent a block in the lower level of the quarry, although it is included in the upper-bench data by Schofield (1980, Appendix 11c). Because of its isolation and its unremarkable analysis, it will have almost no effect on the simulation, even though it is left in the data set.

planning and mine-simulation programs can be more easily devised to work on constant-sized blocks.



(a) Original data (b) Simulated data

Figure 66: Maps showing the locations of the 129 original Plymouth data and the 206 simulated data. The scale is in meters. Numerals indicate multiple locations close together.

The five variables to be simulated are SiO_2, Al_2O_3, Fe_2O_3, CaO, and MgO. Some statistics for the 129 original data are provided in Table 7(a). These five variables, plus loss on ignition (mostly CO_2 and H_2O), which was not determined, should account for nearly 100% of the

limestone. The most important undetermined oxides are likely to be K_2O, Na_2O, and TiO_2 (in clays or other silicates), and SO_3 (possibly as pyrite, FeS_2). Schofield does not describe the mineral composition or detailed geology of the limestone, but the proportions of SiO_2, Al_2O_3, and Fe_2O_3 in the analyses suggest that these three oxides might occur in silicate minerals of roughly illite composition. There appears to be little excess SiO_2 to suggest the presence of much sand or chert in the limestone. The small amount of MgO probably occurs primarily as dolomite ($CaMg(CO_3)_2$). The remainder of the limestone is likely to be almost entirely calcite ($CaCO_3$). As the percentage of calcite, dolomite, and silicates should add up to nearly 100% in all of the samples analyzed, it is likely that, among the five measured variables, there are two significant sources of variation: the proportion of carbonates to silicates, and the proportion of dolomite to calcite within the carbonates. This interpretation is supported by the correlation matrix in Table 7(a), which reveals very strong positive correlations among SiO_2, Al_2O_3, and Fe_2O_3 (constituents of the silicate fraction), strong negative correlations between these oxides and CaO (the major constituent of the carbonate fraction), and relatively uncorrelated behavior for MgO.

4.1.2 Check for a Constant-Sum Constraint

The implied mineralogical makeup of the limestone suggests some potentially useful "normative" (in the petrologic sense) transformations of the chemical data. Suppose we allocate all SiO_2, Al_2O_3, and Fe_2O_3 to the silicate fraction, plus some of the CaO and MgO, and also some

unanalyzed H_2O, K_2O, Na_2O, and TiO_2. Then we might obtain an approximate percentage of "clay" in the limestone by the formula

$$\text{"CLAY"} = 1.2 \left(Si_2O + Al_2O_3 + Fe_2O_3 \right)$$

where the available SiO_2, Al_2O_3, and Fe_2O_3 are determined from the analyses, and the factor 1.2 accounts for a proportionate amount of CaO, MgO, K_2O, Na_2O, TiO_2, and H_2O in the clay. This provides a normative percentage of a "clay" that is consistent with the composition of illite. Because some CaO and MgO are allocated to the clay, they must be subtracted from the analyzed CaO and MgO available to the carbonate minerals. Thus the total calcium carbonate (in calcite and dolomite) and magnesium carbonate (in dolomite) are obtained by

$$\text{"}CaCO_3\text{"} = 1.785 \left(CaO - 0.02 \text{ "CLAY"} \right)$$

$$\text{"}MgCO_3\text{"} = 2.092 \left(MgO - 0.01 \text{ "CLAY"} \right)$$

where the factors 1.785 and 2.092 are the ratios of the molecular weights of the corresponding carbonates and oxides. <u>If</u> these mineralogical assumptions are reasonable, and if the oxide analyses are accurate, the total of "CLAY" + "$CaCO_3$" + "$MgCO_3$" should be essentially 100%, or a little less. Transformations to remove linear constraints, described in Section 3.5.2.4, might then be applied to these data before using them in a conditional simulation. A plot of "CARBONATE" = "$CaCO_3$" + "$MgCO_3$" versus "CLAY" is provided in Figure 68(a). It is clear from this plot that the relatively high-carbonate samples seem to contain <u>too much</u> carbonate. A revised set of transformations cooked up to provide a better fit,

$$\text{"CLAY"} = 1.25 \left(SiO_2 + Al_2O_3 + Fe_2O_3 \right)$$

$$\text{"}CaCO_3\text{"} = 1.73 \left(CaO - 0.02 \text{ "CLAY"} \right)$$

$$\text{"}MgCO_3\text{"} = 2.03 \left(MgO - 0.01 \text{ "CLAY"} \right)$$

$$\text{"CARBONATE"} = \text{"}CaCO_3\text{"} + \text{"}MgCO_3\text{"}$$

is plotted in Figure 68(b). These transformations come closer to satisfying the 100%-maximum constraint, but there is still a great deal of scatter beneath (and some above) the constraint. The elliptical clouds of points in Figures 68(a) and (b) could just as easily represent a sample from a bivariate normal distribution. The second set of transformations is also more difficult to explain mineralogically. Perhaps the mineral composition is rather different from what was supposed (a possibility, because some rocks in the region have been considerably metamorphosed), or perhaps the analyses are simply lacking in accuracy and precision. Whatever the case, the constraint is not manifest in the data, so no effort is made to reproduce it in this case study. The second case study in this chapter will provide a much clearer example of the treatment of linear constraints.

4.1.3 Description of the Simulation Procedure

An outline of the simulation procedure is provided in Table 2. This procedure conforms closely to that illustrated in Figure 35. The general idea is to transform the raw data into approximately normally distributed (at least at h=0) data, conditionally simulate the principal components of those data, then invert the principal-components transformation and all other transformations in reverse order to obtain a conditional simulation of the original phenomenon.

TABLE 2

Outline of the simulation procedure for the Plymouth case study.

CHECKS	STEPS (refer to subheadings in Section 4.1.3)
+---->	1. Statistics, plots, and variograms of original data
	2. Regression to obtain linear trends
	3. Subtraction of trends to obtain residuals
+--->	4. Statistics, plots, and variograms of residuals
	5. Normal scores of residuals
+-->	6. Statistics, plots, and variograms of normal scores
|	7. Principal components of normal scores
++>	8. Statistics, plots, and variograms of principal components
||	9. Variogram models for principal components
||	10. Independent simulations of circular structures
||	11. Principal components and linear model for correlated nugget structures
||	12. Independent simulations of nugget structures
||	13. Linear combinations of nugget structures
||	14. Combinations of correlated nugget and independent circular structures to get unconditionally simulated principal components
+>	15. Statistics, plots, and variograms of simulated components
|	16. Ordinary kriging with principal-component data
|	17. Ordinary kriging with simulated components
|	18. Combination of unconditional simulation and two krigings to get conditional simulation of components
+->	19. Statistics, plots, and variograms of conditionally simulated components
|	20. Inversion of principal-components transformation to get conditionally simulated normal scores
+-->	21. Statistics, plots, and variograms of simulated normal scores
|	22. Inversion of normal-scores transformation by linear interpolation to get conditionally simulated residuals
+--->	23. Statistics, plots, and variograms of simulated residuals
|	24. Restoration of linear trends
+---->	25. Statistics, plots, and variograms of final conditional simulation

Statistics, semivariograms, scatterplots, and maps of the data or simulations are obtained after each transformation or simulation so errors in the transformations or poor reproduction of the data characteristics by the simulations can be detected as early as possible. For this small simulation, the easy-to-use statistical package MINITAB (Ryan et al., 1976) is employed for checking statistics, plotting data and semivariograms, and performing some transformations and manipulations of data. All of the illustrations presented in this case study were constructed with MINITAB. The BMDP4M ("Factor Analysis") program (Dixon and Brown, 1979) is used to obtain principal components and a linear model for correlated nugget structures. The only other routines used in the study are subroutine CS2D (Appendix A) to perform unconditional circular-variogram simulations in two dimensions; unpublished Stanford subroutine OKB2D, used to perform ordinary kriging in two dimensions; and a modification of Verly's (1984b) subroutine LINT to invert the normal-scores transformation performed by MINITAB.

The titles of the following sections refer to the numbered steps in Table 2.

4.1.3.1 Drifts and Geologic Structure: Steps 1-4

Schofield (1980, p. 200) states that the limestone has been folded, but the quarry benches are essentially horizontal. The result is that different sets of strata intersect the sampled blast holes at different locations in the quarry. It is not surprising then that the analyses exhibit noticeable drifts, roughly linear in form, which appear strongest in a roughly north-south direction. The drifts are so marked

in the semivariograms of the raw data (Figure 71) that the local structures to be reproduced by the unconditional simulations are largely obscured, except in the east-west direction. Thus a set of linear trends, obtained by regression of each variable on the spatial coordinates (i.e., by "trend-surface analysis"), will be removed from the data prior to further analysis. No claims of optimality or of significance (statistical or geological) are made for these linear trends; they are simply regarded as deterministic functions to be subtracted from the data at the beginning of the simulation procedure and restored at the end, in order to make the intermediate steps easier. Semivariograms for some of the residuals from these trends are displayed in Figure 72.

4.1.3.2 Normal Scores and Principal Components: Steps 5-9

The "nscore" command[96] in the MINITAB system is used in Step 5 to obtain the normal scores of the residuals obtained in Step 3. The residuals corresponding to each of the five variables are transformed separately, as described in Section 3.5.2.2. All ten bivariate scatterplots, representing all possible pairs of the five variables, display roughly elliptical forms, confirming that the variables are at least approximately binormally distributed at h=0. It is then hypothesized that the normal scores are approximately multinormally distributed among all five variables at all possible lags h.

---

[96] This command does not arbitrarily rank or despike any tied values that appear in the input data, as described in Section 3.5.2.2, but rather assigns the same normal score to each tied value. It is thus undesirable if the input distribution contains large spikes.

Plots of the direct and cross semivariograms of the normal scores exhibit a variety of shapes and anisotropies (based upon inspection of sixty semivariogram plots, not shown here -- four directions for each of five direct and ten cross semivariograms). The fitting of a positive-definite matrix of fifteen nested, anisotropic semivariogram models to such plots "by hand" would be a daunting task, so program BMDP4M ("Factor Analysis") of the Biomedical Computer Programs series (Dixon and Brown, 1979) is used to obtain principal-components scores (Table 3) for these data, in an effort to simplify the modeling job by simplifying the cross semivariograms. As described in Section 3.5.3, the cross semivariograms among the principal components must be zero if the normal scores are intrinsically coregionalized; otherwise, some cross correlation is likely to persist among at least some of the pairs of components at some lags. Omnidirectional plots of the direct and cross semivariograms of the first three principal components are illustrated in Figure 67. As the normal-score semivariograms exhibit a wide variety of shapes, it is not surprising that the cross semivariograms of the components exhibit a few small departures from zero, indicating that the normal scores are not intrinsically coregionalized. Nevertheless, these plots (and the directional plots, not shown) do appear simpler in structure than those of the normal scores (also not shown).

Because the plots in Figure 67 are rather ragged-looking (even more so for the directional plots) owing to the paucity of data, no great effort should be expended to fit the models very carefully. An easy approximation (justified below) is adopted here: all five components are fitted by <u>independent</u> single circular semivariogram structures to

TABLE 3

Formulas for the principal components of the normal scores of the residuals.

Each of the five principal-component scores for each of the 129 observations is a linear combination of the five normal scores for that observation weighted by a vector of principal-component score coefficients. In the notation of Section 3.5.3,

$$y_j = \sum_{i=1}^{k} a_{ij} z_i \quad j = 1,2,\ldots,k$$

where in this case k=5, the y_j's are the five principal-component scores, the z_i's are the five normal scores, and each a_{ij} is an element of the 5x5 matrix of coefficients. In this case:

COMPONENT — COEFFICIENTS x NORMAL SCORES of residuals

	SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO
y_1 =	$0.26762z_1$	$+0.27276z_2$	$+0.26090z_3$	$-0.25293z_4$	$-0.08114z_5$
y_2 =	$-0.05233z_1$	$-0.04690z_2$	$+0.08599z_3$	$-0.31547z_4$	$+0.93040z_5$
y_3 =	$-1.03444z_1$	$+0.33895z_2$	$+1.56268z_3$	$+0.85473z_4$	$+0.10381z_5$
y_4 =	$1.32438z_1$	$+1.07375z_2$	$-0.42621z_3$	$+1.86176z_4$	$+0.79929z_5$
y_5 =	$-1.63935z_1$	$+2.39657z_2$	$-1.32049z_3$	$-0.50981z_4$	$-0.02183z_5$

These coefficients are obtained from the output of program BMDP4M, under the heading "factor score coefficients" (as below). The coefficients printed by BMDP4M are strictly valid only after the means of the z_i's have been subtracted, but as the normal scores are already approximately standardized, the coefficients can be used "as is" with little loss of accuracy. The relevant output of program BMDP4M looks like the following:

FACTOR SCORE COEFFICIENTS

THESE COEFFICIENTS ARE FOR THE VARIABLES AFTER THEIR MEANS HAVE BEEN SUBTRACTED.

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
SIO2	1	0.26762	-0.05233	-1.03444	1.32438	-1.63935
AL2O3	2	0.27276	-0.04690	0.33895	1.07375	2.39657
FE2O3	3	0.26090	0.08599	1.56268	-0.42621	-1.32049
CAO	4	-0.25293	-0.31547	0.85473	1.86176	-0.50981
MGO	5	-0.08114	0.93040	0.10381	0.79929	-0.02183

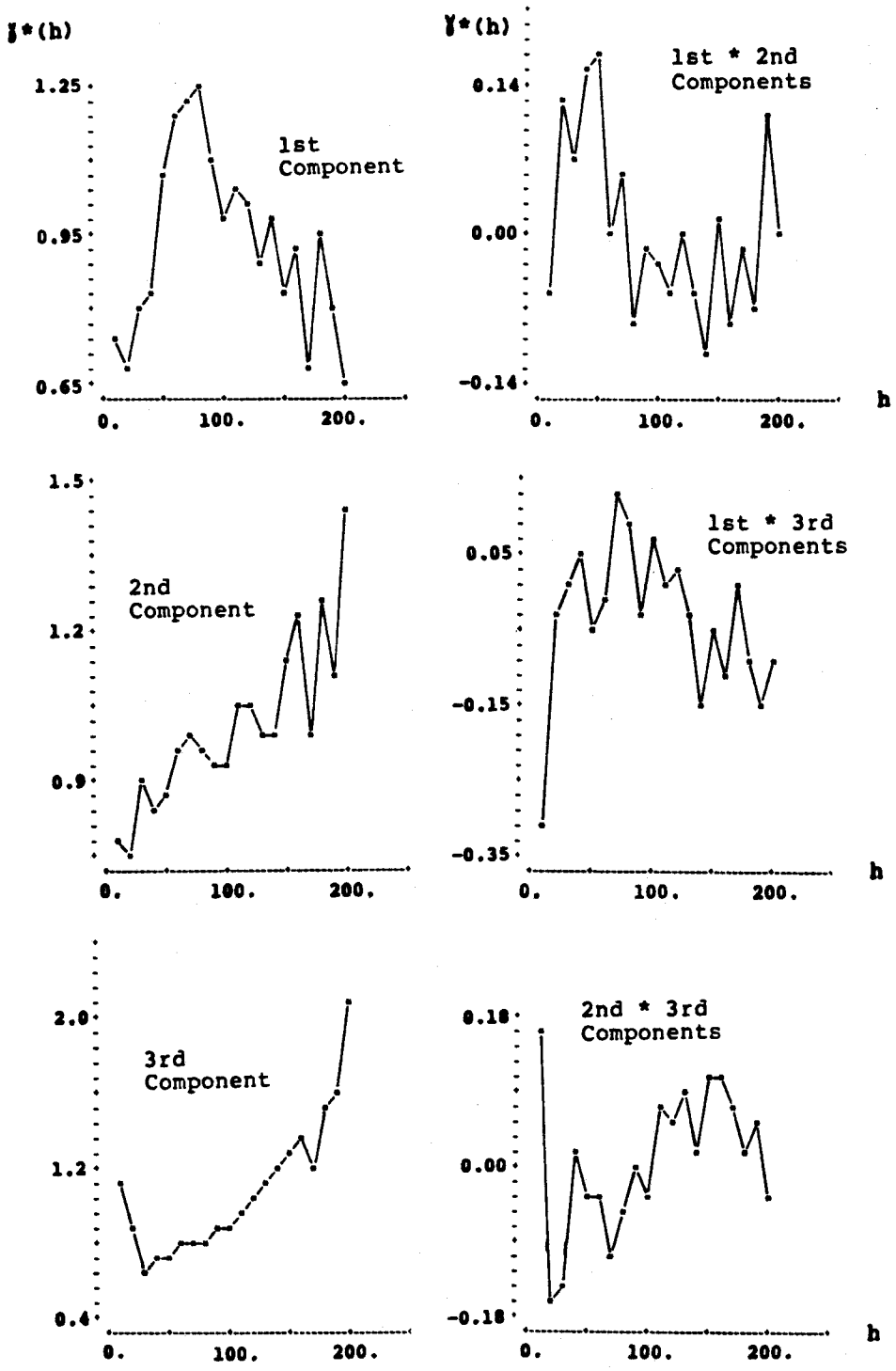


Figure 67: Omnidirectional direct and cross semivariograms of the first three principal components for 129 observations. Notice that the horizontal scale (in meters) is constant, but the vertical scales vary.

account for the short- to medium-scale behavior (up to about 200 meters) of the directional and omnidirectional plots, and dependent nugget structures to account for the very short-scale behavior of both the direct and cross semivariograms. The sample plots are so ragged (especially at small h, where there are few pairs) that it is difficult to distinguish nugget effects from short-scale transition behavior. Nevertheless, it is important to retain some fidelity to the short-scale sample structure in the unconditional simulation. The solution adopted here allocates any observed short-scale departures of the cross semivariograms from zero entirely to the nugget constants and ignores the longer-scale behavior of the cross semivariograms, trusting conditioning to restore the unmodeled features to the conditional simulation. The final matrix of semivariogram models used for the unconditional simulation is summarized in Table 4.

4.1.3.3 Unconditional Simulations: Steps 10-15

The coordinate locations of the data set (x increasing to the east, y increasing to the north) are clustered near the center of a square area bounded by x= 400 to 700 meters, y= 50 to 350 meters. For the unconditional simulations, data are generated on a matrix of 60x60 points spaced 5x5 meters apart. The southwesternmost point is assigned to (x=405,y=55). The northeasternmost point is thus (x=700,y=350).

Subroutine CS2D (Appendix A) is used to generate five independent unconditional circular simulations as described in Table 4. IMSL subroutine GGNPM[97] (IMSL, 1982) is used to generate five independent

[97] Many other normal generators are available; e.g., IMSL subroutine GGNML, or the "nran" command in MINITAB, or any of the

TABLE 4

Nested nugget and circular semivariogram models for the five principal components to be simulated.

INDEPENDENT ANISOTROPIC CIRCULAR STRUCTURES

VARIABLE	P.C. 1	P.C. 2	P.C. 3	P.C. 4	P.C. 5	
SILL	0.6	0.3	0.2	0.1	0.15	
N-S RANGE	150	200	80	50	150	(meters)
E-W RANGE	80	80	80	50	80	(meters)

DEPENDENT NUGGET STRUCTURES -- matrix of sills

VARIABLE	P.C. 1	P.C. 2	P.C. 3	P.C. 4	P.C. 5
P.C. 1	0.6				
P.C. 2	0.1	0.7			
P.C. 3	0.0	-0.1	0.6		
P.C. 4	0.1	-0.1	0.0	0.7	
P.C. 5	0.0	0.1	0.1	0.0	0.75

sets of standard normal data for the nugget structures. These nugget data are combined using a linear model obtained from another run of program BMDP4M.

To obtain a linear model for the nuggets from BMDP4M, one treats the matrix of nugget constants as a covariance matrix. Program BMDP4M calculates the eigenvalues of this matrix, providing an immediate check

---

transformations of uniform random numbers suggested by Rubinstein (1981, pp. 86-91).

on positive definiteness (all eigenvalues must be positive), and provides a linear model in the form of a matrix of "unrotated factor loadings (pattern)", reproduced in Table 5. The loadings are then treated as the coefficients of a linear model of coregionalization for the nugget simulations, as illustrated in Table 5. (This amounts to the "eigenvalue-eigenvector approach" to multinormal simulations discussed in Section 3.3.1.1.[98])

To obtain the final unconditional simulation for this case study, the independent circular simulations are multiplied by the square roots of the sill values listed in Table 4 to obtain the correct circular sills, and the coregionalized nugget data are simply added on.

4.1.3.4 Conditioning: Steps 16-19

The unconditionally simulated principal-component scores are distributed within a grid of 60x60 locations at 5x5-meter spacings. However, the original data are irregularly spaced, with integer locations nearly always more than five meters apart but not necessarily falling on the grid. Because the CS2D subroutine executes very slowly when the simulation grid is very dense, it is impractical to create a 300x300-meter simulation grid at 1x1-meter spacings. Thus an approximation must be made to obtain simulated values at the exact data locations for kriging purposes: the coordinate location of each real datum is simply truncated down to the nearest 5-meter grid location, and

---

[98] An even simpler approach that might be used is to employ a routine such as IMSL subroutine GGNSM (IMSL, 1982), which employs a factorization (also described in Section 3.3.1.1) of a specified covariance matrix to generate independent vectors of correlated normal random numbers.

TABLE 5

A linear model of coregionalization for the nugget simulation, obtained by principal components.

The matrix of nugget sills from Table 4 is (multiplied by 10):

$$\begin{array}{ccccc} 6 & & & & \\ 1 & 7 & & & \\ 0 & -1 & 6 & & \\ 1 & -1 & 0 & 7 & \\ 0 & 1 & 1 & 0 & 7\frac{1}{2} \end{array}$$

Program BMDP4M obtains principal-components loadings and eigenvalues for this matrix, reported as below:

Output from BMDP4M:

UNROTATED FACTOR LOADINGS (PATTERN)

FOR PRINCIPAL COMPONENTS

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
PC1	1	0.392	0.265	1.71	1.36	-1.00
PC2	2	2.06	-0.839	0.855	0.714D-01	1.14
PC3	3	-0.115	1.44	-0.855	1.52	0.934
PC4	4	-1.04	1.32	1.71	-0.826	0.745
PC5	5	1.76	1.80	-0.428	-0.774	-0.614
	VP	8.605	7.833	7.500	5.444	4.118

The positive eigenvalues ("VP") confirm positive definiteness of the matrix. To obtain nuggets z_i, $i = 1$ to 5, with a covariance matrix equal to the above (divided by 10), we then perform the following transformations to a set of five independent standard-normal nugget simulations y_j, $j = 1$ to 5:

$$\begin{aligned} z_1 &= 0.3162\,[\ 0.392y_1+0.265y_2+1.710y_3+1.360y_4-1.000y_5\ ] \\ z_2 &= 0.3162\,[\ 2.060y_1-0.839y_2+0.855y_3+0.071y_4+1.140y_5\ ] \\ z_3 &= 0.3162\,[-0.115y_1+1.440y_2-0.855y_3+1.520y_4+0.934y_5\ ] \\ z_4 &= 0.3162\,[-1.040y_1+1.320y_2+1.710y_3-0.826y_4+0.745y_5\ ] \\ z_5 &= 0.3162\,[\ 1.760y_1+1.800y_2-0.428y_3-0.774y_4-0.614y_5\ ] \end{aligned}$$

where $0.3162 = \sqrt{(1/10)}$.

the simulated datum at that grid location is attributed to the original location. For example, a real data location of (603,297) would receive the simulated data located at (600,295). As the data are more widely spaced than 5x5 meters and kriging is done on a 10x10-meter grid, this approximation should be acceptable.

A two-dimensional ordinary-kriging routine for irregularly spaced data (Stanford subroutine OKB2D) is used to obtain kriged values on a 30x30 grid with 10x10-meter spacings. These spacings correspond to simulated square blocks comparable in size to the average size of the real, roughly rectangular blocks. The kriging program is instructed not to krige grid locations for which fewer than four conditioning data exist within a 20-meter radius. Only 206 of the 900 grid locations examined satisfy this requirement, effectively limiting the conditional simulation to the vicinity of heavily sampled areas, as illustrated in Figure 66. This limitation of the simulation domain prevents domain errors (Section 3.7.2) from becoming too large.

The program used to perform the kriging does not provide explicitly for kriging with circular variograms. However, within the relatively short 20-meter search radius being used in this case, the circular and spherical models have nearly linear forms, differing only in slope (Figure 13), so appropriately adjusted spherical models are used as approximations to the circular models in Table 4.

Ordinary kriging is used to condition each simulated principal component independently. As the components are only weakly cross correlated and the simulation is confined to well conditioned areas, ordinary kriging can be expected to differ little from cokriging.

Furthermore, much of the cross correlation embodied in the real data is restored when the simulated principal components are inverted to obtain simulated cross-correlated normal scores, as explained in the next section.

4.1.3.5 Inversion of Transformations: Steps 20-25

The principal-component loadings furnished by program BMDP4M in Step 7 can now be used to transform the conditionally simulated principal-component scores back into simulated normal scores. The procedure is the same one used in Step 13 to obtain a linear model of coregionalization for the unconditionally simulated nugget structures. Details of this back transformation are provided in Table 6. Notice that the first two principal components, which explain 90.7% of the total variance in the normal-scores data, essentially correspond to the two principal geological sources of variation described in Section 4.1.1. The first component ("Factor 1" in Table 6) contributes heavily to SiO_2, Al_2O_3, Fe_2O_3, and (negatively) CaO, thus corresponding to the influence of the proportion of carbonates to silicates in the rock. The second component contributes heavily to MgO, and also negatively to CaO, thus corresponding to the influence of the proportion of dolomite to calcite within the carbonates.

The conditionally simulated normal scores are inverted to the corresponding residuals from linear trends by linear interpolation, as described in Section 3.5.2.2, using a simple program similar to Verly's (1984b) subroutine LINT. A great deal of subjectivity creeps in at this point, because some of the simulated normal scores lie outside the range

TABLE 6

A linear model to transform conditionally simulated principal-component scores back into normal scores.

Program BMDP4M obtained the following principal-component loadings and eigenvalues from the calculated covariance matrix of the 129 normal-score data:

Output from BMDP4M:

UNROTATED FACTOR LOADINGS (PATTERN)

FOR PRINCIPAL COMPONENTS

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
SIO2	1	0.931	-0.530D-01	-0.234	0.182	-0.156
AL2O3	2	0.948	-0.475D-01	0.767D-01	0.148	0.227
FE2O3	3	0.909	0.873D-01	0.354	-0.588D-01	-0.125
CAO	4	-0.880	-0.320	0.193	0.256	-0.484D-01
MGO	5	-0.282	0.943	0.235D-01	0.110	-0.207D-02
	VP	3.445	1.004	0.224	0.136	0.094

Thus the simulated normal scores z_i, i = 1 to 5, can be obtained from the components y_j, j = 1 to 5, as follows:

Normal score

$$SiO_2 = z_1 = 0.9310y_1 - 0.0530y_2 - 0.2340y_3 + 0.1820y_4 - 0.1560y_5$$
$$Al_2O_3 = z_2 = 0.9480y_1 - 0.0475y_2 + 0.0767y_3 + 0.1480y_4 + 0.2270y_5$$
$$Fe_2O_3 = z_3 = 0.9090y_1 + 0.0873y_2 + 0.3540y_3 - 0.0588y_4 - 0.1250y_5$$
$$CaO = z_4 = -0.8800y_1 - 0.3200y_2 + 0.1930y_3 + 0.2560y_4 - 0.0484y_5$$
$$MgO = z_5 = -0.2820y_1 + 0.9430y_2 + 0.0235y_3 + 0.1100y_4 - 0.0021y_5$$

of the real normal scores. In this case one must furnish additional extreme "real" values to correspond to the extreme simulated scores. In practice, one might choose to transform all out-of-bounds simulated

scores to the most extreme of the observed real values -- but this is a too-easy solution that tends to produce awkward spikes of simulated values at the extremes of the simulated sample distributions. A better approach is to examine the histograms of the real data and try to imagine where their extremes would be if more data were available. In some cases the appropriate extremes are dictated by natural constraints on the data; e.g., minimum percentage data should never fall below 0%. But in this case study the back-transformed normal scores are not original percentages but residuals from linear trends, so natural constraints on the original data values cannot be enforced in the simulation unless the data locations as well as their values are taken into account in the transformation. Nothing so fancy is attempted in this case study, with the result that, among the 206 final simulated SiO_2 values obtained after the linear trend is restored, four values in the southern part of the simulation domain end up with slightly negative SiO_2 percentages, with the lowest being -0.71%. If this simulation were being used to solve a real mining problem, such a violation would be unforgivable, so either a second try at the back transformation of the SiO_2 normal scores would be necessary or some final fixup would have to be made to the low-SiO_2 values in the southern part of the domain. This can be viewed as an important practical drawback to simulating residuals from deterministic trends.

4.1.4 Results

Figures 68 to 71 show histograms, selected scattergrams, and selected directional semivariograms for the original and simulated data. Statistics are compared in Table 7. Because the linear trends obscure some of the subtle local features of the semivariograms, Figure 72 illustrates the semivariograms of the residuals for the same variables plotted in Figure 71.

It is clear that the major features of the data have been reproduced in the simulation. Even the hypothesized linear constraint of Figure 68 seems as well obeyed by the simulation as by the original data, even though it has not been taken into account in the transformations. Presumably reproduction would have been better if the spatial domains of the two data sets had coincided more exactly, and if the real data points had all coincided exactly with simulated points. Some trial-and-error work with the back transformation of the normal scores might also have helped, especially in the reproduction of the SiO_2 minimum. (It would not be a good idea to perform the normal-scores transformation before removing the trend from the data, as then the detrended normal scores would no longer be normal, and the constraints finally reproduced by the simulation would probably not be valid if the simulation domain were at all different from the domain of the data. The whole problem here is that histograms of nonstationary data do not mean much outside the domain of the data.)

From inspection of Table 7 and plots such as those in Figures 68-72, it appears that the behavior of MgO is less well reproduced than that of the other variables (e.g., SiO_2, shown in all of the figures). This is

TABLE 7

Statistical summaries of the Plymouth limestone raw data and of the conditional simulation based upon the data.

(a) 129 ORIGINAL DATA

VARIABLE	MEAN	STD. DEV.	MIN.	MAX.	CORRELATION WITH SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO
SiO_2	4.00	2.22	0.33	10.07	1.00				
Al_2O_3	1.43	0.54	0.45	2.98	0.94	1.00			
Fe_2O_3	0.57	0.20	0.22	1.07	0.91	0.95	1.00		
CaO	51.07	1.83	46.22	54.39	-0.91	-0.88	-0.85	1.00	
MgO	1.40	0.62	0.57	4.96	-0.54	-0.53	-0.47	0.28	1.00

(b) 206 CONDITIONALLY SIMULATED DATA

VARIABLE	MEAN	STD. DEV.	MIN.	MAX.	CORRELATION WITH SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO
SiO_2	3.97	2.22	-0.71	10.21	1.00				
Al_2O_3	1.40	0.56	0.20	2.59	0.91	1.00			
Fe_2O_3	0.55	0.20	0.12	1.00	0.87	0.93	1.00		
CaO	51.23	1.85	46.74	55.23	-0.88	-0.87	-0.84	1.00	
MgO	1.33	0.56	0.27	3.72	-0.58	-0.58	-0.51	0.36	1.00

particularly apparent in the departures between the east-west semivariograms of real and simulated MgO shown in Figures 71 and 72. This problem apparently arose mostly at the conditioning stage, where some unusually high-MgO data at the southern edge of the domain did not achieve much influence in the kriging.

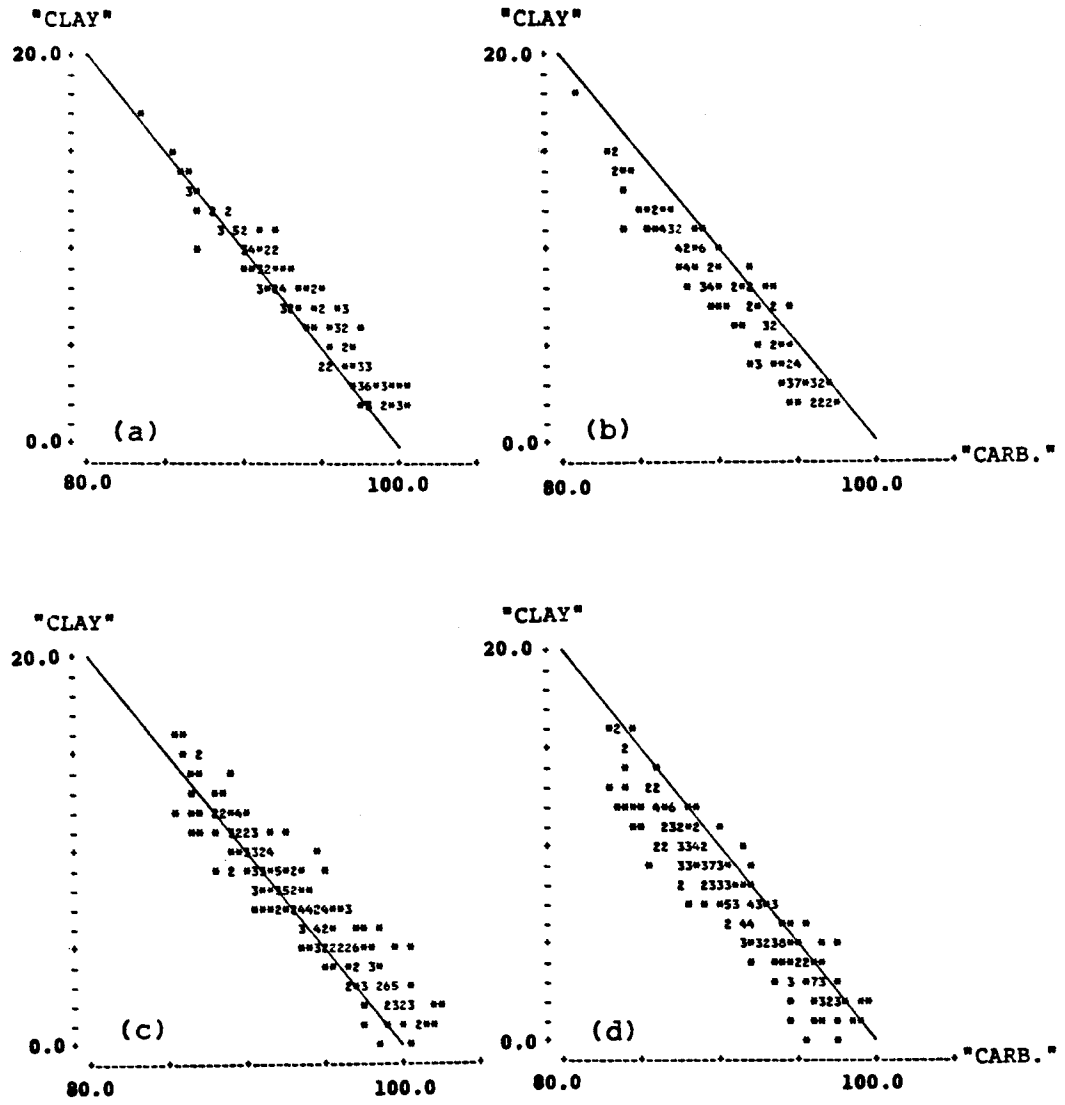


Figure 68: Scattergrams of normative "CLAY" versus "CARBONATE" in the original limestone data and in the simulated data. Some of the coefficients in the transformations used in (a) have been slightly altered, as shown in (b), to obtain a better fit of the original data to the 100%-maximum constraint (straight line). Corresponding plots of the simulated data are shown in (c) and (d).

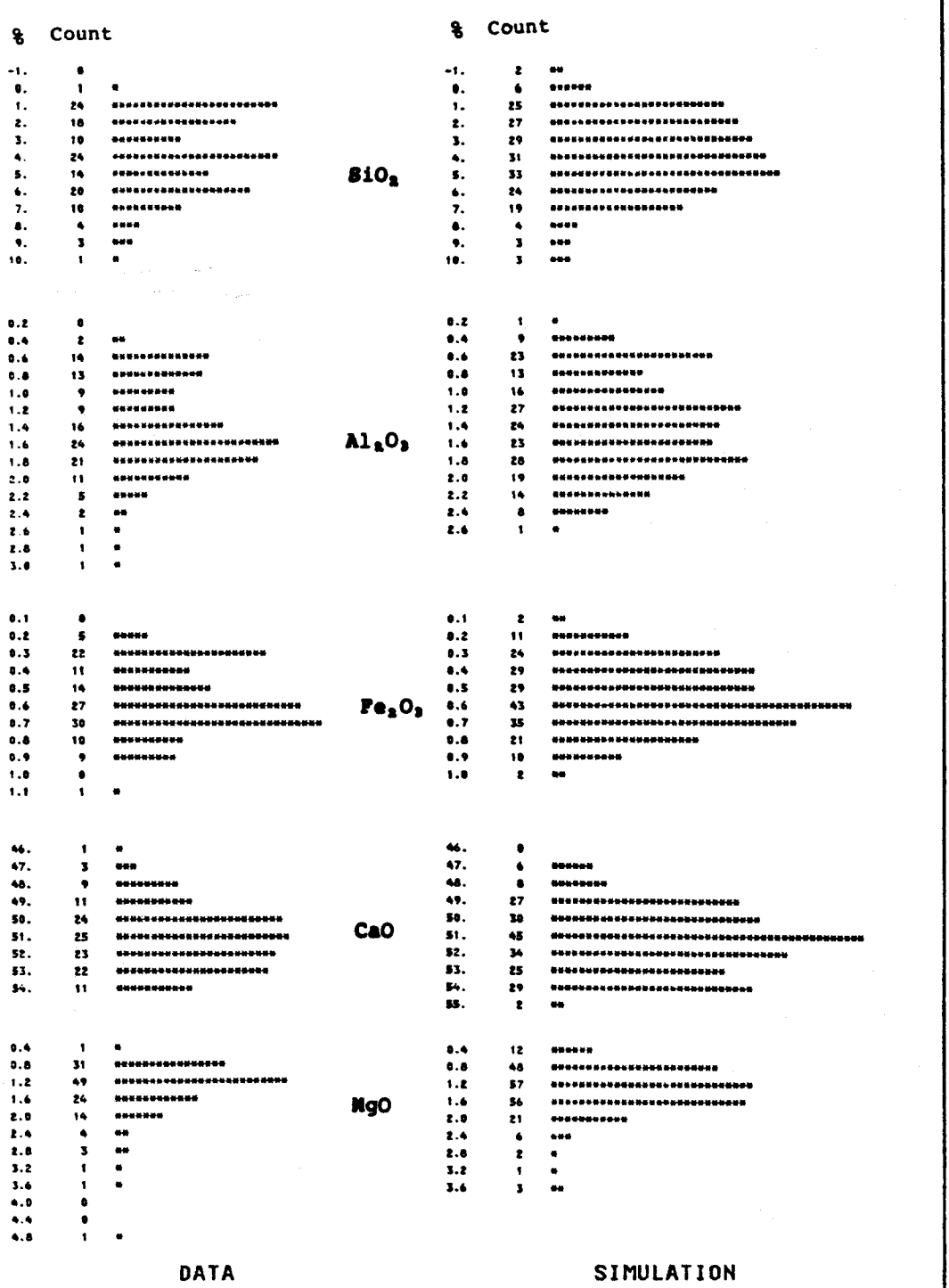


Figure 69: Histograms of the original 129 data and of the final 206 simulated data.

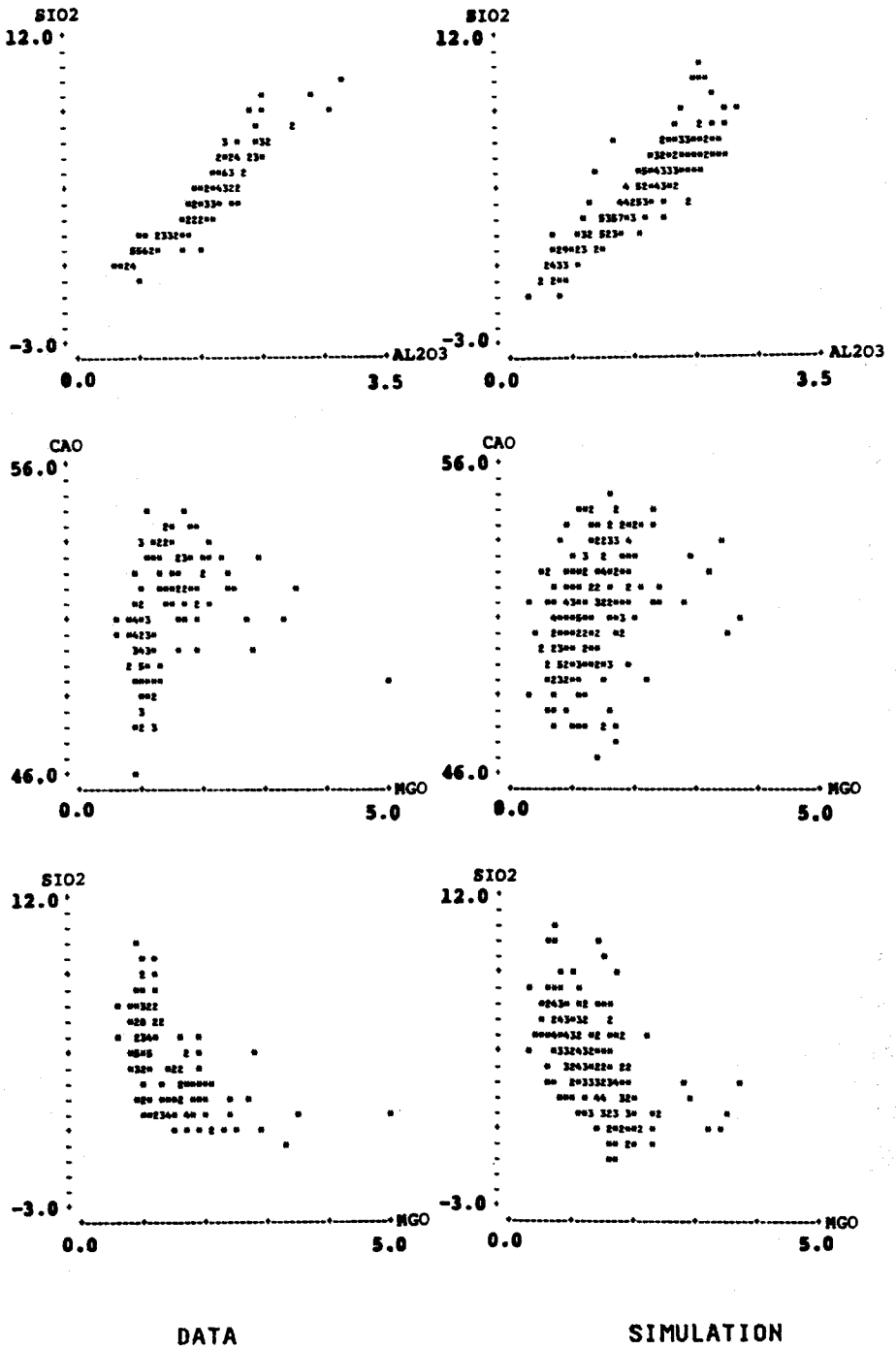


Figure 70: Selected scatterplots of the original 129 data and of the final 206 simulated data.

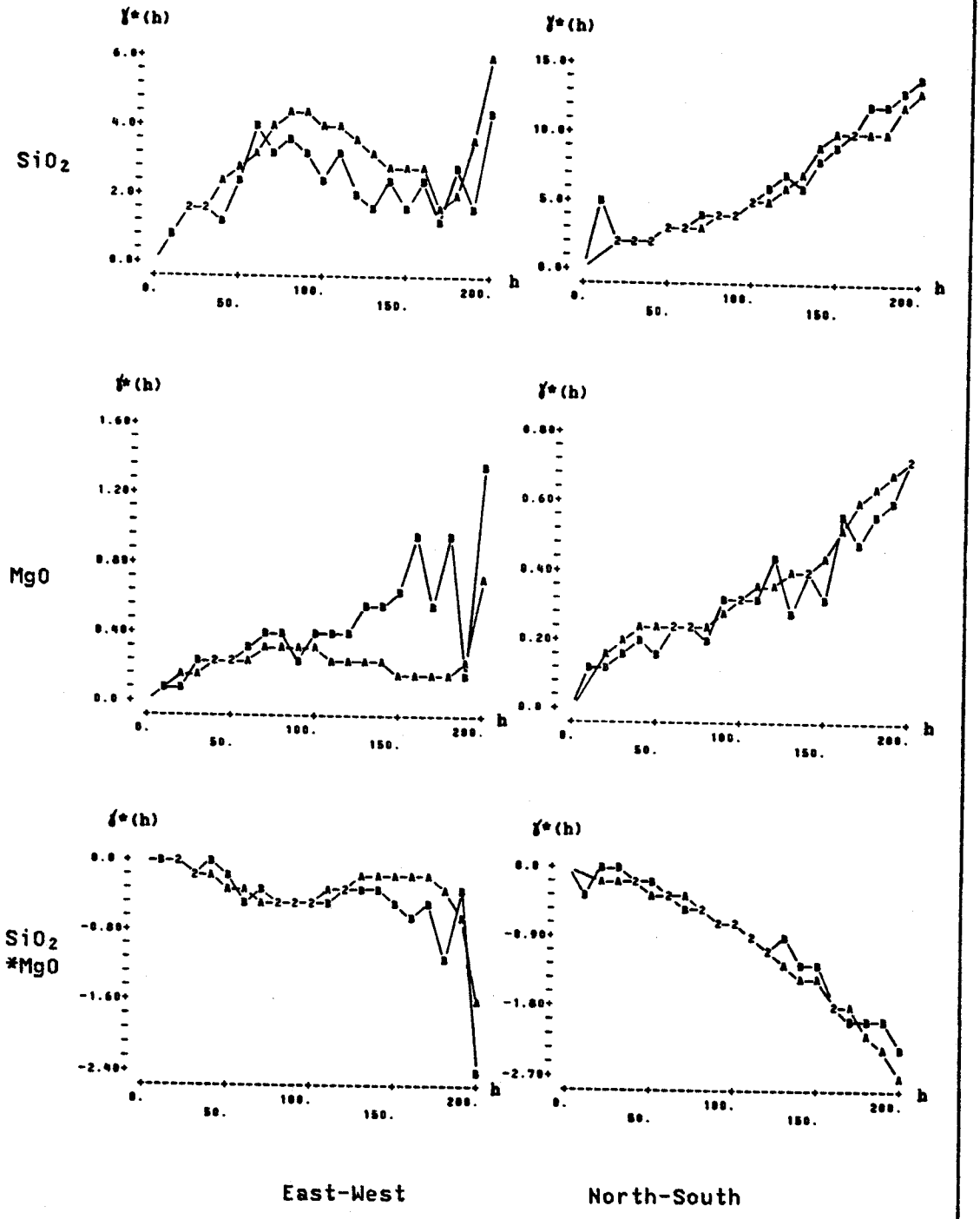


Figure 71: Direct and cross semivariograms of the original 129 SiO_2 and MgO data and of the final 206 simulated data. A=simulation, B=data, 2=both. Notice that the vertical scales vary.

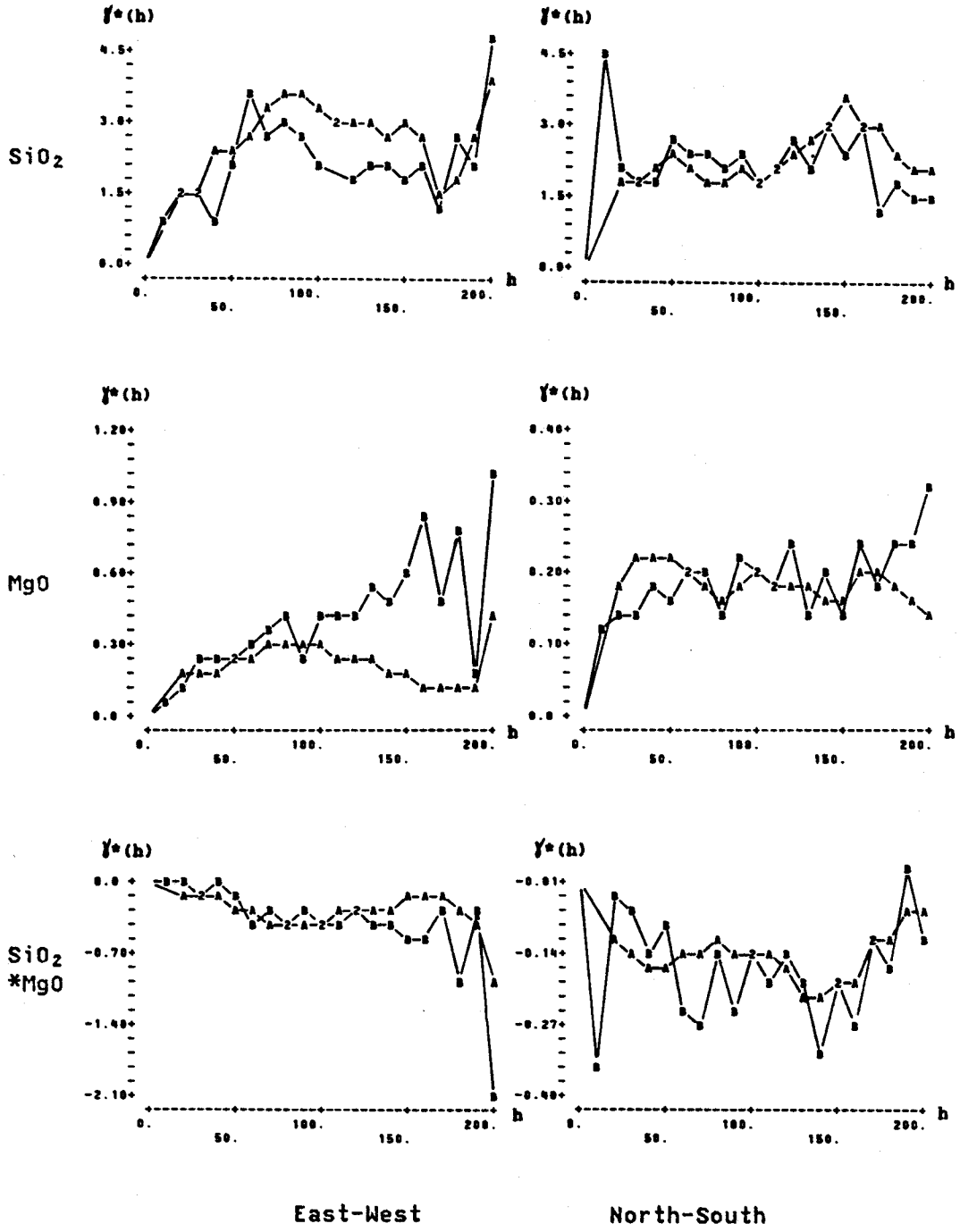


Figure 72: Direct and cross semivariograms of the original 129 SiO_2 and MgO residuals and of the 206 simulated residuals. A=simulation, B=data, 2=both. Notice that the vertical scales vary.

One should always confirm that any data falling exactly on the simulation grid have been reproduced, for all practical purposes, exactly. In this case study, only one of the original 129 data points fell exactly on one of the 206 10x10-meter block centers. A comparison of the real and simulated values at this point is provided below:

Coordinates: x=535, y=285 meters

	SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO
SIMULATION	5.95024	1.84027	0.74001	49.71967	0.80993
DATA	5.95	1.84	0.74	49.72	0.81
DIFFERENCE	0.00024	0.00027	0.00001	-0.00034	-0.00007

4.2 FORECASTING RAW-MATERIAL VARIABILITY FOR A CEMENT PLANT

The case study presented here is much more complicated than the one in Section 4.1. In fact, it has deliberately been made more elaborate than necessary for the practical problem to which it is applied, so that a wide variety of potentially useful techniques can be illustrated. A summary of ways in which the study might have been completed more easily and quickly in a "production" situation is included in Section 4.2.6.

4.2.1 The Problem

This case study involves the simulation of a limestone-dolostone deposit in north-central Iowa and the use of this simulation to choose among alternative mining and mineral-processing systems for a cement plant. For several years the Lehigh Portland Cement Company has operated an open-pit limestone quarry near Mason City, Iowa, primarily in the Devonian Cedar Valley Limestone (Coralville Member), but locally including a small amount of the overlying Shell Rock Formation (Nora

Member). The bedrock stratigraphic section in the quarry, described by Koch (1970, pp. 38-39 and 94-95) and in Anderson (1984), is typically about fifty feet[99] thick and is overlain by Pleistocene outwash and till. The deposit consists mostly of fine-grained, light-grey, high-calcium limestones (averaging about 96% $CaCO_3$, 1% $MgCO_3$), interbedded with coarser-grained, brownish-grey dolostone beds (much more variable in composition, and averaging about 60% $CaCO_3$, 34% $MgCO_3$). The dolostone beds, which appear as irregular bodies interfingering with the limestones, are concentrated mostly in the upper third of the section. The rocks just beneath the quarry floor also are dolomitic.

The limestones dip gently southwestward on a regional scale, although within the quarry area the regional dip is obscured by several local warps in the structure. Much of the bedrock surface in the quarry area appears to lie at nearly the same stratigraphic position, such that the bedrock topography tends to follow the structure. The bedrock surface may thus represent a glacially stripped contact, perhaps the unconformity separating the Cedar Valley and Shell Rock carbonates from the overlying shale of the Lime Creek Formation (Juniper Hill Member).[100]

The high $MgCO_3$ content of the dolostone makes it unsuitable as a cement raw material, except in small, controlled quantities.[101] Areas of

---

[99] English units are used throughout this case study, as all data furnished by Lehigh were recorded that way.

[100] Koch (1970) considers the Shell Rock to be only four feet thick in the southern part of the quarry and missing in the northern part. Regionally, the Lime Creek Formation appears to rest on an angular unconformity above the limestones.

[101] The maximum MgO content of a cement is 6% (Table 1).

the deposit where dolostone appears are thus more costly to quarry, because either the upper dolomitic "caprock" must be completely stripped and wasted, or more complicated mining procedures must be employed -- possibly including small-scale selective quarrying and blending of materials from several quarry faces to reduce fluctuations in $MgCO_3$. For most of the quarry's history, these problems have been avoided by simply "high-grading" the deposit on a large scale, i.e., by limiting mining operations to broad areas of the property where there is little or no dolostone in the section. However, these high-grade areas will eventually be depleted; then the plant will be forced either to strip and waste all of the dolomitic caprock present in the remaining areas (shortening the life of the quarry), or to devise methods to use as much of the dolostone as possible. If some dolostone is used, the variable thicknesses and compositions of the dolostone beds could cause quality-control problems in the plant, unless new procedures are adopted to control the amount of $MgCO_3$ in the stone delivered to the plant. The new procedures might include small-scale selective mining or blending within the quarry, or installation of a stacker-reclaimer system in the plant to smooth out limestone variability and possibly to blend higher- and lower-grade stone from different parts of the quarry. (An approach that appears to work well, based on simulation results, is suggested in Section 4.2.5.3.)

Because the quarry has historically operated in a high-grading mode, there has been no sustained experience in dealing with large fluctuations in the $MgCO_3$ content of the limestone. In fact much of the variability in $MgCO_3$ that the plant has faced in the past has been

attributed to variable amounts of dolostone within the shale of the Lime Creek Formation, which is also used as a cement raw material. Thus the choice of mining procedures, as well as the choice of homogenization and proportioning facilities that may be needed to deal with more variable limestone compositions in the future, must be made without the aid of relevant historical data on limestone variability. In such a "dataless" situation, managers and plant designers have traditionally relied on their own (sometimes mistaken) judgement to anticipate future problems, and relied even more on the inclusion of liberal "safety factors", sometimes at great cost, to enable the plant to accommodate worst-case situations.

Fortunately, completely subjective judgements will not be necessary in the present case. The entire quarry area has been sampled by an unusually dense (by cement-industry standards) grid of core-drill holes -- an excellent data set for use in conditional simulation. The remainder of this chapter describes a complete, full-scale simulation study directed toward a better solution to this problem. Qualitative and statistical descriptions and models of the data, as well as descriptions of numerous adjustments and transformations that had to be made, are provided in Section 4.2.2. Steps in the conditional simulation of $CaCO_3$ and $MgCO_3$ for the partly dolomitized western part of the quarry property (a likely area for quarry expansion) are described in Section 4.2.3, and the results are summarized in Section 4.2.4. Simulations of alternative mining procedures applied to the simulated raw materials are described in Section 4.2.5.1, and homogenization alternatives in Section 4.2.5.2. Section 4.2.5.3 presents the results

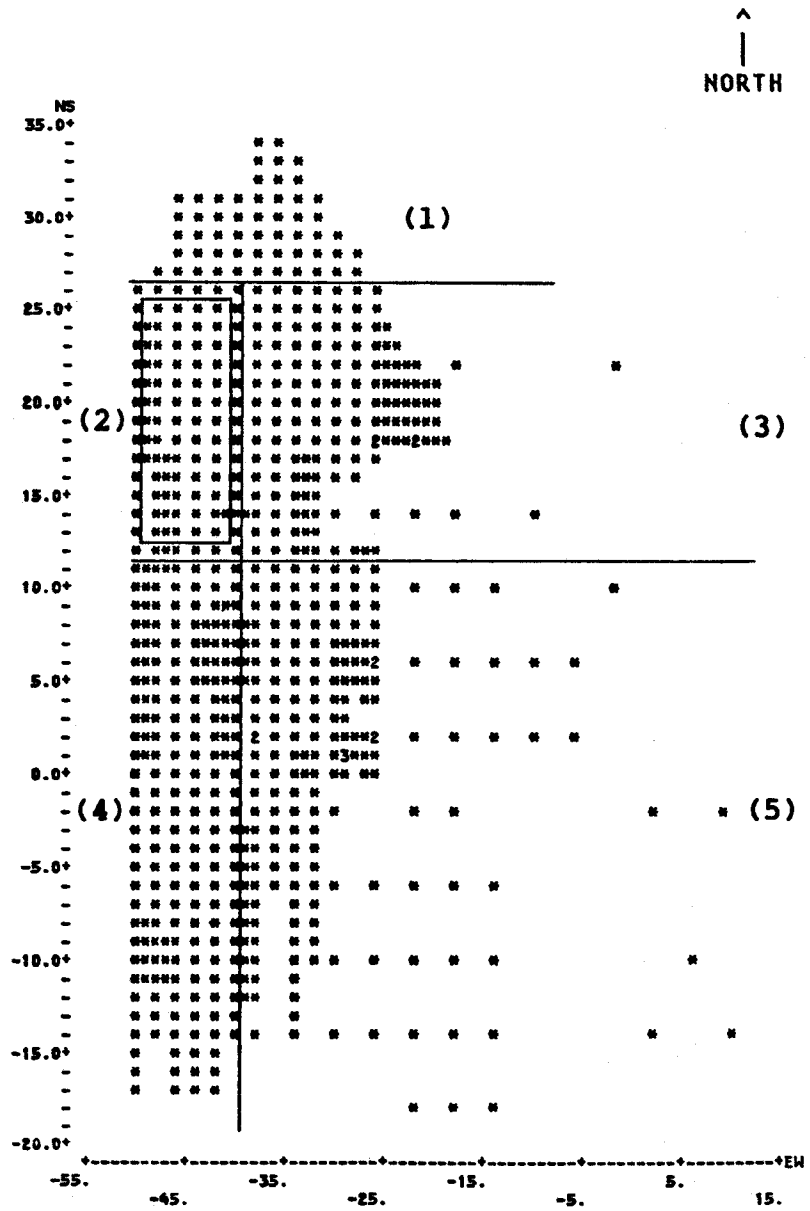
and discusses how these results can be used to select appropriate mining methods and homogenization facilities. Finally, Section 4.2.6 discusses how this complicated simulation might have been simplified with little loss of applicability.

4.2.2 Data Analysis and Structural Analysis

4.2.2.1 Description of the Data

The quarry property has been investigated by four drilling programs, concluded in 1953, 1955, 1960, and 1967. The 1955 and 1967 data sets are much more extensive than the others. The holes from these two programs were drilled on a nearly regular grid covering nearly all of the quarry area (Figure 73). Each core sample was analyzed for $CaCO_3$ and $MgCO_3$, and geologic descriptions of the core were recorded. Geologic descriptions are not available for the 1953 and 1960 drilling programs, and the $MgCO_3$ contents of the limestone samples collected from those holes seem high, suggesting that the limestone and dolostone beds were not well separated during sampling. Because data from the 1953 and 1960 programs are incomplete, possibly less reliable than data from the 1955 and 1967 programs, and distributed thinly over the same area covered by the 1955 and 1967 programs, they were excluded from the final data set used in this case study.

As an example of the types of data available for this study, the geologic section in hole J193, drilled on the quarry property in 1967, is depicted in Figure 74, and the associated entry in the original data file compiled for the case study is shown in Figure 75. The data available from each hole include surface elevation, depth of overburden,



SCALE: HUNDREDS OF FEET

Figure 73: Plot showing the locations of core-drill holes used in this case study. "*" = one hole, "numeral" = number of holes close together. The five numbered regions are the five regions discussed in Section 4.2.2.5. Region 2 is the domain of the unconditional simulation of Section 4.2.3. The inner rectangle in Region 2 is the domain of the conditional simulation.

depths to the top and bottom of each sample, depths to the tops and bottoms of lithologically distinct beds encountered in the core, the total depth of the hole, the $CaCO_3$, $MgCO_3$, and core recovery for each sample, and descriptions of the color and any outstanding features of each sample. (The core-recovery data are not important for this study, although they may be extremely valuable in other situations, as described in Section 3.10.2.) Depths were recorded originally in feet and inches, but converted to feet and tenths for this study. The surface elevations for most holes were recorded to the nearest foot. All data were recorded originally on handwritten log sheets, most of which were subsequently typed. In these cases only the typed logs are now available.

At the start of this case study, the data from the logs were typed into computer disk files (as in Figure 75) and carefully inspected for errors. Numerous obvious errors were found in the typed data sheets during data entry, and several more subtle errors were detected by inspection of histograms and scatterplots of the chemical data, by running a computer program to check the continuity of sample footages down the holes, and by inspection of a contour map of surface elevations and line-printer cross sections through the holes (Figure 82). The fact that many errors were found suggests that many others were not, but the density of data within the property at least assures that the influence of the remaining errors will be kept local.

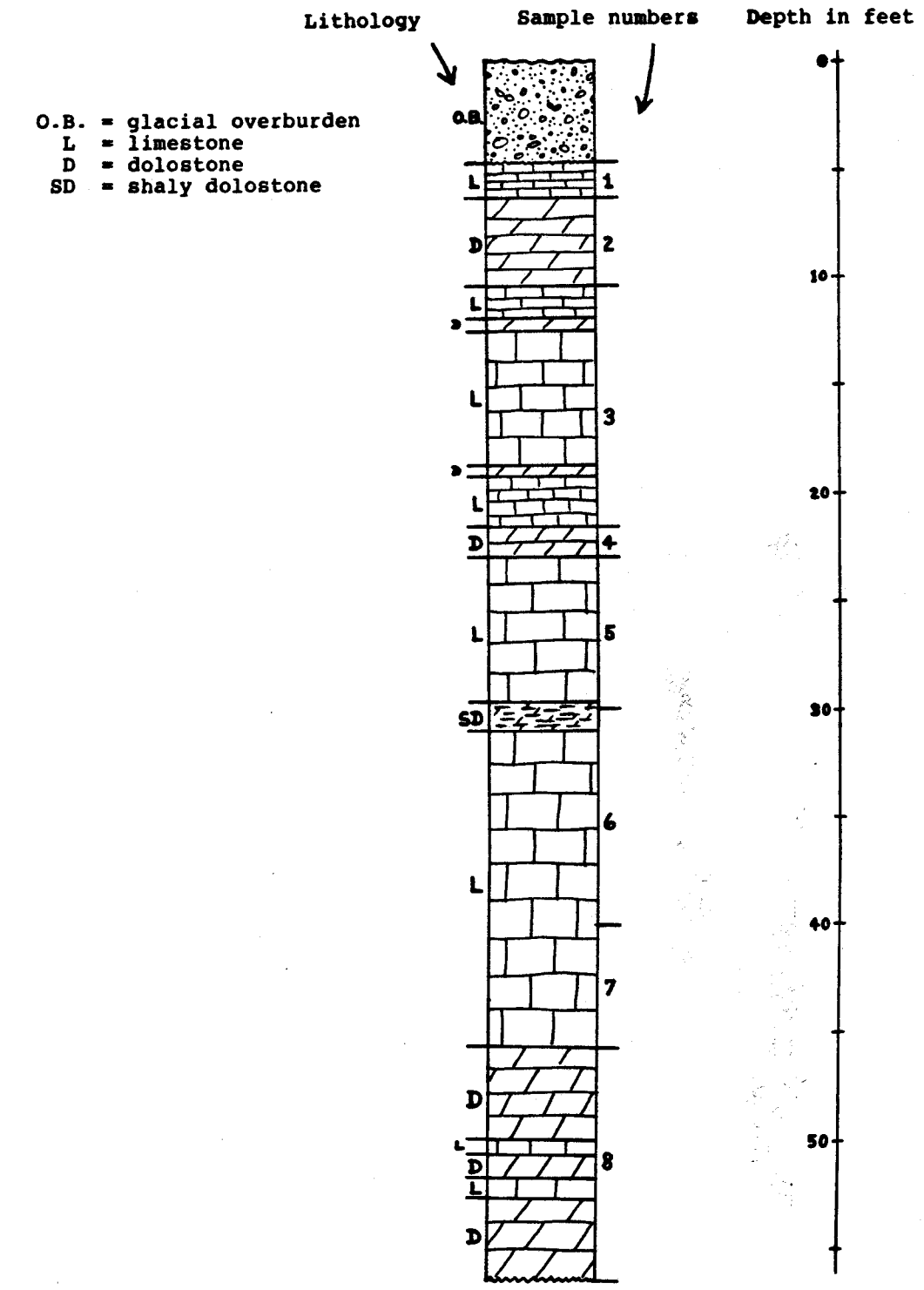


Figure 74: Geologic section in hole J193, showing the intervals of drill core submitted for chemical analysis. The corresponding entry in the data file is shown in Figure 75.

```
J193 IZ 45 1124  56.4   4.7  45.7 08
GWL  4.7   6.3  96.2   0.0
BBD  6.3   10.4  65.0  23.8
GW* 10.4   21.6  89.8   7.7     [DOLOMITE 12-12.5 AND 18.8-19.2
GWL 10.4   12.0
NND 12.0   12.5
GWL 12.5   18.8
NND 18.8   19.2
GWL 19.2   21.6
GBD 21.6   23.0  59.9  36.5
GW@ 23.0   30.0  96.4   1.3     [SEE BELOW
GWL 23.0   29.7
NNX 29.7   30.0
GW@ 30.0   40.0  92.5   3.4     [DOLOMITE AND CLAY 29.7-31.0
NNX 30.0   31.0
GWL 31.0   40.0
GWL 40.0   45.7  95.5   0.4
GB$ 45.7   56.4  66.4  29.9     [LIMESTONE 50-50.7 AND 51.7-52.7
GBD 45.7   50.0
NNL 50.0   50.7
GBD 50.7   51.7
NNL 51.7   52.7
GBD 52.7   56.4
```

Figure 75: Data from hole J193, as entered in the original computer data file. The first line contains the hole number, east-west and north-south coordinates (as cross-section names "IZ" and "45"), surface elevation (nearest foot), total depth of hole, depth of overburden, depth to a fairly continuous "marker" dolostone beneath the quarry floor (used to "dewarp" the local geologic structure), and number of chemically analyzed samples in the hole. The next line contains information on the topmost sample, including color ("GW" = gray-white"), lithology ("L" = "limestone"), depths in feet to the top and bottom of the sample, and the $CaCO_3$ and $MgCO_3$ contents of the sample. The fourth line describes a contaminated sample (indicated by "*", meaning a limestone sample containing a small amount of dolostone), which is broken down in subsequent lines into subsample intervals. These subsamples are eventually assigned "synthetic" chemical compositions, as described in Section 4.2.2.2. In all lines of the data file, the character strings to the right of the numerical data are reserved for geologic descriptions, and originally for subsample intervals, as shown here.

4.2.2.2 Assignment of Chemical Data to Lithologic Subsamples

It is clear from Figure 74 that the support of the chemical data in the vertical direction is variable, and that the support of the lithologic data (i.e., indicators of limestone or dolostone) is smaller than that of the chemical data. In fact, the lithologic support can be made arbitrarily small in the vertical direction, as the contacts between rock types were measured by a geologist in the field, independent of the chemical compositions subsequently determined from the associated samples.[102]

Because the conditional simulation must be performed on a constant support, some conversion of the available chemical data into a constant-support format must be performed. The first question that arises is whether adjacent limestones and dolostones can be averaged together in support adjustments, or whether they are so different and nongradational that they should be simulated separately, and therefore must be identified and separated before the support adjustments are made. The data provided in Figure 75 are typical and provide an immediate answer to this question. The limestones tend to be extremely low in $MgCO_3$ (Figure 76), even where in direct contact with high-$MgCO_3$ dolostones. The $MgCO_3$ contents of the "limestone" samples rise above 2% very rarely, unless thin layers of dolostone have been included in the sample. Dolostone samples, on the other hand, usually exhibit $MgCO_3$ values in the 30's (Figure 77) unless contaminated by limestone. (Samples classified by the geologist as "limey dolomite" or "dolomitic limestone"

---

[102] Actually, a few rock types assigned by the geologist, or possibly typed wrong, have been reassigned for this study because of serious inconsistencies with the chemical data.

are very rare and separate distinctly into two groups; for this study, these samples were considered limestone if the $MgCO_3$ was below 10%.) Furthermore, there is no apparent relationship between the chemical compositions of samples and their thicknesses; e.g., samples of thin limestones encased in dolostone show no consistent tendency toward higher $MgCO_3$, although a few anomalously high values do occur. Thus the limestone and dolostone populations are distinct and almost totally nongradational. Clearly they cannot be realistically modeled as a realization of a single random process, and any averaging of the two populations prior to their simulation should be avoided.[103]

This means that the samples that contain both limestone and dolostone (5.2% of the total footage, plotted in Figure 78) present a serious problem: although we should expect the final simulation of this deposit to be conditioned to the analyses of these samples, they are actually far from representative of either population. Fortunately, the geologic logs of the drill cores provide information on the locations and thicknesses of the contaminating materials, and this information is useful in finding a practical though imperfect solution to the problem.

To approach the contaminated-sample problem rigorously, one would have to condition the simulations of the two rock types in such a way that the average of the analyses of the two conditional simulations, weighted by the thicknesses of the respective rock types, would equal the analysis of the original contaminated sample at each location of contaminated data. One possible approach would be to use a Monte Carlo method entailing no spatial considerations. For instance, for a mixture

---

[103] But see Section 4.2.6.

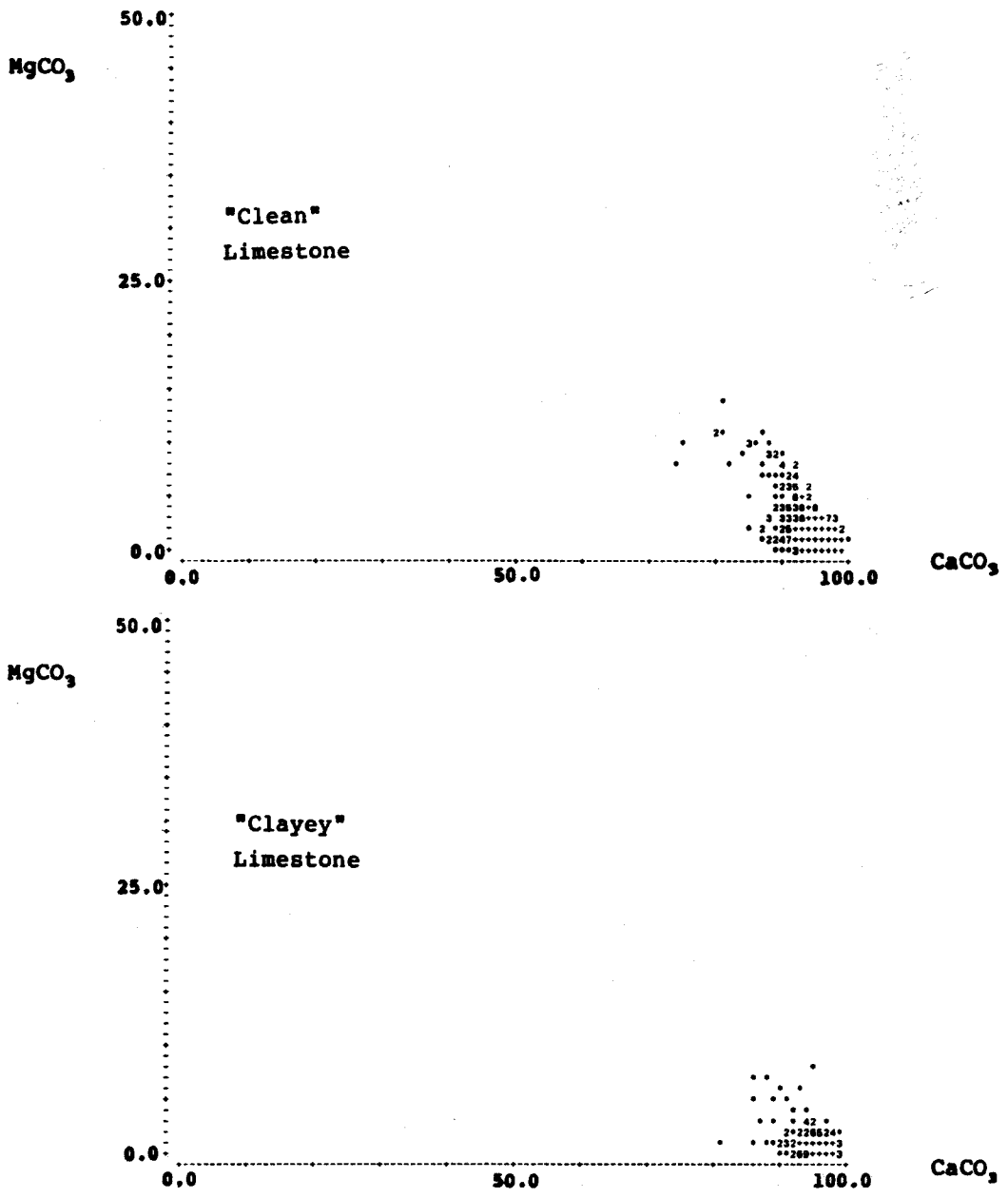


Figure 76: Scatterplots of data from limestone samples uncontaminated by dolostone. Samples containing "clay", as described in the geologic logs, are plotted separately here but appear to be indistinguishable from the "clean" limestones, so they are subsequently treated together. In the scatterplots, "*" = one observation, "numeral" = number of observations, "+" = more than nine observations.

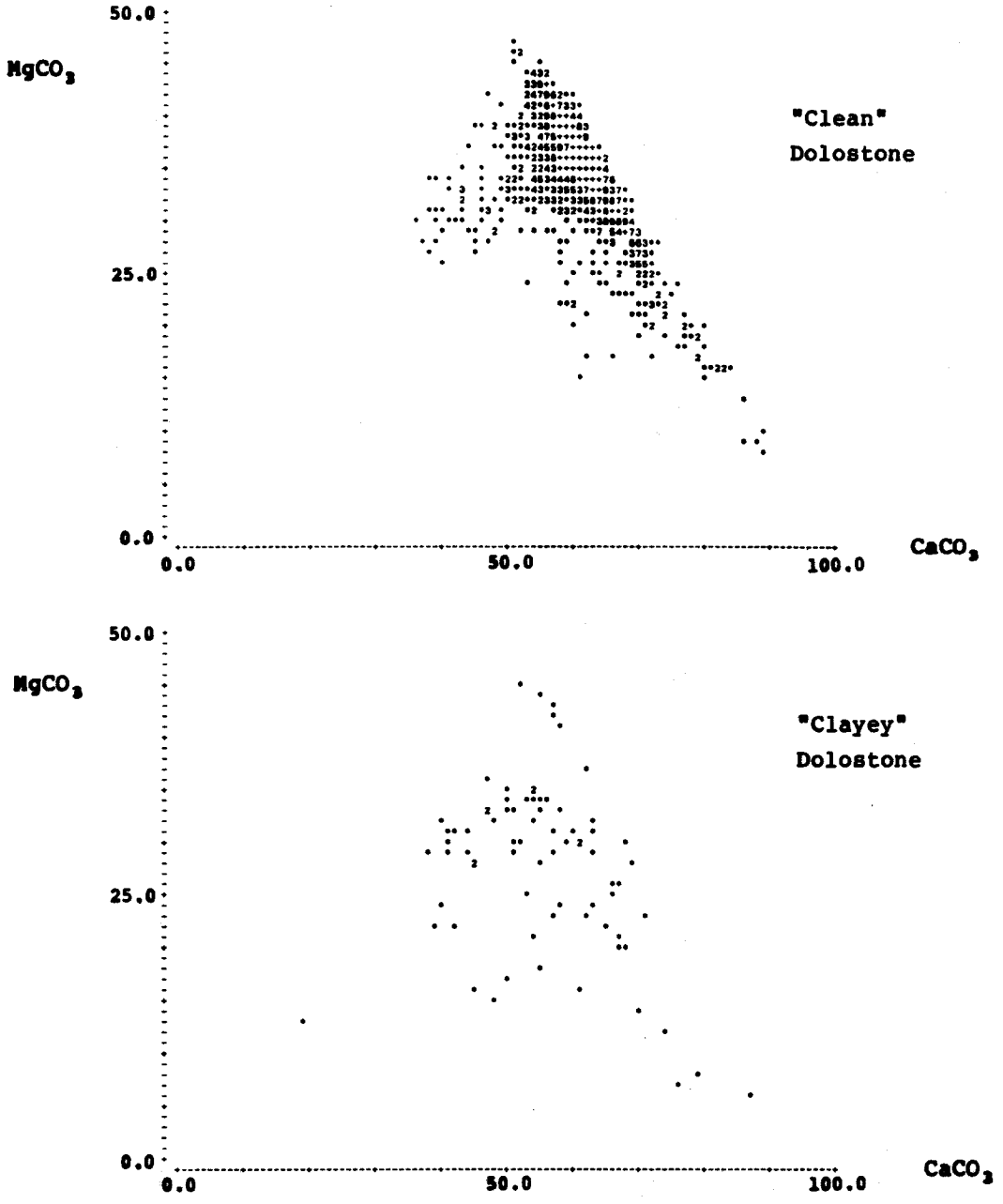


Figure 77: Scatterplots of data from dolostone samples uncontaminated by limestone. Samples identified as "dolomite and clay" in the geologic logs are different from the "clean" dolostones on the average, but it is clear from the scattergrams above that these populations overlap, so they are subsequently treated together.

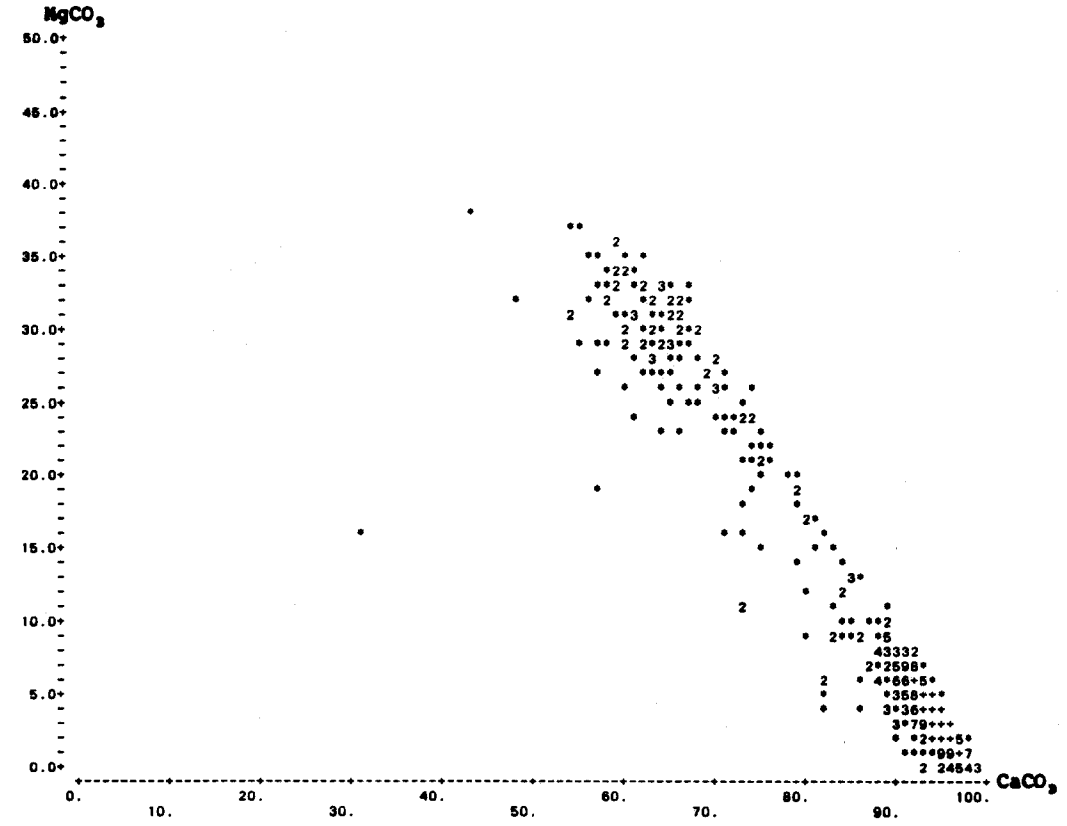


Figure 78: Scatterplot of data from samples containing both limestone and dolostone. Samples described in the geologic logs as containing "clay" are combined with the "clean" samples here. The chemical contents of these samples have been allocated to synthetic subsamples, summarized in Figures 80 and 81.

of limestone and dolostone, one could estimate the four-variate joint probability density function of $CaCO_3$ and $MgCO_3$ in limestone and $CaCO_3$ and $MgCO_3$ in dolostone (assumed independent of limestone) and draw samples from the conditional density defined by the intersection of this density with the intersection of

$$T_l Ca_l + T_d Ca_d = T_t Ca_t$$

$$T_l Mg_l + T_d Mg_d = T_t Mg_t$$

where T=thickness[104], Ca=$CaCO_3$, Mg=$MgCO_3$, and the subscripts (l,d,t) = (limestone, dolostone, total sample). The great disadvantage of this approach is that it would involve a mighty amount of programming for such an exotic application.

A better and simpler solution would be to use a stepwise conditional-simulation approach, first simulating one rock type and conditioning the analyses subject to inequality constraints (Section 3.7.5). The constraints would be the maximum and minimum values that an analysis of a rock type could have without violating the rock's definition and without rendering the known analysis of the total sample impossible. For example, in a sample containing 50% limestone and 50% dolostone and having an analysis of 10% $MgCO_3$, the dolostone certainly could not have over 20% $MgCO_3$ and presumably could not have less than 10% $MgCO_3$ if the dolostone is really more dolomitic than the limestone. Once the conditional simulation of the dolostone is completed, the conditioning values for the associated limestones are fixed. They can then be used to condition the limestone simulation by ordinary kriging.

The data set used for this case study contains so many contaminated analyses that implementation of the stepwise approach above would involve a great number of constraints and be costly and complicated to perform, unless the area being simulated were very localized. To make things easier, a deterministic approximation is used here that ensures that the conditioning values are reasonable but does not take spatial information or randomness of any kind into account in obtaining them. This approach still involves mathematical programming (which probably

---

[104] The limestone and dolostone have approximately the same density, so thickness substitutes for weight.

would be used in an implementation of the previous approach), but in the form of many small problems instead of one very large one.

Suppose that we are given a sample containing four rock types: limestone, clayey limestone, dolostone, and clayey dolostone. In Figure 76, it is shown that clayey limestones are virtually identical to "clean" limestones in this deposit, so they can be safely lumped together. Clayey dolostones do, however, differ on the average from clean dolostones, although they appear to be overlapping variants of a single population (Figure 77). At this stage of the study, the clean and clayey dolostones are treated separately; subsequently they will be combined into a single population. As an approximate solution to the contaminated-sample problem, the analysis of the total sample is partitioned into "synthetic" ("fake", if the reader prefers) analyses of the individual rock types on a sample-by-sample basis (taking no spatially nearby data into account), in the following manner.

First, a mineralogical constraint on dolostone composition (discussed at the beginning of Section 4.2.2.6) is removed by transforming $CaCO_3$ and $MgCO_3$ to stoichiometric "CALCITE" and "DOLOMITE" as follows:

$$DOLOMITE = MgCO_3 / 0.457$$

$$CALCITE = MgCO_3 + CaCO_3 - DOLOMITE$$

Then we can identify several linear constraints that the analyses of the individual rock types should follow. Within any analyzed sample, the compositions of the component rock types should satisfy the two material-balance constraints

$$T_l C_l + T_d C_d + T_c C_c = T_t C_t$$

$$T_l D_l + T_d D_d + T_c D_c = T_t D_t$$

where T=thickness, C=CALCITE, D=DOLOMITE, and the subscripts (l,d,c,t) = (limestone, dolostone, clayey dolostone, total sample). There are also maximum-sum constraints on the analyses of the individual rock types, i.e.,

$$C_l + D_l \leq \text{Max}(100\%, C_t+D_t)$$

$$C_d + D_d \leq \text{Max}(100\%, C_t+D_t)$$

$$C_c + D_c \leq 100\%$$

where the bound Max(100%, C_t+D_t) allows for the fact that scattered limestone and dolostone analyses total to slightly more than 100%. Finally, the rock-type analyses should make comparative lithologic sense; e.g., the calcite content of the limestone should exceed that of the dolostone:

$$C_l - C_d \geq 0$$

And similarly,

$$C_l - C_c \geq 0$$

$$D_d - D_l \geq 0$$

$$D_c - D_l \geq 0$$

$$D_d - D_c \geq 0$$

And finally, we have the nonnegativity[105] constraints:

$$C_l, D_l, C_d, D_d, C_c, D_c \geq 0$$

These constraints still permit an infinite number of solutions for (C_l,D_l,C_d,D_d,C_c,D_c), so some rule must be devised to select one solution from those available. One could devise a complicated sequence of

---

[105] Actually, nonnegativity was violated in one of the samples, because the ratio $MgCO_3/CaCO_3$ in the total sample exceeded the ratio in stoichiometric dolomite, resulting in a negative value of C_t for that sample. This problem was circumvented by temporarily setting C_t=0.

if-then decisions to pick a "geologically reasonable" solution, but it is easier and perhaps just as realistic to pick a single "geologically reasonable" objective function (a "subjective objective function") to be optimized by mathematical programming. The question is what sort of function to pick.

If a reasonable linear objective function can be devised, the problem can be quickly solved by linear programming. For instance, it might be thought geologically reasonable to maximize the calcite in the limestone, the dolomite in the dolostone, and the negative of calcite in the dolomitic clay, i.e.,

$$\text{MAXIMIZE:} \quad X\ C_l + Y\ D_d - Z\ C_c$$

where X, Y, and Z are positive weights subjectively chosen. The problem with this sort of objective is that the solution found will invariably lie at a vertex defined by a set of constraints; i.e., the limestone will tend to become 100% calcite, the dolostone will seek to be 100% dolomite, and the clayey dolostone will tend to be devoid of calcite, and perhaps of dolomite as well if the sample contains clean dolostone. The unfortunate effect of this is illustrated by the plot of dolostone analyses[106] in Figure 79, which contrasts sharply with the plot of real dolostone analyses provided at the top of Figure 77. The influences of the maximum-sum constraint,

$$CaCO_3 + MgCO_3 \leq 100\%$$

and the mineralogical constraint,

$$MgCO_3 / (MgCO_3 + CaCO_3) \leq 0.457$$

---

[106] The linear-programming routine used to obtain these analyses is IMSL subroutine ZX3LP (IMSL, 1982).

are pronounced in the synthetic data but subdued in the real data, most of which represent samples of rocks that are not so mineralogically pure. An attempt was made to get around this problem by averaging the solutions obtained using several objective functions. For Figure 79, six objectives were used, obtained from the six permutations of (1,2,3) for the coefficients (X,Y,Z). In general, different objectives will tend to shift a solution to different vertices of the convex feasible region, and the averages of solutions lying at different vertices should lie at interior (feasible) points of the region rather than on the constraints. However for this problem most solutions appear to be not very sensitive to changes in the objective coefficients, as no noticeable retreat from the constraints appeared unless fairly unrealistic (zero or negative) choices for (X,Y,Z) were included, and these choices introduced other undesirable artifacts into the result.

Ideally, the scatterplots of the synthetic analyses should resemble those of the real analyses of "pure" samples presented in Figures 76 and 77. We should at least ask that the mean values of the synthetic analyses be close to those of the real ones. This suggests a nonlinear objective function, similar to one described on page 24, that minimizes squared deviations from an "aim" analysis. Using the sample mean compositions listed in Figures 76 and 77 as the aim, i.e.,

$$(C_l^*, D_l^*, C_d^*, D_d^*, C_c^*, D_c^*) = (95.1,\ 2.2,\ 20.0,\ 75.1,\ 22.5,\ 60.6)$$

the objective becomes:

$$\text{MINIMIZE: } U(C_l-95.1)^2 + V(D_l-2.2)^2 + W(C_d-20.0)^2 + X(D_d-75.1)^2 + Y(C_c-22.5)^2 + Z(D_c-60.6)^2$$

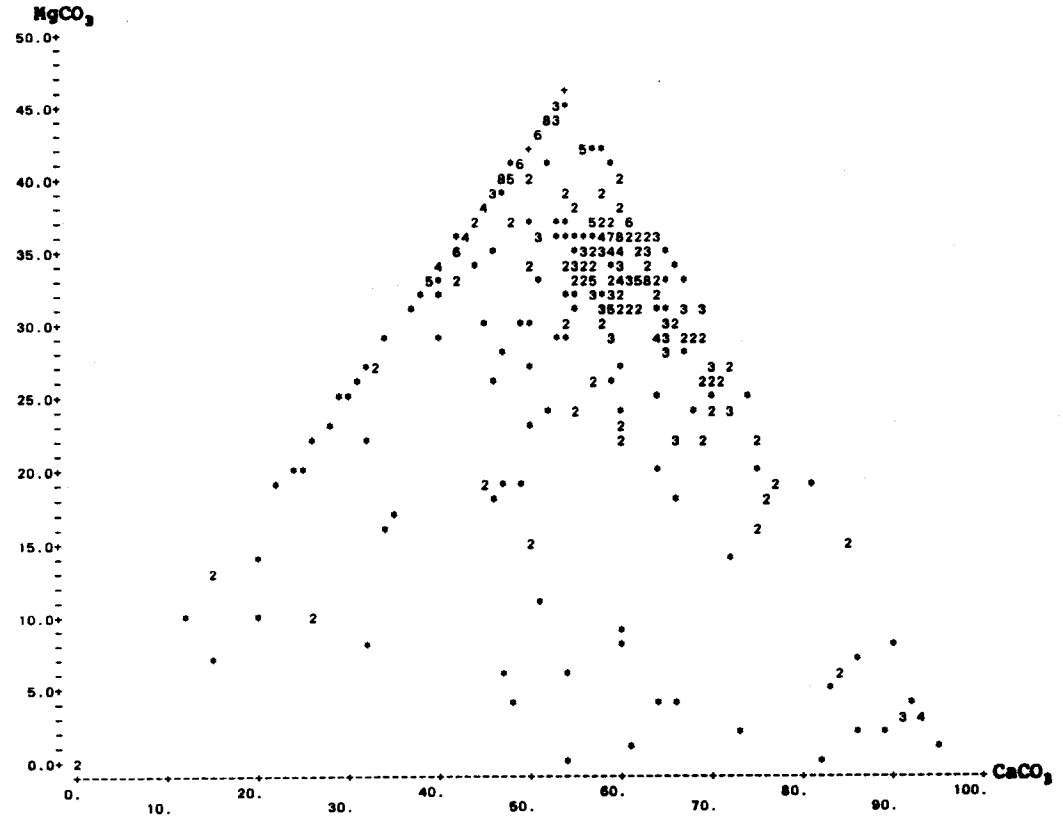


Figure 79: Summary of synthetic dolostone analyses created from samples in Figure 78 by means of linear programming.

We might expect this objective to produce a set of synthetic analyses with approximately the same means as the real analyses but with more peaked distributions, owing to the fact that dispersion around the means is being minimized, subject to the constraints. The choice of weights U through Z is problematical. In this case study two sets of weights were tried: all weights equal to one, and weights equal to the reciprocals of the sample variances calculated from the pure analyses. The reciprocals were used in an attempt to keep the additive terms in the objective function approximately equal in influence.

The results of the two nonlinear-programming runs[107] turned out to be nearly identical, and in most respects not much different from the linear-programming results. This suggests that the linear constraints, which were common to all runs, exerted a considerable influence on the solutions; it further suggests that any reasonable method used to allocate these sample analyses among the component rock types probably would have produced similar results for most samples. Results from the nonlinear-programming run with inverse-variance weighting were narrowly judged to be the best of all runs, and are reproduced in Figures 80 and 81. The advantage of the nonlinear objective function is clear by comparing the results at the top of Figure 81 with Figure 79. This improvement was evident only in the dolostones and clayey dolostones. The importance of the constraints is evident when one notices that many of the more widely scattered points in Figures 79 and 81 are in identical positions, regardless of the objective functions.

The synthetic analyses produced by this method are not designed to account for spatial correlations and in a few cases are not typical of their rock types, even though they satisfy the constraints, so they are not used in the estimation of semivariograms and other characteristics that are described in succeeding sections. They are used only for conditioning the simulations, to assure that the overall compositions of the contaminated samples in the data set are reproduced in the simulations.

---

[107] Stanford subroutine package LCMNA ("Linearly Constrained Modified Newton Algorithm"), using a method described by Gill and Murray (1972), was used in these runs.

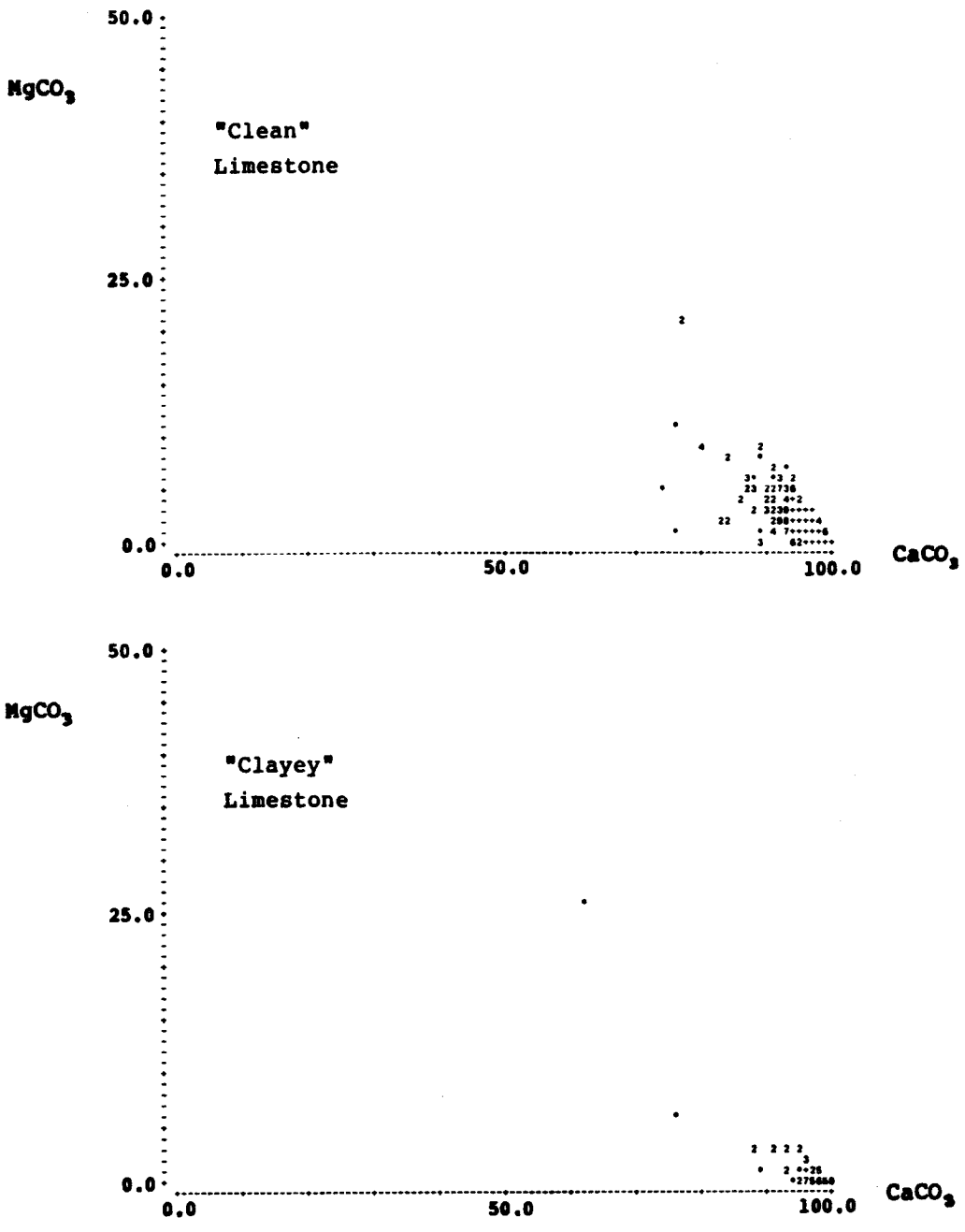


Figure 80: Summary of synthetic limestone and clayey limestone analyses created by nonlinear programming. These two rock types were treated as one by the program but are plotted separately here. Compare these results with the "pure" analyses summarized in Figure 76.

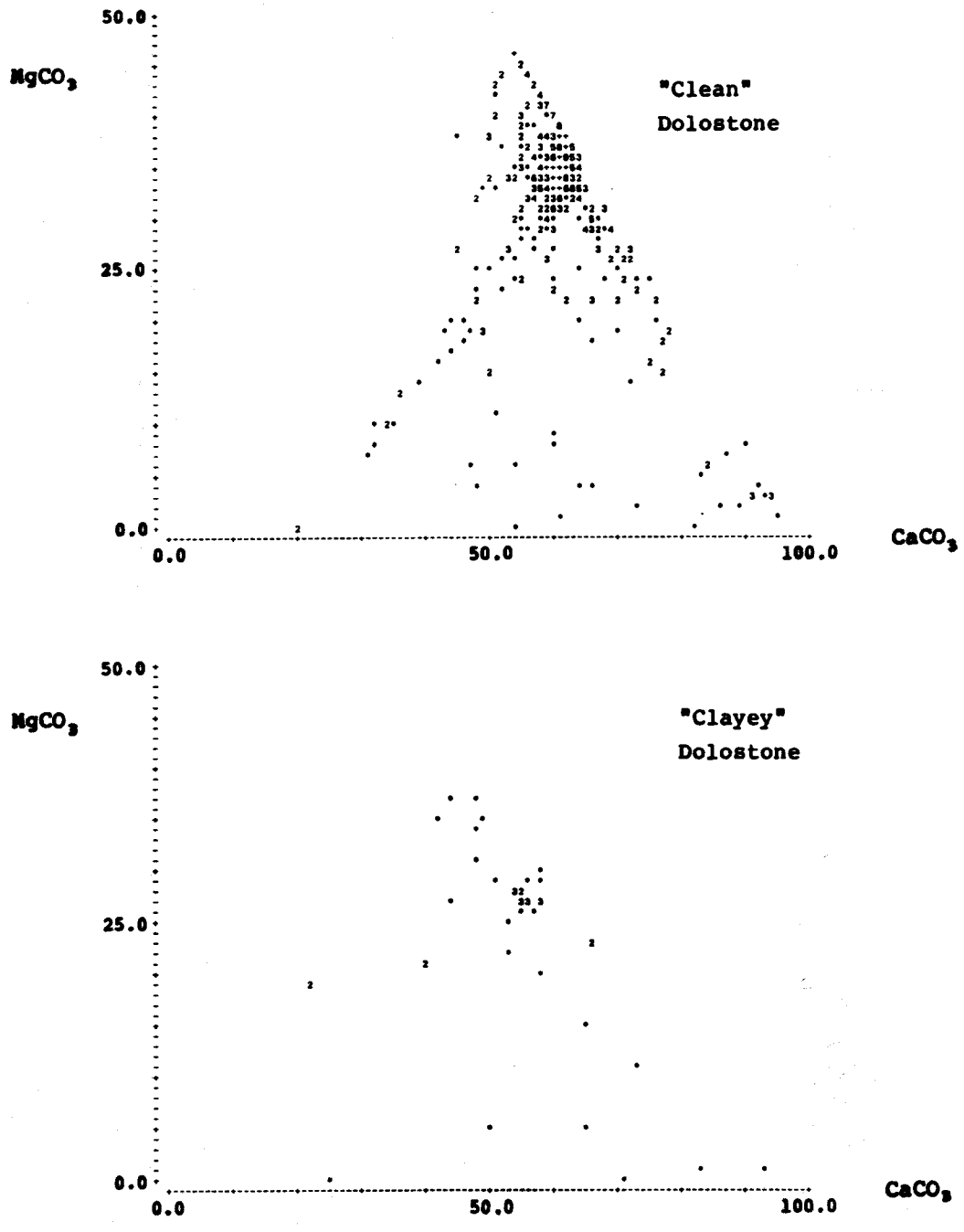


Figure 81: Summary of synthetic dolostone and clayey dolostone analyses created by nonlinear programming. Compare these results with the "pure" analyses summarized in Figure 77.

4.2.2.3 Adjustments for Geologic Structure

Semivariograms obtained from sedimentary deposits typically reveal a strong (usually zonal) anisotropy, with a short range or a rapidly fluctuating drift in the direction vertical to bedding, and much longer ranges in directions parallel to bedding. It is clearly desirable to capture this anisotropy in simulations of sedimentary deposits. However, in folded or faulted deposits bedding orientations may change so rapidly that the geologic structure obscures the variographic structure. In such cases it is necessary somehow to flatten out the geologic structure before computing sample semivariograms.

The geologic structure in the Mason City quarry is dominated by several broad warps in the bedding, commonly several hundred feet across but usually less than twenty feet in amplitude. Nevertheless, these structures are severe enough to obscure the large-scale structure of horizontal sample semivariograms. The short-scale variographic structure may also be complicated by observed channeling, lenticular bedding, and paleocollapse features within the limestone. The drill holes are too widely spaced to reveal most of the local features clearly if at all, but the broad warps are evident in geologic cross sections through the property (e.g., Figure 82(a)). Quarry personnel attempt to recover nearly all of the limestone above the lowermost dolostone bed that is evident in Figure 82, with the result that the quarry floor tends to follow most of the warps in the bedding. Thus both the structural analysis of the variables to be simulated and the eventual mining simulations to be performed on the conditionally simulated

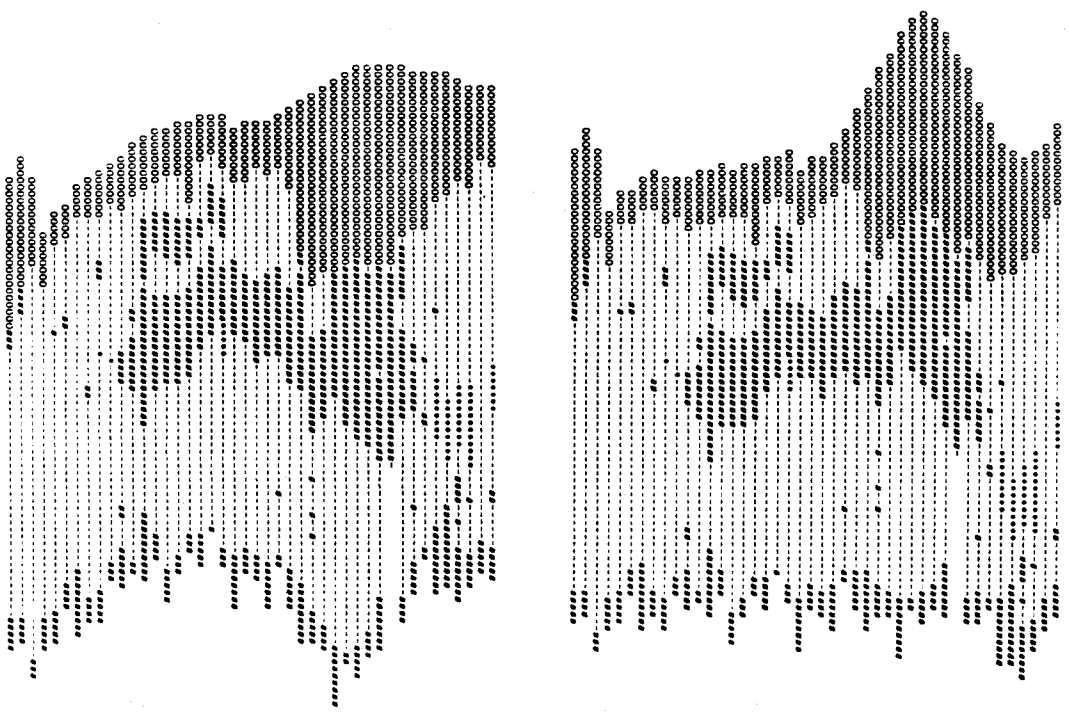
deposit will be made simpler and more accurate if the effects of the warps in the geologic structure are removed from the data set.[108]

Ideally, the effects of broad structural warps can be removed simply by subtracting the elevation of some "datum" surface -- a marker bed in the geologic section -- from all drill-hole elevations. The dolostone bed just beneath the quarry floor is continuous across the entire property, so its top would seem to be a convenient datum to use. Unfortunately, the top is rather irregular in some areas (Figure 82), and in one area the dolostone appears to interfinger with the overlying limestone. Furthermore, a few holes stopped short of the datum. Hence it is not possible to locate this datum with certainty in all holes. Furthermore it is not clear that the local irregularities in the datum elevation extend upward through the whole section, so exact adjustments according to this elevation in each hole might add short-scale variability to the overall structure while subtracting the effects of the broad warps. To get rid of just the large-scale effects, a smoother datum is required.

A smooth datum for this study was obtained by kriging. (Any surface-fitting method might have been used just as well in this application; the kriging software was simply convenient.) Elevations of the top of the lowest dolostone were obtained from drill holes on a 400-foot square grid. (Recall that the property was drilled essentially on a 100-foot grid.) Holes in which the position of the datum was unclear were excluded. A two-dimensional kriging program (Stanford subroutine OKB2D) was used to estimate datum elevations at all other locations, using a smooth power-model semivariogram, $\gamma(r)=r^{3/2}$. The elevations from this

---

[108] But see Section 4.2.6.



(a) Before adjustment (b) After adjustment

NORTH --->

Figure 82: Cross section along the west boundary of the quarry property, before and after adjustments for geologic structure. Each vertical string represents one drill hole within a north(right)-south(left) line of holes. Section (a) shows the original warped structure, and section (b) shows the same structure after removal of the broad trend that is evident in the top of the lowermost dolostone bed in the section. The spacing between holes is 100 feet. Within each hole, the characters represent rock types observed at one-foot intervals down the hole. "0" represents overburden, "-" is limestone, "#" is dolostone, and "*" (present mainly in the lower north end of the section) is clayey dolostone (lumped with dolostone in the subsequent simulations). The line of holes depicted in this section is located along the western (left) edges of Figures 73 and 83.

smooth kriging (Figure 83) were then subtracted from the surface elevations at each hole, which are used to obtain the elevations of all core samples. The resulting "dewarped" structure is exemplified by Figure 82(b).

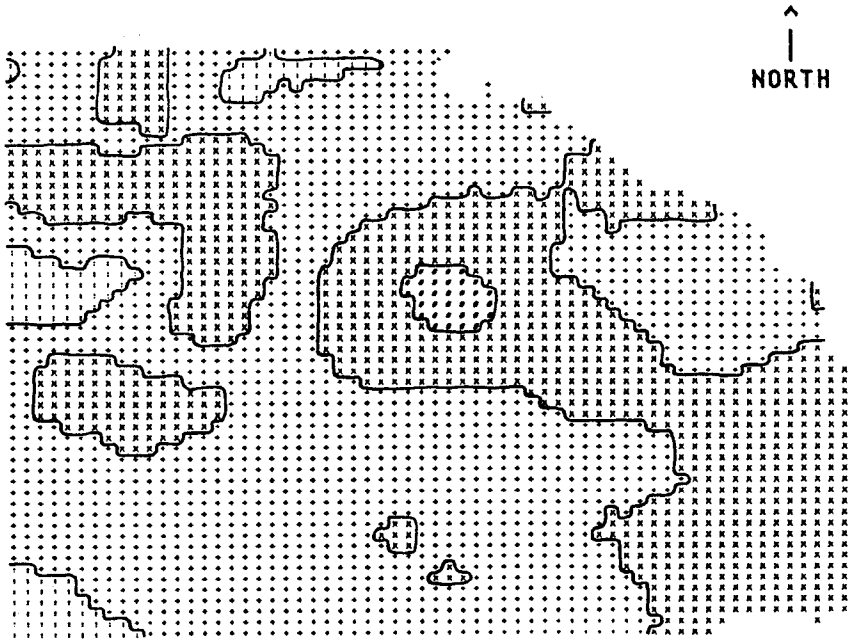


Figure 83: Map of the kriged surface used to adjust hole elevations for geologic structure. The warped structure evident in Figure 82(a) can be seen along the western (left) side of this map. (The map extends a bit farther north than the cross section.) Each character on the map represents a value to be subtracted from the surface elevation at the corresponding 100-foot grid node. ":" = less than 15; "+" = 15 to 25; "X" = 25 to 35; "#" = more than 35.

4.2.2.4 Support Adjustments

A conditionally simulated realization exists only at discrete locations in space -- usually grid nodes. Each simulated value may represent either a point-support observation at that grid node or a non-point volume of material centered at that node. Although it is possible in principle to construct a conditional simulation on one support from data observed on another support (rigorously only if the data support is smaller), the procedure is greatly simplified if both supports are the same. Variable support is particularly difficult to manage in any geostatistical study, so it is common practice to perform some kind of "reconstitution" of the variable-support data to create a revised data set with constant support. For this case study, core samples of different lengths must be reconstituted into samples of constant length. The method is described by Journel and Huijbregts (1978, p. 202). To do this it is assumed (outrageously) that the material within a given analyzed section of core has a constant composition represented by the core analysis. Then to obtain a new data set on a constant support, one simply divides the core into a series of new constant-length samples and calculates a composition for each new sample from the average of all old samples that overlap the domain of the new sample, weighting the analysis of each old sample by the proportion of its overlap. Because some old samples will contribute to more than one new sample, this operation has a smoothing effect that is most pronounced in the short-scale behavior of vertical semivariograms.

The original Mason City data consisted of three observations: $CaCO_3$, $MgCO_3$, and rock type. Within thick beds of limestone or dolostone, the

core was usually sampled at ten-foot intervals; i.e., the core was divided at depths of 10, 20, 30, etc., feet from the surface of the hole. Departures from this regular sampling pattern occurred wherever there was a change in rock type that persisted for more than a foot or so. Thus a two-foot-thick bed of dolostone would be sampled separately, but a ½-foot-thick bed usually would be included in the sample of the surrounding limestone. (Problems resulting from these "polluted" samples are resolved in Section 4.2.2.2.) These variable-support data were converted to a constant-support data set with five variables: $CaCO_3$ in limestone (CALS), $MgCO_3$ in limestone (MGLS), $CaCO_3$ in dolostone (CADOL), $MgCO_3$ in dolostone (MGDOL), and percentage of dolostone in each new sample (PCTDOL). Wherever the percentage of dolostone was 0% or 100%, only two of the four chemical variables would be observed.

The vertical thickness chosen for the new samples was <u>three feet</u>. This choice was a compromise. On the one hand, the very important rock-type variable (responsible for most of the variability in the deposit, as noted in Section 4.2.6) was observed effectively at point support in the original data set, because rock types were defined in the geologist's logs by their contacts, rather than by their presence or absence within a sample interval. Thus the constant-support data set could have been defined on even a 0.1-foot vertical interval without "creating new information" on rock type. On the other hand, the chemical data were observed on larger supports: commonly about three feet for dolostone analyses (as dolostone beds usually are thin in this deposit), and commonly ten feet for limestone, except where limestone

and dolostone are interbedded. Thus a great amount of vertical smoothing can be expected in the reconstituted chemical data. In the limestone, especially, this smoothing results not only from local reconstitution across the sample boundaries but from the fact that the average support has been greatly reduced by the reconstitution, such that strings of two or three reconstituted limestone samples commonly will have identical analyses.

Another consideration in the choice of this spacing is its effect on the size of the simulation grid (Section 4.2.3). With a three-foot spacing, the unconditional simulations will contain about one-half million five-variate simulated data -- about as large as can be conveniently and cheaply handled with the computer facilities being used. A smaller spacing, and thus a larger data set, would have been too cumbersome and expensive to handle. Finally, selective mining operations that might be performed on this deposit could not be expected to separate one rock type from another very efficiently if the rock units were less than three feet thick. From an economic standpoint, selection even at this level would be prohibitive using the current quarrying methods.

The vertical smoothing of the data would be unacceptable if small-scale vertical selective mining were economically feasible, but it is not. In the mining simulations of Section 4.2.5.1, only selectivity at a much larger scale than three feet is considered, so this smoothing is of no consequence in practice.

4.2.2.5 Structural Analysis of the Untransformed Data

Much of the limestone deposit represented by the constant-support data set has been mined out already, and the remaining areas, from casual inspection of logs and cross sections and from field observations, appear to differ somewhat from one another. As a preliminary check on the stationarity of the five variables to be simulated, the quarry property, including mined-out areas, was divided into five large regions (Figure 73) and summary statistics, histograms, and scattergrams among the variables were calculated for each region. Synthetic chemical data (Section 4.2.2.2) were excluded. Some of these statistics are reported in Table 8. It is evident from inspection of these statistics and the associated plots (not reproduced) that the five regions are really not much different from one another. Thus most conclusions to be drawn from a simulation study of one region can, with some caution, be applied in a general way to the other regions. In this study, only Region 2 (Figure 73), on the west side of the property, was analyzed and simulated in detail. This region was chosen because it appears to exhibit a wide variety of geologic and mining situations, and because it is a likely area for quarry expansion within the next several years. The whole deposit was not simulated, simply because a simulation of such a large volume of material in suitable detail would involve several million simulated analyses and would be hopelessly time-consuming and expensive to perform with the time-sharing computer facilities being used here.

Within Region 2, each of the fifteen sample direct and cross semivariograms among the five constant-support variables was calculated in each of thirteen directions, using a cross-semivariogram program

TABLE 8

Statistical summary of the constant-support data.

Variable	Region	Count	Minimum	Mean	Maximum	Std. Dev.
	1	1115	0.0	24.102	100.0	39.9
	2	2048	0.0	29.519	100.0	42.2
PCTDOL	3	2785	0.0	26.410	100.0	39.5
	4	3782	0.0	24.273	100.0	38.9
	5	3476	0.0	28.950	100.0	41.2
	1	834	87.1	96.513	99.8	1.92
	2	1383	74.7	95.838	99.3	2.58
CALS	3	1933	81.3	96.288	99.0	2.04
	4	2620	79.8	96.386	99.4	1.96
	5	2374	82.4	96.760	99.4	1.83
	1	834	0.0	0.864	9.6	0.95
	2	1383	0.0	1.158	11.3	1.66
MGLS	3	1933	0.0	0.880	13.8	1.01
	4	2620	0.0	1.053	11.4	1.30
	5	2374	0.0	0.873	9.2	0.97
	1	294	50.0	63.187	81.7	6.52
	2	655	38.2	59.217	88.1	7.24
CADOL	3	835	44.8	62.351	86.9	5.21
	4	1054	36.8	57.381	86.4	7.47
	5	1060	18.8	61.144	89.2	6.02
	1	294	6.7	31.265	43.0	7.45
	2	655	0.3	33.962	44.6	5.83
MGDOL	3	835	2.2	33.597	47.5	5.60
	4	1054	8.2	33.448	45.9	5.30
	5	1060	7.8	33.551	43.7	5.01

Simplified correlation coefficients (r) for all five regions:

	PCTDOL	CALS	MGLS	CADOL
	-0			
CALS	-0 +0			
	+0 -0			
	+0	-5		
MGLS	+0 +0	-7 -5		
	+0 +0	-6 -5		
	+0	+2	-0	
CADOL	+0 +0	+0 +1	+1 +0	
	-0 -0	+0 +1	+0 +0	
	+0	+0	-0	-5
MGDOL	-0 +0	+0 +0	-0 +0	-4 -6
	+0 +1	+0 +0	+0 +0	-1 -4

Key to Regions
1
2 3
4 5

Key to Absolute Values at Left	
0	$0.0 < \lvert r \rvert < 0.1$
1	$0.1 < \lvert r \rvert < 0.2$
etc.	

Most data sets contain some "anomalous" values; e.g., some minimum MGDOL values represent "clayey dolostone", which has compositions ranging from mostly clay to pure dolomite.

designed for data on a regular grid (and allowing for the many missing data in the drilling grid). The thirteen directions are those connecting all opposite sides, edges, and vertices of a cube with edges pointing north-south, east-west, and vertical.[109] A unit lag in each direction is equal to

$$\sqrt{[(NS \cdot 100)^2 + (EW \cdot 100)^2 + (VERT \cdot 3)^2]}$$

where NS, EW, and VERT are the direction indicators -1, 0, or +1, and the multipliers 100, 100, and 3 are the distances between holes on the drilling grid and between adjacent down-hole sample centers for the constant-support data. The north-south (1,0,0) direct semivariograms of the five variables are displayed in Figure 84. The north-south semivariograms tend to be better estimated (more pairs) than those with east-west components, because Region 2 is longer in the north-south direction and because the holes were drilled 100 feet apart north-south, usually 200 feet apart east-west. The rock-type semivariogram (Figure 84(a)) is nicer than the others because it was calculated with more pairs (the rock type being observed at <u>all</u> sample locations), and because the distribution of dolostone in the rock is a fairly regular-looking phenomenon anyway (Figure 82).

The smoothing effect of the support correction can be seen clearly in the vertical semivariograms of the chemical data, such as the CALS semivariogram displayed in Figure 85. The first three lags, representing distances of three, six, and nine vertical feet, all lie within the typical ten-foot length of the original limestone core

---

[109] Components (south,east,down) of these directions are: (1,0,0), (1,1,0), (0,1,0), (-1,1,0), (1,0,1), (1,1,1), (0,1,1), (-1,1,1), (1,0,-1), (1,1,-1), (0,1,-1), (-1,1,-1), (0,0,1).

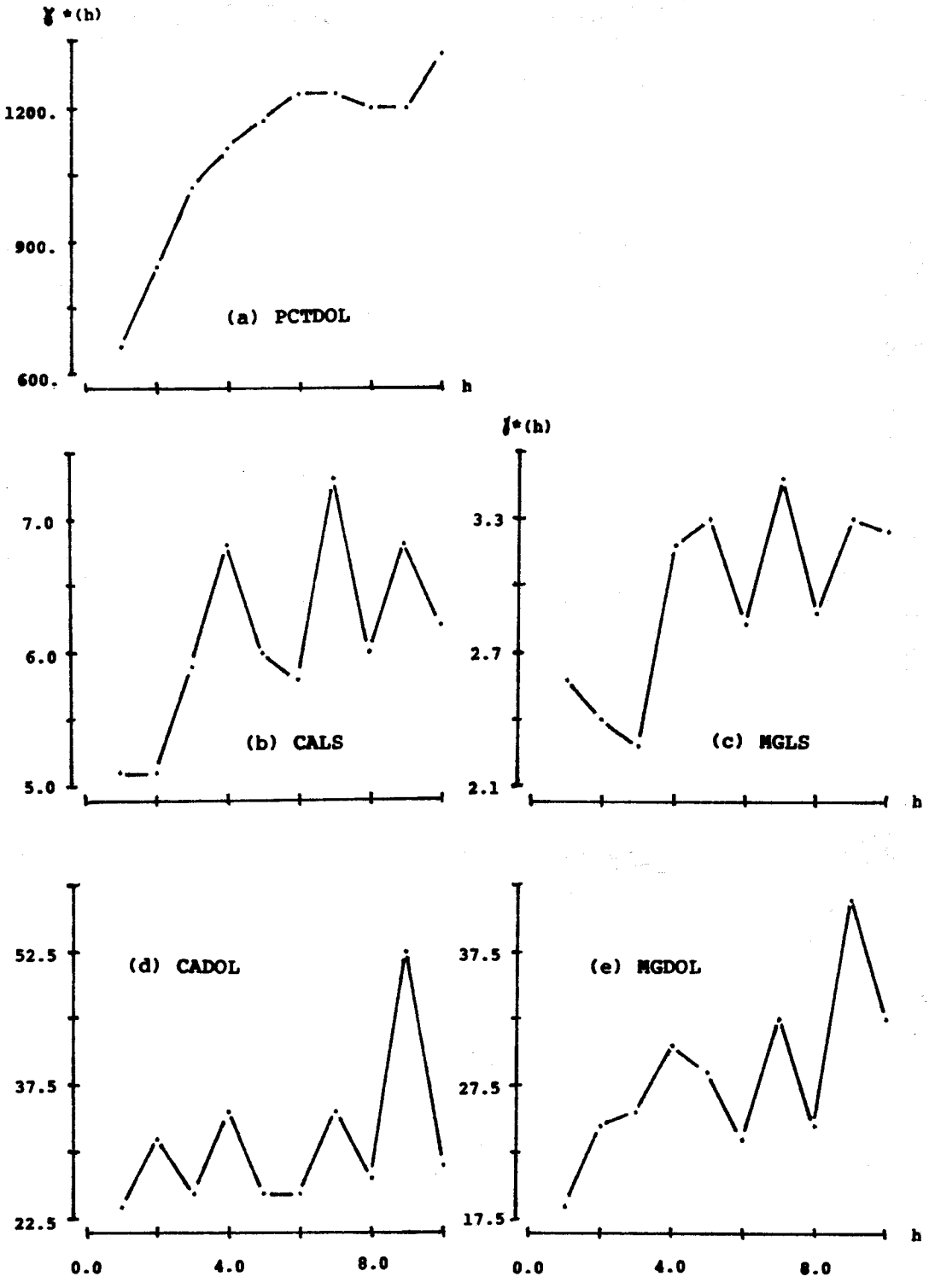


Figure 84: Sample north-south semivariograms of the five variables to be simulated.

samples. Little if any structure is evident beyond nine feet. Notice in particular that the north-south semivariogram of CALS (Figure 84(a)) displays a large nugget effect, which disappears in the vertical semivariogram.

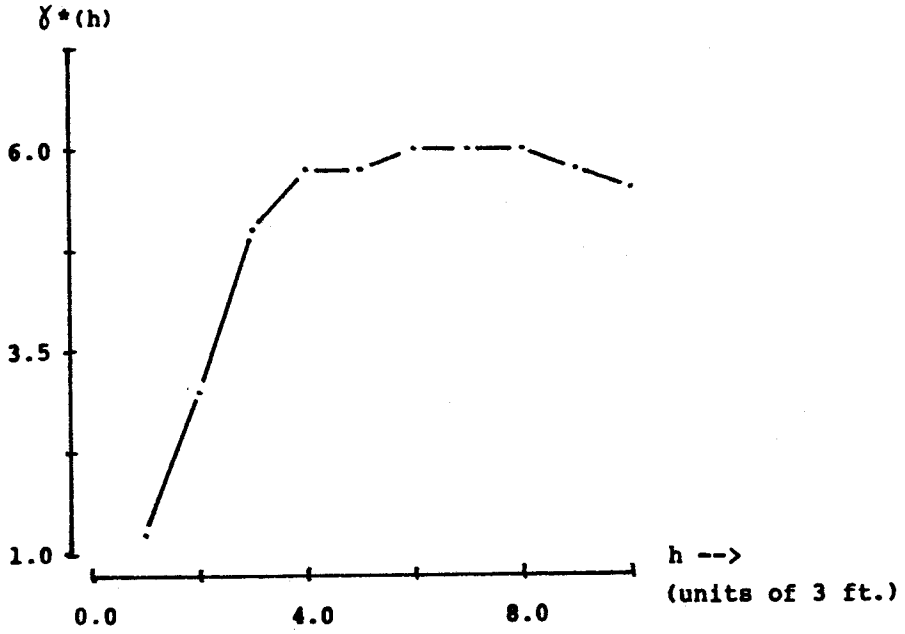


Figure 85: Vertical sample semivariogram of $CaCO_3$ in the limestone.

An interesting phenomenon appears in the sample semivariograms with east-west components (Figure 86). Even-numbered lags have been calculated with many more pairs than odd-numbered lags, because the usual east-west hole spacing is 200 feet. Thus the sample semivariograms look much smoother if the odd-numbered lags are ignored. Furthermore, the few areas within Region 2 where additional north-south

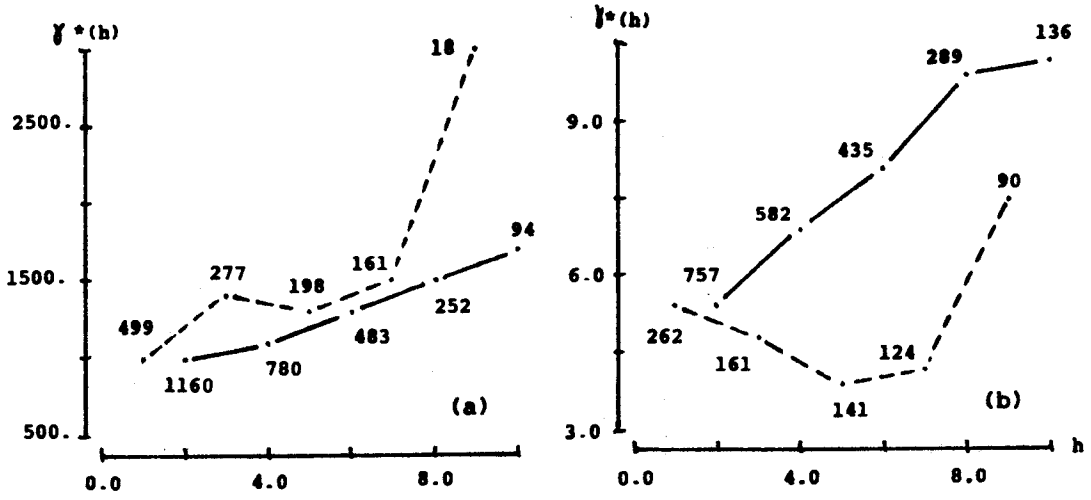


Figure 86: Some sample semivariograms with east-west components, exhibiting the "two-sample" effect. (a) Semivariogram of PCTDOL in the NE-SW direction. (b) Semivariogram of CALS in the E-W direction. The number adjacent to each point represents the number of pairs used in the calculation.

columns of holes at 100-foot east-west spacings were inserted are nearly isolated from one another (Figure 73), so that these additional holes contribute to the sample semivariograms almost exclusively at odd-numbered lags. Thus the even and odd lags have been calculated from very different, though overlapping, data sets, and in some cases seem to suggest different models. The effect persists after repeated transformations. In Section 4.2.2.7, model semivariograms are fitted to sample semivariograms of the normal scores of transformations of these data. The even-numbered lags were considered more reliable for fitting (except in the north-south semivariograms), because they were calculated from more pairs, because they represented an even 100x200-foot grid of

data over the whole region, and because they were not calculated with data from subjectively chosen additional holes.[110]

4.2.2.6 Transformations

Transformations to account for constraints. The raw chemical data are subject to a maximum-sum constraint of approximately 100%. Furthermore, the dolostone analyses should ordinarily respect the constraint:

$$MgCO_3/(CaCO_3+MgCO_3) \leq 0.457$$

Analyses violating this constraint would suggest that magnesite and dolomite, rather than calcite and dolomite, were the dominant carbonate phases in the dolostone; geologically, this would be a rare occurrence. So that these constraints can be taken into account in the simulation, the following transformations were performed, as explained in Section 3.5.2.4:

```
TOTLS = CALS + MGLS

TOTDOL = CADOL + MGDOL

RATLS = 100.0 * (MGLS+0.01)/(TOTLS+0.01)

RATDOL = 100.0 * (MGDOL+0.01)/(TOTDOL+0.01)
```

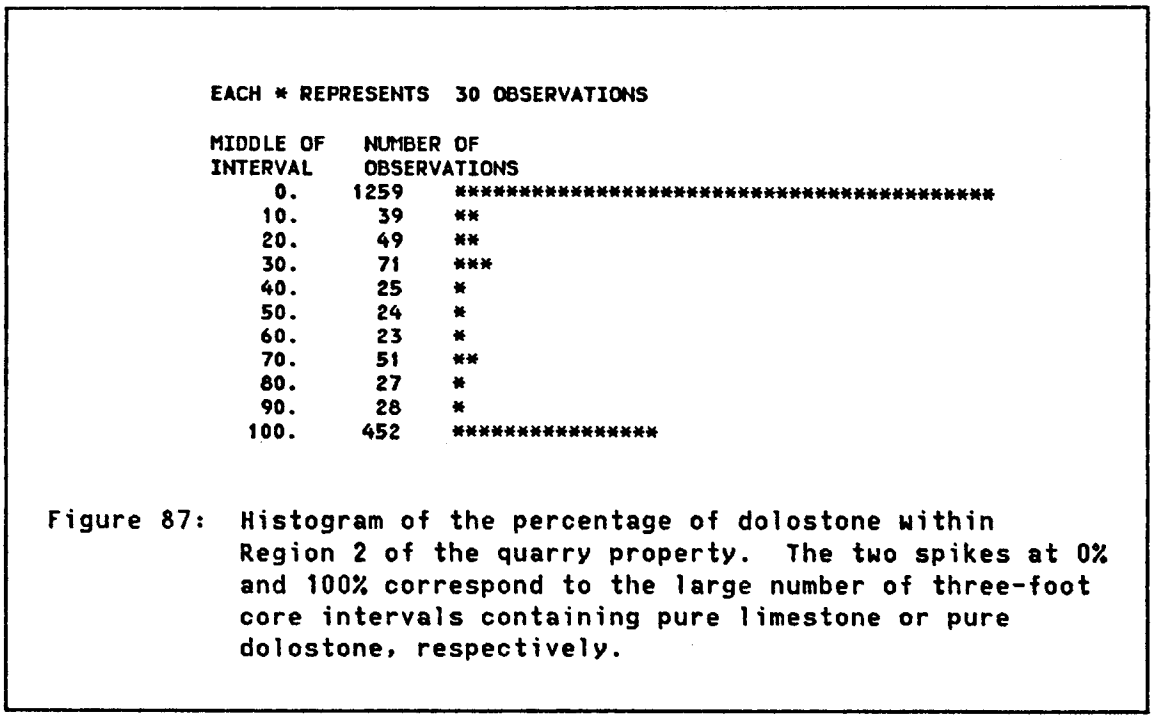
The numerators and denominators of the ratio transformations RATLS and RATDOL were "started" by 0.01 for two reasons. First, it is a good idea always to start the denominators of ratio transformations by a small value so the ratios do not fly off toward infinity if the major term in the denominator approaches zero (actually not a problem in this data set). The numerators were started as a partial "despiking" measure

---

[110] The geologist asked that the additional north-south columns of holes be drilled in areas where he was having trouble correlating the stratigraphy between holes at 200-foot spacings.

for the limestone data. The MGLS data contain many zero values, which would produce a spike of zero RATLS values as well, regardless of the associated CALS values. But it was felt that a zero MGLS and a large CALS really should transform to a smaller RATLS value than a zero MGLS and a small CALS, because a zero value for a chemical variable really means "below detection limit", not zero. This despiking is accomplished by adding 0.01 to the numerator. (Actually, this measure alone is not sufficient to remove the ill effects of the zero spike, as we shall see later in this section.)

Normal-scores transformation. The next step is to transform the four new chemical variables, plus the untransformed PCTDOL variable, into univariate gaussian data by means of the normal-scores transformation, discussed in Section 3.5.2.2. As in the case study of Section 4.1, the convenient MINITAB "nscores" function was used to obtain the normal scores, although indirectly. A straightforward plug-in of the data to obtain this transformation would be unsatisfactory, for two reasons: first, the PCTDOL distribution contains two huge spikes at 0% and 100% (Figure 87), corresponding to three-foot thick samples of pure limestone or pure dolostone, respectively; and second, there are so many data to be transformed that the inverse transformation of the simulated data by linear interpolation (Section 3.5.2.2) would become extremely time-consuming because of the necessity of sifting through a large number of ordered bounds to find the appropriate interval for interpolation. Although ordering the simulated data would eliminate the second problem, the task of ordering hundreds of thousands of data presents some problems of its own. Both of these problems are discussed in the paragraphs that follow.



EACH * REPRESENTS 30 OBSERVATIONS

MIDDLE OF INTERVAL	NUMBER OF OBSERVATIONS	
0.	1259	**********
10.	39	**
20.	49	**
30.	71	***
40.	25	*
50.	24	*
60.	23	*
70.	51	**
80.	27	*
90.	28	*
100.	452	**********

Figure 87: Histogram of the percentage of dolostone within Region 2 of the quarry property. The two spikes at 0% and 100% correspond to the large number of three-foot core intervals containing pure limestone or pure dolostone, respectively.

Clearly we cannot allow MINITAB to assign the same normal score to all data within each spike in the PCTDOL distribution, as the resulting "gaussian" distribution would still contain two spikes. Some method must be adopted for ordering the tied data or at least for decomposing the large spikes into several much smaller ones. Clues to the kinds of despiking methods that would be appropriate can be obtained from the discussion of indicator simulations in Section 3.8.3. Indicator simulations are relevant here because the PCTDOL data really consist of a regularization of point-support indicator data (0=limestone, 100=dolostone) over a three-foot vertical support. This support is small enough to allow the regularized data to retain much of the indicator character (i.e., the spikes) of the original data. Unfortunately, there seems to have been little theoretical work done on

the derivation of appropriate gaussian conditioning data for simulations of this type (Sections 3.8.3.4), so the approach used here is very much ad hoc. This is unfortunate, as the percentage of dolostone in the raw material is the major contributor to variations in the amount of $MgCO_3$ in the material delivered to the cement plant (Section 4.2.6). However, the density of drilling on the property should prevent the simulation from going far astray, even if the model employed for the unconditional simulations is inappropriate.[111]

The PCTDOL data must be transformed into a data set that appears to have been drawn from a realization of a three-dimensional gaussian process. This process probably should have a very regular character, because the percentage of dolostone in the geologic section is a very regular-looking phenomenon, at least at short scales. Dolostone beds observed in the quarry faces (Figure 88) normally change thicknesses very gradually, and the PCTDOL sample semivariograms (e.g., in Figures 84 and 86) show fairly smooth behavior with low relative nugget constants (at least, low compared to those of the chemical semivariograms). This regular character is reflected in the "distance" despiking method, described below, which was applied to the 0% and 100% PCTDOL data.

For each 0% or 100% datum, a "structural distance" to all other data with different PCTDOL values was calculated by

$$\text{DISTANCE} = \sqrt{(NS^2 + EW^2 + 4*VERT^2)}$$

---

[111] However, in Section 4.2.4 it is shown that the despiking procedure used here results in some unfortunate smoothing of the short-scale PCTDOL semivariogram structure, which cannot be corrected by conditioning. However, this does not mean that the procedure would be inappropriate for any data set.

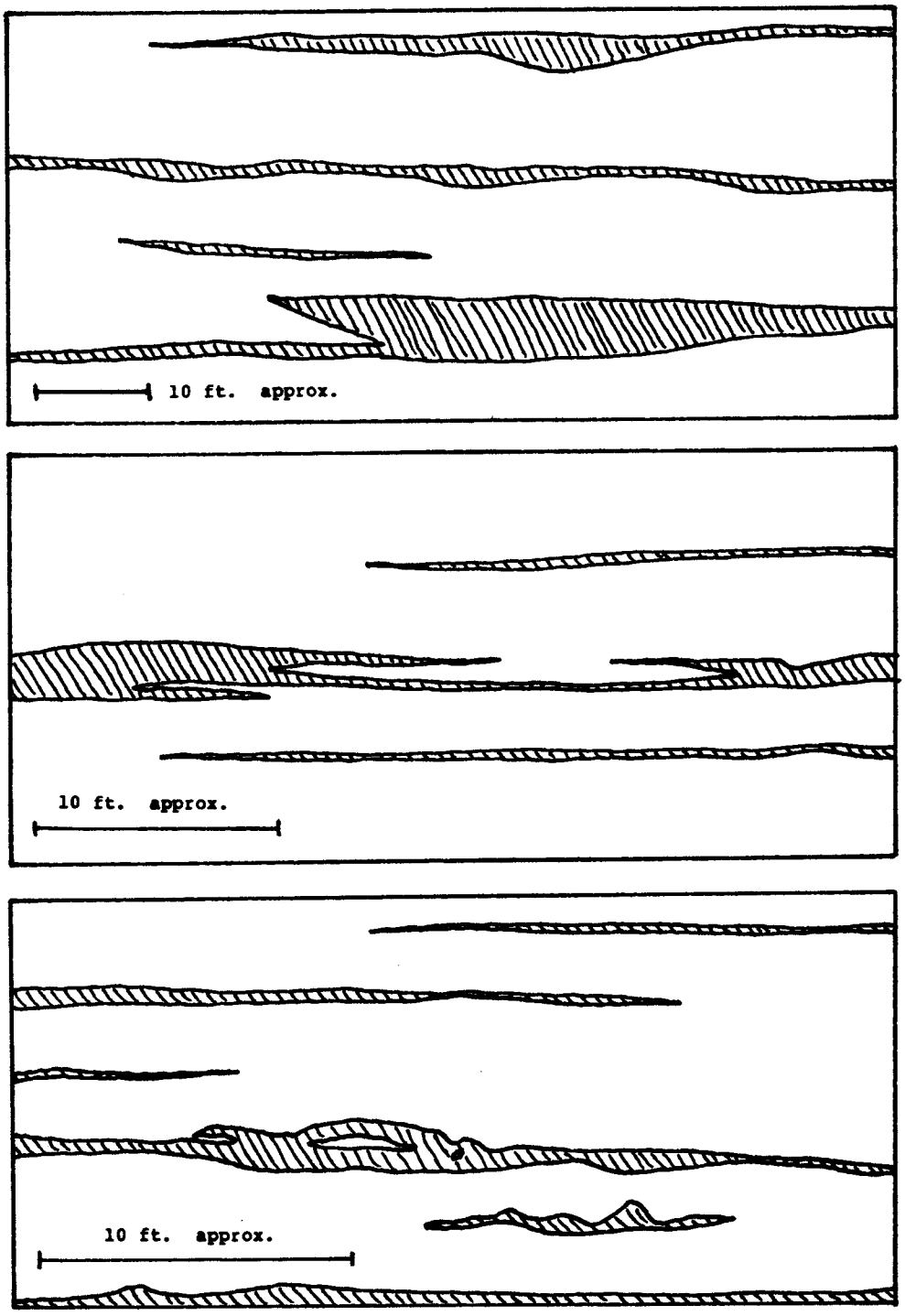


Figure 88: Sketches of limestone and dolostone beds exposed in the quarry faces. The dolostone beds are indicated by hachures.

where NS, EW, and VERT are "data-grid" distances in the corresponding directions (1 grid unit equals 100 feet horizontally, 3 feet vertically). The vertical grid distance was effectively doubled to account for the lesser long-range continuity of the phenomenon in the vertical direction, as observed in the PCTDOL semivariograms. The minimum ("MINDIST") of all such distances, i.e., the distance to the structurally closest datum that is not in the same spike, was stored and used subsequently for despiking. The idea here is that 0% data (pure limestone) lying structurally very far from any non-zero data should be transformed to a lower gaussian value than data close to a dolostone contact. Similarly, 100% data (pure dolostone) should receive a higher gaussian value if no limestone is nearby.

Use of this distance criterion alone will not do a very thorough job of despiking, because many data will have the same distance value, thus the big spikes will simply be broken into a set of smaller spikes. To decompose these smaller spikes further, a second despiking criterion, a "moving-window" approach (Section 3.8.3.4), is introduced. This criterion ("COUNT") is the number[112] of PCTDOL data within a structural distance of 5.0 that are different from the datum being despiked. In other words, a 0% value surrounded by nonzero values would receive a higher count than a 0% value with only a few nonzero values nearby. This secondary despiking operates most effectively when two data are located at the same short minimum distance from a contact, but one is surrounded by the contact on most sides (large COUNT), whereas the other is not.

---

[112] Actually the proportion would have worked better, as the number always decreases near the edges of the sampled domain.

The final despiking is then performed as follows. If the original PCTDOL is 0.0, the despiked PCTDOL is:

$$\text{New PCTDOL} = 0.0 - 0.01 \, (\, \text{MINDIST} - 0.01 \, \text{COUNT} \,) - 0.2$$

If the original PCTDOL is 100.0:

$$\text{New PCTDOL} = 100.0 + 0.01 \, (\, \text{MINDIST} - 0.01 \, \text{COUNT} \,) + 0.2$$

Thus each despiked value is given a more extreme value if MINDIST is large, but is then rendered a little less extreme if COUNT is also large. The most extreme values will thus be those that are a long way from any data from outside the spike. The least extreme values will be those completely surrounded by data from outside the same spike. These alternatives are depicted in Figure 89. The effect of the total despiking procedure is to put deep "valleys" in the normal scores of the PCTDOL data wherever there is a large volume of uncluttered limestone, high "mountains" wherever there is a large volume of unadulterated dolostone, and middling values near the contacts. The constants 0.2 are subtracted from or added to the 0% or 100% data, respectively, to ensure that all 0% values are despiked to values less than 0%, and all 100% values are despiked to values greater than 100%. In this way, the despiked values cannot overlap data lying inside the (0%,100%) interval.

In the exposition of the normal-scores transformation in Section 3.5.2.2, and in the previous case study of Section 4.1, <u>all</u> of the original data and their normal scores were used to define a series of linear-interpolation bounds for the inverse transformation of the simulated data. In this case study, this approach presents a problem, because there are between 655 and 2048 original data for each variable, and about 350,000 conditionally simulated data to be back-transformed.

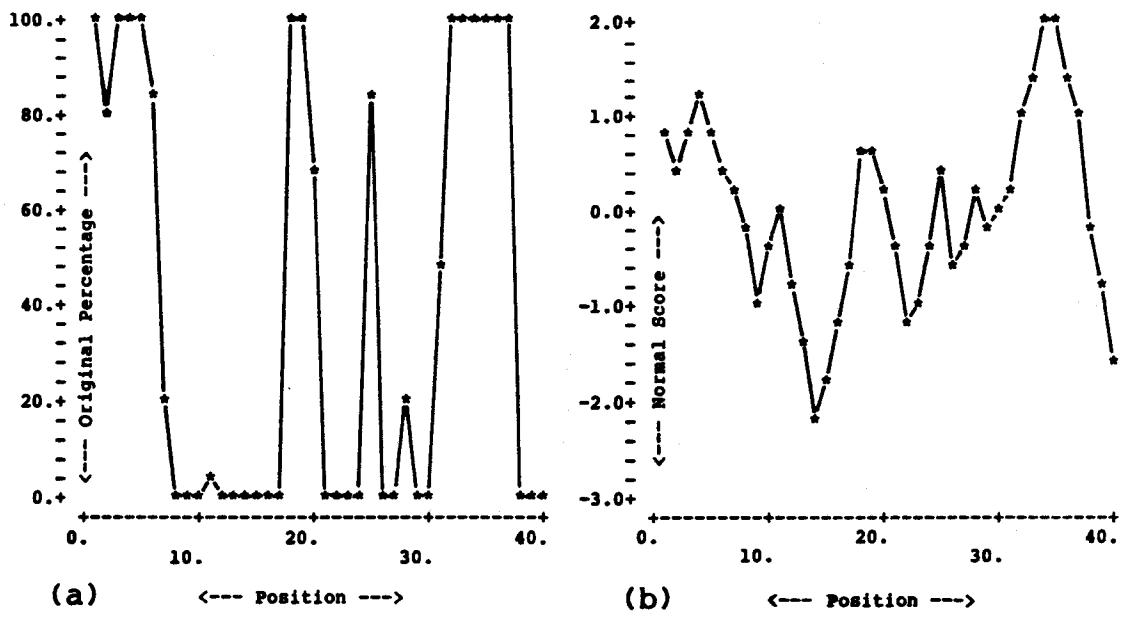


Figure 89: Despiking of percentage data by a combination of "distance" and "moving-window" methods. (a) Original data. (b) Normal scores of despiked data.

If we simplistically sifted through an average of about 300-1000 bounds (about half of the original data) to find the right pairs for each interpolation, the 100 million to 350 million comparisons involved for each of the five variables would consume an inordinate amount of computer time. The search is greatly simplified if the simulated data are ordered first (as in subroutine LINT, listed by Verly, 1984b), but the sorting of 350,000 data for each of five variables also consumes a lot of time and storage. One of several simple ways to reduce the amount of time needed for the inverse transformation is to reduce the total number of bounds used in the inverse transformation. This requires that linear interpolation be used for both the inverse and

forward gaussian transformations, if the original data values are to be reproduced by the simulation. The forward transformation then does not produce "exact" normal scores, but rather linearly interpolated approximations of normal scores. The exact normal scores are used only at the bounds. The method is illustrated in Figure 90. First the normal scores of all data are obtained (with MINITAB in this example). Then several pairs of original and transformed data are picked out of the resulting data sets to serve as forward and backward linear-interpolation bounds. The number of pairs selected is greater for data whose original distributions do not look very smooth, such that a more detailed piecewise-linear approximation is needed. Also, more bounds are used in the tails of each distribution than in the middle, so that the sensitive tail portions of the gaussian distribution are better reproduced. For instance, in this case study, the following numbers of original and transformed data were selected:

Variable	Number of Original Data Pairs	Number of Pairs Selected as Bounds
PCTDOL	2048	61
TOTLS	1383	108
RATLS	1383	123
TOTDOL	655	51
RATDOL	655	62

The final step in the forward transformation is to re-transform the original data to a nearly gaussian data set by linear interpolation. This ensures that the eventual inverse transformation of the conditionally simulated data will reproduce the original data exactly.

Despiked 0% and 100% data in the tails of the PCTDOL distribution were not transformed by interpolation; the original normal scores generated by MINITAB were retained instead. In the eventual inverse

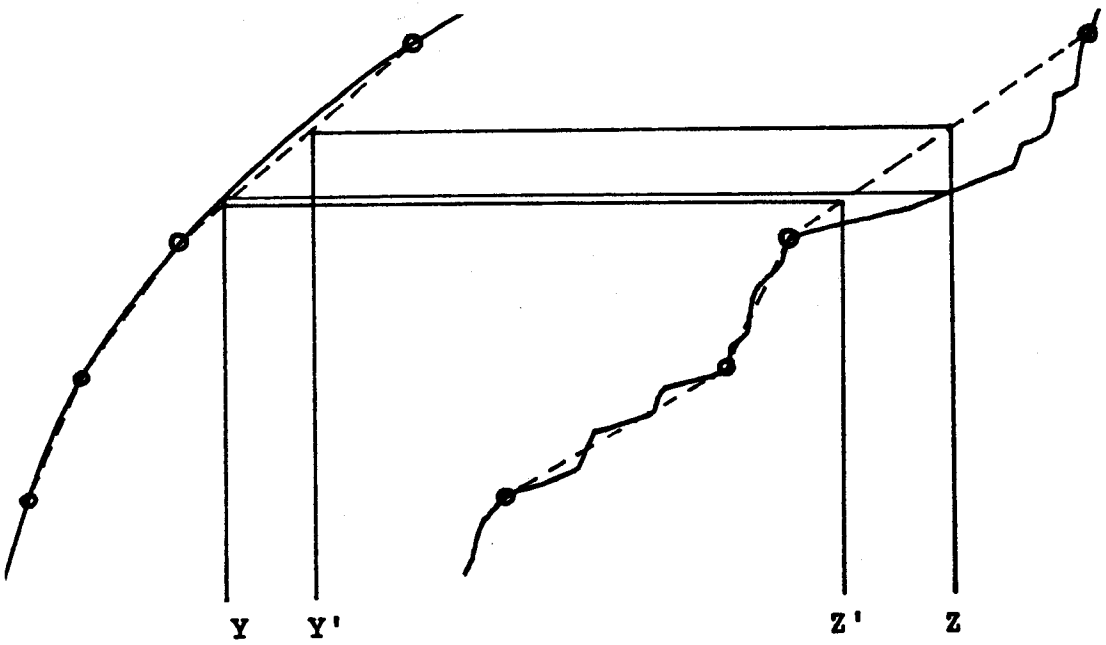


Figure 90: Normal-scores transformation using using both forward and inverse linear interpolation. Datum Z would be transformed to interpolated normal score Y′ instead of exact score Y. A simulated score Y would be back-transformed to interpolated datum Z′ instead of exact datum Z. These inaccuracies decrease as the linear approximations approach the true distribution functions.

transformation of the simulated gaussian PCTDOL data, all simulated data in the tails are simply transformed back to 0% and 100%, respectively. Thus the spikes are restored automatically by the inverse transformation.

These transformations were first performed only on the "real" data from Region 2 -- not on the synthetic data -- to get data for a structural analysis of the transformed phenomena. To get data for conditioning, the same sequence of transformations, including a forward

linear interpolation using the same bounds, was performed on the entire set of data from Region 2.

Results of the transformations. Statistics of the "real" transformed data from Region 2 are summarized in Table 9,[113] and scatterplots of selected pairs of data are displayed in Figure 91. The histograms, not shown, of course appear very gaussian. The statistics are not so nice as those that would be obtained from a set of true normal scores of fully despiked data: such data would have mean zero, variance very close to 1.0 (but not exactly -- see Section 3.5.2.2), and equal maximum and minimum values. Nevertheless, these data are certainly close enough to the correct univariate gaussian statistics to be treated as standard normal data for simulation purposes.

The scatterplots reveal some initially shocking features, however. Under the multigaussian hypothesis (Section 3.5.1), one would expect all scattergrams to look something like that between PGT(TOTLS) and PGT(TOTDOL), depicted in Figure 91(d). This plot (and in fact most plots, not shown here) exhibits the elliptical character typical of bivariate normal data. However, all scatterplots containing PGT(PCTDOL) exhibit the "truncated-normal" pattern of Figures 91(a) and (b), and the PGT(RATLS) - PGT(TOTLS) scatterplot in Figure 91(c) contains a strange "tail" corresponding to the lowest values of RATLS.

The peculiar appearance of the scatterplots containing PGT(PCTDOL) can be understood when we recall that the four chemical variables have been "preferentially sampled" with respect to PCTDOL. When PCTDOL=0

---

[113] In Table 9 and subsequently, the notation "PGT( )" denotes the "piecewise gaussian transformation" discussed above. It is approximately equal to "$G^{-1}F^*( )$" as in Section 3.5.2.2, except that both G^{-1} and F^* are approximated by a coarse linear interpolation.

TABLE 9

Statistical summary of the linearly interpolated normal scores from Region 2.

Variable	Count	Minimum	Mean	Maximum	Std. Dev.
PGT(PCTDOL)	2048	-3.3289	-0.002	3.3288	0.997
PGT(TOTLS)	1383	-3.2394	-0.002	2.9999	1.00
PGT(RATLS)	1383	-3.2395	-0.009	2.9413	1.01
PGT(TOTDOL)	655	-3.0549	-0.019	2.6402	0.991
PGT(RATDOL)	655	-3.0336	0.000	2.7884	0.998

Correlation Matrix

	PGT(PCTDOL)	PGT(TOTLS)	PGT(RATLS)	PGT(TOTDOL)
PGT(TOTLS)	-0.054			
PGT(RATLS)	0.119	-0.025		
PGT(TOTDOL)	0.277	0.377	0.045	
PGT(RATDOL)	-0.066	-0.125	-0.062	-0.077

(<0, despiked), there can be no observations of TOTDOL and RATDOL; similarly, when PCTDOL=100 (>100, despiked), there are no observations of TOTLS and RATLS. Nevertheless the simulation model described in the next section is formulated as though three feet of limestone and three feet of dolostone could coexist at each grid node, regardless of the value of PCTDOL. (The simulated values of PCTDOL are used only in the final stage of the simulation to proportion the simulated compositions of the two rock types into a final simulated composition for each three-foot interval.) Thus the data set behaves as though we had taken some of the normal scores of the chemical data corresponding to each spike of the PCTDOL distribution and simply thrown them away. The correlation coefficients in Tables 8 and 9, and any other measures of joint

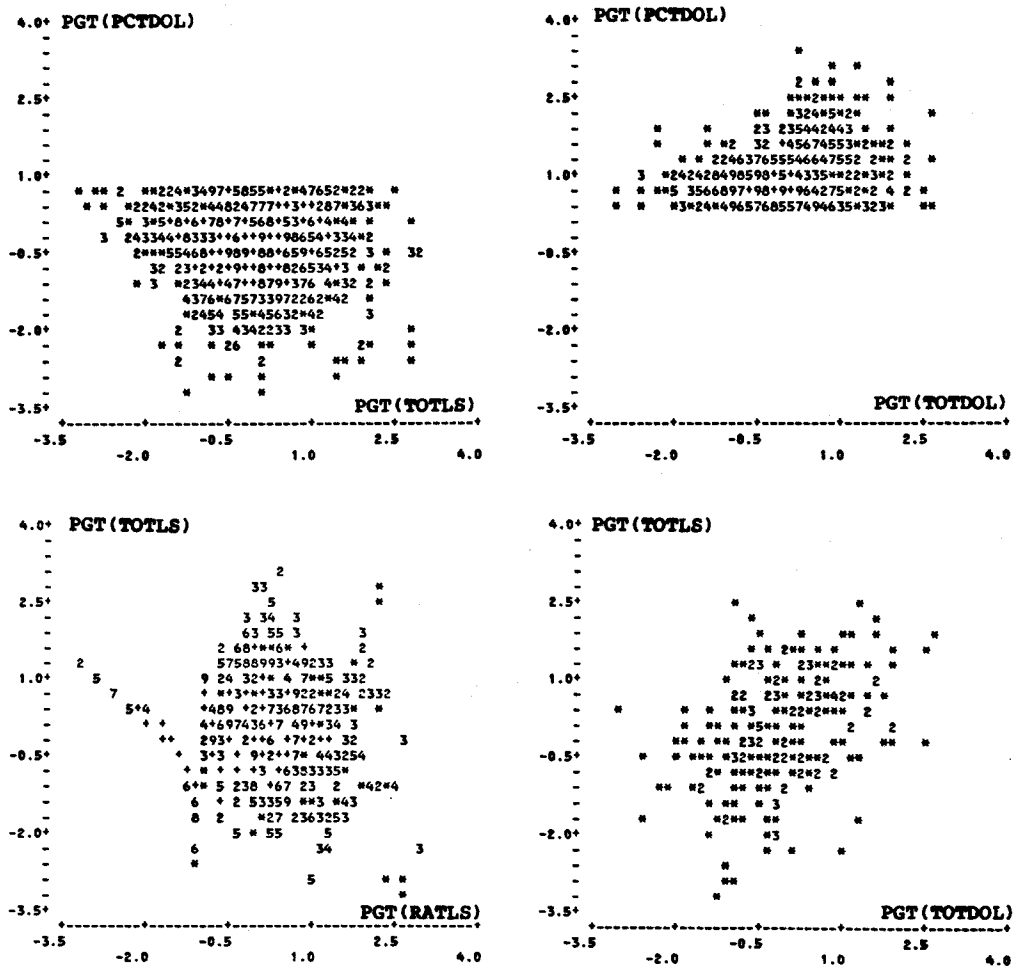


Figure 91: Scattergrams of selected pairs of normal scores.

variability among the data, such as cross semivariograms, actually measure the behavior of this hypothetical simulation model for five joint-normal variables only if we hypothesize that the behavior of the "preferentially sampled" data accurately reflects the behavior of the total (unrealizable) phenomenon being simulated.

The same hypothesis must be invoked whenever we measure the joint behavior of chemical variables from two different rock types. For

example, sample cross semivariograms between PGT(TOTLS) and PGT(TOTDOL) can be calculated using only data from samples in which 0<PCTDOL<100 -- only 225 out of 2048 samples from Region 2. If this were thought to be seriously biasing the calculations, the noncentered cross-covariance function between these variables might be calculated instead of the cross semivariogram, as explained in Section 3.8.4. However, the structural analysis of these sample semivariograms presented in the next section (and the correlation matrix in Table 9) strongly suggest that in this data set the chemical compositions of the two rock types are essentially uncorrelated. In any case, the conditioning data for these simulations are so ideally located that careful modeling of their joint variability will not be critical.

The peculiar appearance of the plot in Figure 91(c) is a consequence of the spike of zero MGLS values, producing a similar concentration of near-zero RATLS values. These values are not exactly zero because the numerator of RATLS was "started" by a constant value of 0.01 in the transformation:

$$RATLS = 100.0 * (MGLS+0.01) / (CALS+MGLS+0.01)$$

Among the samples with zero MGLS values, the calculated RATLS values are exactly determined by, and inversely proportional to, the CALS values entered in the denominator of the transformation. The values of

$$TOTLS = CALS + MGLS$$

are also exactly determined by the CALS values. Thus the large spike of zero MGLS values is dissolved not into a cloud of bivariate-normal-looking values but into a series of small spikes in which RATLS and TOTLS are inversely related. This relationship persists into the normal

scores, as illustrated by the tail on the left side of Figure 91(c). This tail represents 215 of the 1383 limestone analyses in the data set and must therefore be regarded as a nontrivial departure from the assumption of multivariate normality that must be made in order for the simulations to reproduce the original spatial distribution of limestone grades. In practice, however, this departure is really not very important. Although the simulations will not reproduce this peculiar scatterplot, any simulated value falling below the highest PGT(RATLS) value in the spike will still automatically be assigned to MGLS=0.0, CALS=TOTLS. And although the semivariogram model adopted for PGT(RATLS) may not be exactly appropriate for reproducing the spatial distribution of limestone compositions, this distribution appears to be almost a pure nugget effect (according to the semivariograms of both the raw and the transformed data) with a vary low variance (in the raw data). The composition of the limestone is thus not an important factor in determining the chemical variability of the mine output. The composition of the dolostone, and particularly the amount of dolostone (PCTDOL), are far more important. Exceedingly accurate reproduction of the CALS and MGLS distributions are thus of little practical importance in this study. (Section 4.2.6 carries this line of thought a bit further.)

4.2.2.7 Structural Analysis of the Transformed Data

North-south sample semivariograms of the normal scores are illustrated in Figure 92. Semivariograms in the other horizontal directions are not so clearly defined, for reasons explained in Section 4.2.2.5.

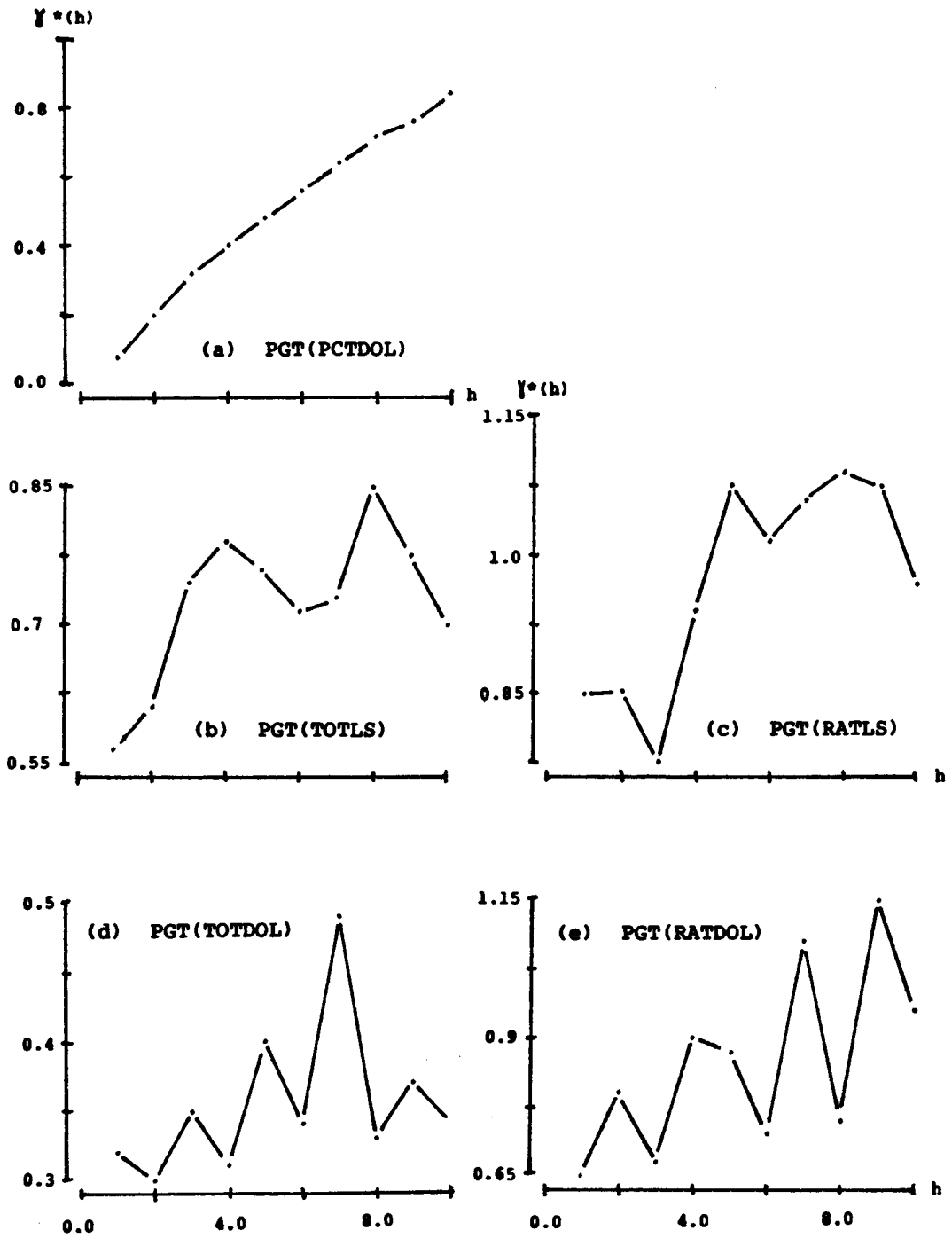


Figure 92: Sample north-south semivariograms of the five normal-score variables to be simulated.

The semivariogram of PGT(PCTDOL) in Figure 92(a) shows no apparent sill, in contrast to the plot of Figure 84(a). This drift in the north-south direction, and in all other horizontal directions except northwest-southeast, is probably the result of long-range trends imposed by the despiking formula (Section 4.2.2.6). The nugget component of the original phenomenon has also disappeared -- also probably the result of the smooth transitions between limestone and dolostone scores imposed by the despiking procedure. (This turns out to be unfortunate, as demonstrated in Section 4.2.4.)

The semivariograms of the chemical variables retain the relatively high nugget components observed in the untransformed data. An attempt was made to gain additional information on the size of the nugget component by examining semivariograms of the interpolated normal scores of five pairs of "twin" holes (shown by the symbol "2" in Figure 73) scattered around the quarry property (none in Region 2). In these five cases, holes were drilled at nominally the same coordinate locations during both the 1955 and 1967 drilling programs, possibly by mistake. Obviously the holes could not have been drilled at <u>exactly</u> the same locations, but assuming reasonable surveying accuracy, it is likely that they were less than ten feet apart, versus the usual 100- to 200-foot spacings in the rest of the data set. Two holes also were in-filled at fifty-foot spacings in one small area in the eastern part of the property (shown by "3" in Figure 73), and short-scale semivariograms using these holes were examined as well. Owing to the small number of pairs available (especially for PGT(TOTDOL) and PGT(RATDOL)), the numerical values of the resulting semivariograms were not reliable.

However, the scattered values from the twin holes did suggest the actual short-scale semivariogram structure is probably not very different from the structure obtained by simply extrapolating larger-scale trends back to the origin. This was the method actually used to obtain nuggets for the semivariogram models used in the simulation. Semivariograms calculated from the fifty-foot holes were dramatically lower than those from the twin holes and the main data set. It was concluded that these results probably reflected only the local structure of the small area where these holes were drilled, so the results were discarded.

Sample direct and cross semivariograms were calculated for the normal scores in the same thirteen directions used for the untransformed data, as described in Section 4.2.2.5. In most of these directions, the sample plots were fairly ragged owing to the small number of pairs, the "two-sample" phenomenon discussed in Section 4.2.2.5, and the peculiarities of the local domain of observation (Region 2). Nevertheless, preliminary models fitted by eye independently to each of these ragged plots generally made sense when compiled together on polar contour plots of the sample semivariogram functions, as illustrated in Figure 93 for the PGT(PCTDOL) semivariogram. The final models for the direct semivariograms were obtained after a careful review of the polar plots. The most questionable fits were those of the nugget constants for the chemical variables, which appeared consistently high but seemed to vary erratically with direction.

In most cases the cross semivariograms were even more ragged than the direct plots, making them difficult to fit. Fortunately, however, these plots nearly always appeared to show zero nugget components and definite

departures from zero only at long lags. The only exception was the cross semivariogram between PGT(TOTLS) and PGT(TOTDOL), which seemed to show a nugget of about 0.1, still small enough to be ignored without seriously biasing the simulation.[114] Because cross associations among the variables become apparent only at longer lags (and in some cases not even there), these associations will automatically be imparted to the simulations by conditioning.[115] Thus it appears that the "total" and "ratio" transformations applied to the chemical data have had the fortunate effect of rendering the four chemical variables practically uncorrelated except for large-scale trends, so in practice they can be modeled and simulated independently. Reliance is placed on conditioning and on the inverse transformations to restore whatever cross associations may be present among the original variables. Fortuitously, it appears that PGT(PCTDOL) is not correlated with the chemical variables, either.

---

[114] Of all pairs of chemical variables, these two show the strongest cross association, even though they were measured in different rock types. They show the highest correlation coefficient in Table 9 and the only noticeably elliptical, rather than circular, cross-chemical scatterplot (Figure 91(d)). The reason for this association may be that the percentage of clay and other noncarbonate material in the rock (approximately equivalent to 100%-TOTDOL in the dolostone, 100%-TOTLS in the limestone) was fixed during deposition of the sediments, whereas patchy dolomitization of the sediments probably occurred later. Neighboring areas of dolomitized and undolomitized limestone would thus retain similar amounts of clay (resulting in cross-correlated TOTDOL and TOTLS) but have radically different and unrelated amounts of $MgCO_3$ (resulting in uncorrelated RATDOL and RATLS). This relationship could be expected to deteriorate rapidly with increasing distance between the limestone and dolostone samples; however, only samples containing <u>both</u> rock types can be used to calculate a cross semivariogram.

[115] However, only 225 of the 2048 data contain both limestone and dolostone analyses, so the cross semivariograms between these rock types will not be very well conditioned.

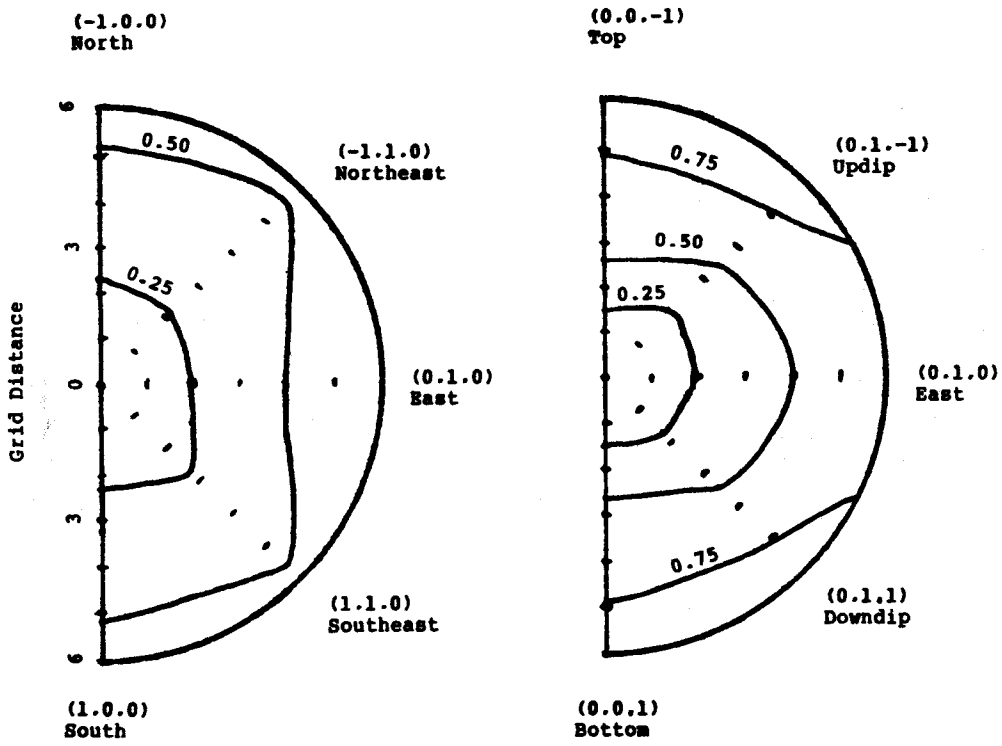


Figure 93: Polar contour plots of the sample semivariogram function for variable PGT(PCTDOL). Only horizontal and east-west sections are shown here.

All five direct semivariograms were fitted independently with isotropic nugget constants and anisotropic linear models, summarized in Table 10. To obtain isotropic nugget constants, the vertical smoothing effect of the reconstitution of core support, visible in the four chemical semivariograms, had to be ignored. Extrapolation of the large-scale behavior of the vertical sample semivariograms to the origin,

ignoring short-scale smoothing, produced estimated nuggets acceptably close to those obtained from the horizontal plots. Conditioning should re-smooth the vertical variability of the simulated data very close to the drill holes but leave them relatively untouched only a short distance away, because of the high nugget components in the data. As the smoothing is artificial anyway, the relative roughness of the simulations should make them more realistic.

For the unconditional simulations, anisotropic spherical models with the same short-scale slopes as the linear models were used. The linear models were used for kriging, except for PGT(RATLS), which has a very short east-west range. The larger-scale discrepancies between the spherical and linear models are unimportant, because in the final simulation, large-scale behavior is imposed by conditioning (Section 3.7.2).

4.2.3 Simulation Procedure

The simulation domain (Region 2) was discretized into a grid of 141 (north-south) by 101 (east-west) by 41 (vertical) points (Figure 94), each of which was associated with a set of five independent unconditionally simulated data. The grid spacing was ten feet horizontally and three feet vertically. Subroutine TB3D (Appendix B) was used to perform the unconditional simulations, using spherical semivariogram models with short-lag slopes equal to the slopes of the linear semivariogram models summarized in Table 10. The larger-scale structures of the simulations in general were different from those inferred from the data, but because of the density of conditioning,

TABLE 10

Nugget constants and linear semivariogram slopes fitted to the normal scores.

Variable	Isotropic Nugget	North-South Slope	East-West Slope	Vertical Slope
PGT(PCTDOL)	0.0	0.106	0.106	0.180
PGT(TOTLS)	0.6	0.040	0.040	0.040
PGT(RATLS)	0.8	0.040	0.400	0.013
PGT(TOTDOL)	0.3	0.011	0.029	0.140
PGT(RATDOL)	0.7	0.030	0.030	0.043

All cross semivariograms were modeled as zero at all lags.

departures at large lags in the unconditional simulations were of no importance (Section 3.7.2). Because the simulations were so large, a modification was made to the subroutine to allow the nugget components of the models to be simulated simultaneously with the single spherical structures. (Otherwise three big files -- spherical, nugget, and total -- of 141x101x41 = 583,881 data would have to be created and stored for each of the four chemical variables, driving up disk-storage requirements.) The nugget components were simulated by IMSL subroutine GGNML (IMSL, 1982), a normal(0,1) generator. A call of this subroutine was inserted into the TB3D code at a point just before the subroutine writes a column of spherical data to the output file (i.e., shortly before statement 4 of the listing shown in Appendix B). The vector of normal(0,1) data returned by GGNML was multiplied by the square root of

the semivariogram nugget constant, added to the column of simulated spherical values, and then written to the output file as usual.

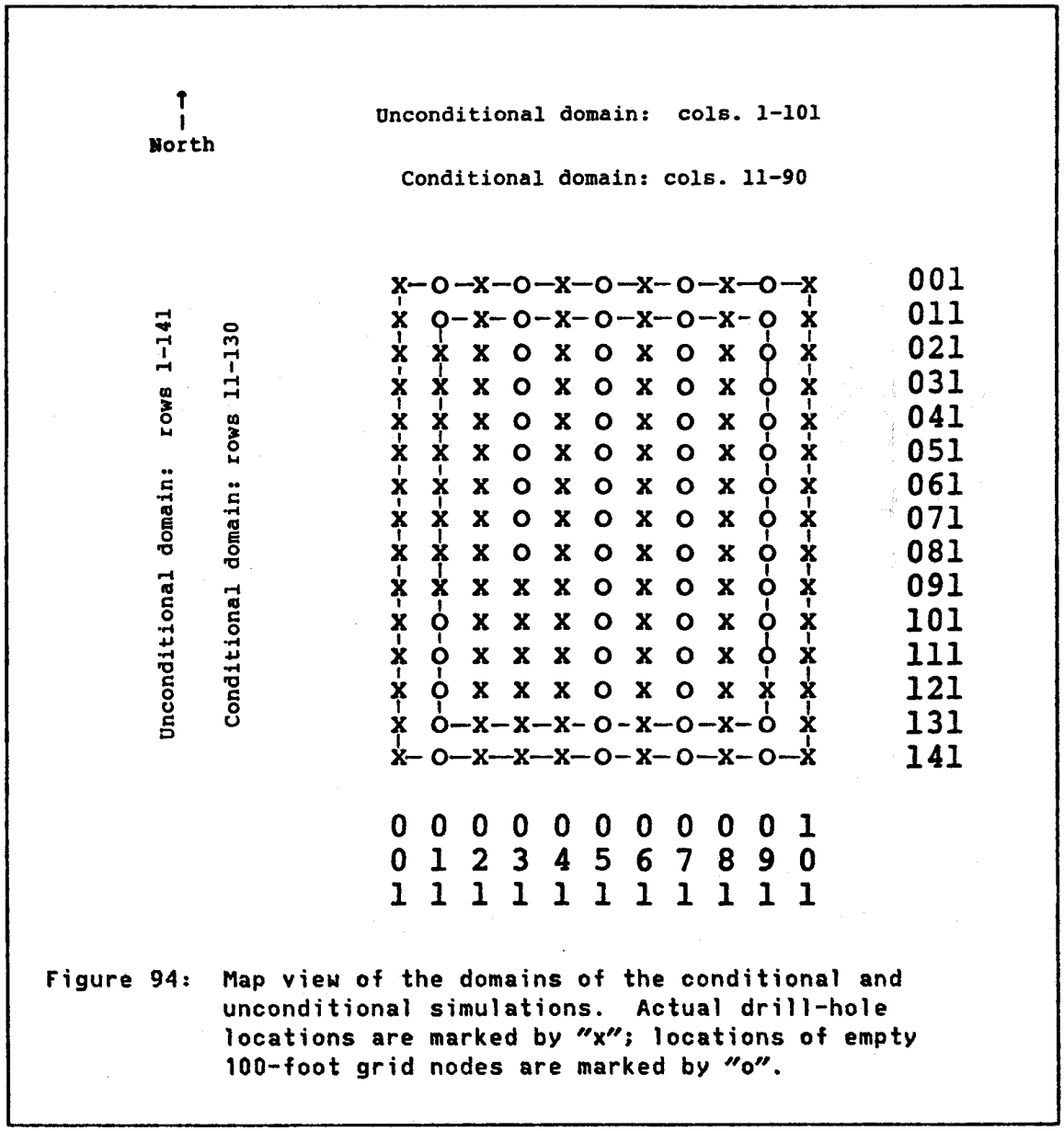


Figure 94: Map view of the domains of the conditional and unconditional simulations. Actual drill-hole locations are marked by "x"; locations of empty 100-foot grid nodes are marked by "o".

Kriging for conditioning was performed using linear semivariogram models, except that variable PGT(RATLS) was kriged using a spherical

model because the range of the sample semivariogram in the east-west direction was too short for an acceptable linear approximation. (A linear model is preferable to a spherical model for a large kriging job because the covariances can be calculated more rapidly.) The fact that the holes were drilled on a regular 100x100-foot grid (with several missing holes) allowed the kriging to be speeded up considerably. To take advantage of the regular grid, kriging was performed in two stages. In the first stage, only locations on the 100x100x3-foot data grid (Figure 94) were kriged (a bit more than 1% of the total points in the domain). Of the 15 (north-south) x 11 (east-west) x 41 (vertical) = 6765 grid nodes in this category, only 2048 were actually occupied by data, and only PGT(PCTDOL) was observed at every one of these 2048 data locations. Thus the first-stage kriging just filled in the gaps in the regular data grid. Then in the second stage the kriged estimates at these locations, combined with the actual data, were used to krige the central part of the domain (Figure 94) on a 10x10x3-foot grid spacing, using a constant data configuration (Figure 95). To obtain a constant kriging configuration over the whole domain, the top and bottom levels, first ten and last eleven columns, and first ten and last eleven rows of the unconditional domain had to be dropped from the conditional simulation, leaving a conditional domain of 39 levels, 80 columns, and 120 rows.

The estimates obtained on the fine grid using this technique were the same as those that would have been obtained if kriging had been done in a single pass using the original irregular data configuration in the same neighborhood (Journel and Huijbregts, 1978, p. 351), but the

regular configuration allowed the kriging weights to be calculated just once, in advance, for the fixed configuration, and then applied repeatedly to krige the entire grid. This saved a vast amount of execution time. For example, on Stanford's IBM 3084 the kriging of the original 6765 drilling-grid locations (including 2048 actual data locations) for variable PGT(PCTDOL) required over an hour[116] of processing ("CPU") time, whereas the final kriging of the 120 (north-south) x 80 (east-west) x 39 (vertical) = 374,400 data on the 10x10x3-foot grid required only about twenty minutes.

The conditionally simulated data were back-transformed to the original variables PCTDOL, CALS, MGLS, CADOL, and MGDOL by applying first a linear-interpolation transformation to invert the simulated normal scores to PCTDOL, TOTLS, RATLS, TOTDOL, and RATDOL, and then inverting the total and ratio transformations described in Section 4.2.2.6. Finally, the PCTDOL data were used to combine CALS and CADOL to get the overall $CaCO_3$ (CA) content of each simulated interval, and MGLS and MGDOL were similarly combined to get the $MgCO_3$ (MG) content. The final data set used in the mining simulations then consisted of three variables: PCTDOL, CA, and MG, describing the total lithologic and chemical makeup of each simulated three-foot-thick section of drill core.

---

[116] Actually much of this time was consumed by an inefficient method of searching for the data, but even with a much better search the 374,400 grid locations could not have been kriged in nearly so fast a time with irregular data. PGT(PCTDOL), having the most data, took the most time because of the search.

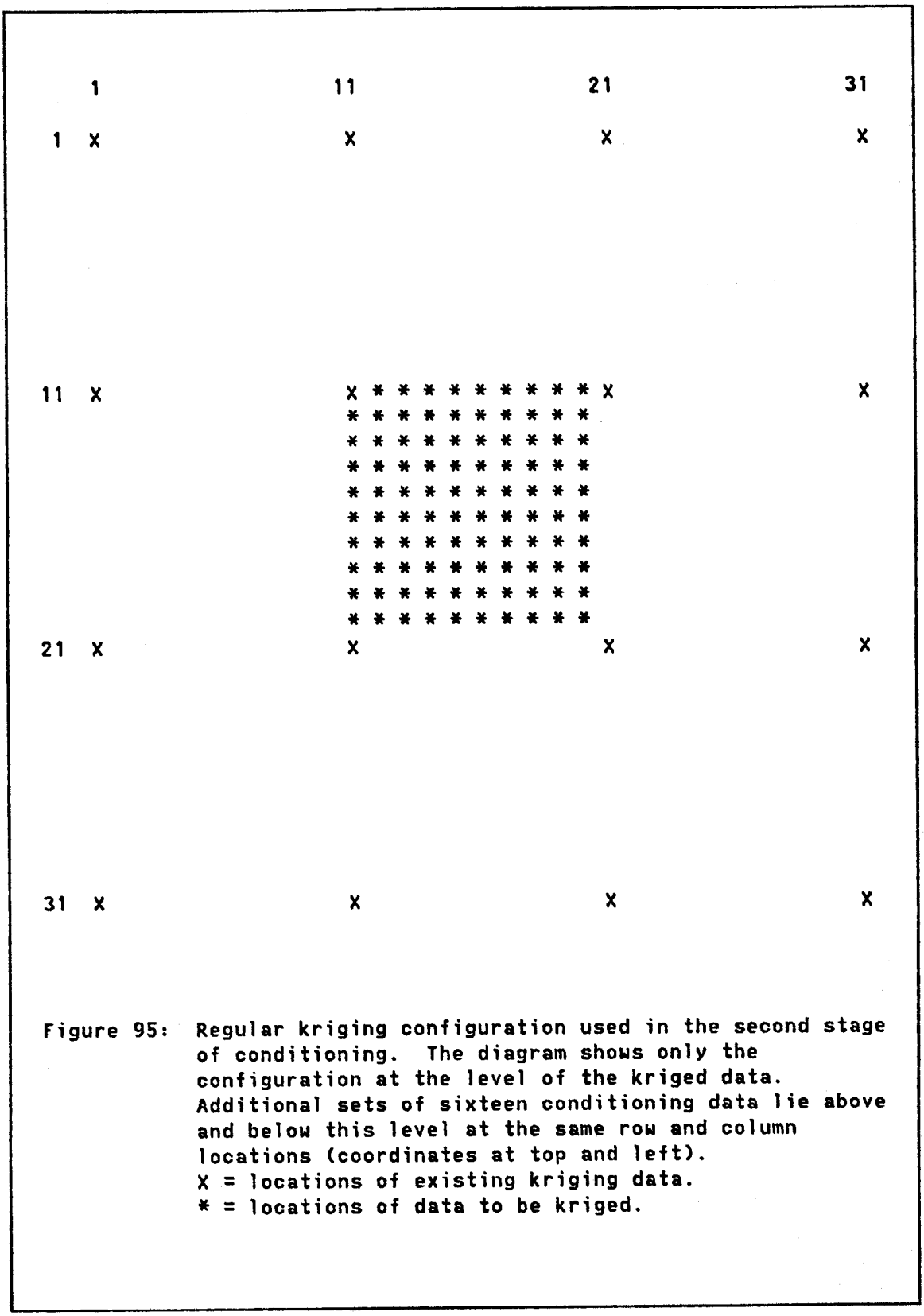


Figure 95: Regular kriging configuration used in the second stage of conditioning. The diagram shows only the configuration at the level of the kriged data. Additional sets of sixteen conditioning data lie above and below this level at the same row and column locations (coordinates at top and left).
X = locations of existing kriging data.
* = locations of data to be kriged.

The domain of the conditional simulation extends well above and below the elevations of the core-drill data used for conditioning, so many data in the upper simulated levels correspond to samples above the actual bedrock surface. The vertical coordinates of the top samples from the core-drill data were used to estimate the elevation of this bedrock contact within the horizontal limits of the simulation domain. Because the contact is naturally very smooth, there would be no point in trying to simulate it, as the differences between estimated and simulated values would typically be less than the three-foot vertical support of the simulated data. Inverse-distance-squared weighting was used to obtain an estimated surface elevation (in grid spacings) at the top of each of the 9600 simulated drill holes. This estimate was rounded to the nearest vertical grid coordinate, and all simulated values above this position were set to a missing-value code. Thus the final simulation extends upward only to the approximate level of the actual bedrock surface. This "simulated" surface is unrealistic in detail, as it varies in discrete three-foot jumps instead of continuously. But because the deposit is typically over fifty feet thick, and because the mining simulations use data averaged from groups of five or more holes, this approximation is of no consequence to the final results.

4.2.4 Simulation Results

Statistics. Table 11 compares statistics of the five conditionally simulated variables and the corresponding five variables observed at the original 2048 sample locations. In most cases the comparisons are

excellent. The larger discrepancies between real and simulated PCTDOL data arise from the influence of domain errors and from some oversmoothing committed by the despiking technique applied to this variable (discussed below). Domain errors result primarily from the fact that the original 2048 data were collected from a domain measuring (in the horizontal directions) 10,000x14,000 feet (101 columns x 141 rows at a 10x10-foot grid spacing in the unconditional simulation), whereas the conditional simulation occupies only columns 11 through 90 and rows 11 through 130. Variable PCTDOL, having a long range, a relatively low nugget component, and a high sill, is more sensitive to domain errors than the chemical variables; thus the statistics for PCTDOL differ more noticeably among the three data sets. The lower standard deviations of the simulated PCTDOL data probably arose mainly from smoothing (described below).

Statistics were not calculated for all simulated data, but rather for two subsets extracted from every tenth row and column starting with coordinate (11,11) for Subset 1 and (16,16) for Subset 2. Subset 1 is "in phase" with the drilling grid and contains many conditioning data; Subset 2 is shifted to the southeast and contains no conditioning data. In the vertical direction, the two subsets of simulated data were collected only between the bedrock surface and an elevation six feet (two three-foot vertical grid spacings) below the top of the lower dolostone unit that underlies the quarry floor. This is about the same vertical interval covered by the core drilling. Elevations of the dolostone contact were determined as described in the discussion of mining simulations in Section 4.2.5.

TABLE 11

Statistical summaries of 2048 real data and two sets of 1987 and 1993 simulated data.

Variable	Real Data		Simulated Data (Subset 1)		Simulated Data (Subset 2)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
PCTDOL	29.5	42.2	21.7	37.3	20.6	35.8
CALS	95.8	2.6	95.8	2.6	96.0	2.3
MGLS	1.2	1.7	1.2	1.8	1.2	1.8
CADOL	59.2	7.2	58.4	7.2	59.0	7.2
MGDOL	34.0	5.8	32.8	6.3	32.9	6.1

Correlation Coefficients (/ Real Data, Subset 1, \ Subset 2)

	PCTDOL	CALS	MGLS	CADOL
	-0.055			
CALS	0.007			
	0.102			
	0.049	-0.727		
MGLS	0.028	-0.722		
	-0.028	-0.729		
	0.052	0.089	0.110	
CADOL	0.057	0.118	-0.032	
	0.101	0.114	-0.043	
	-0.002	0.086	-0.096	-0.434
MGDOL	0.128	0.023	0.000	-0.310
	0.083	0.008	0.023	-0.461

Scatterplots. Figure 96 shows scatterplots of the real data and of the simulated data from Subsets 1 and 2. The general shapes of the bivariate distributions of $CaCO_3$ and $MgCO_3$ in the two rock types have been well reproduced, as a consequence of the "ratio" and "total"

transformations applied to the data. The few anomalies that appear in the simulated data, particularly in the dolostone compositions, owe their existence to the influence of a few synthetic conditioning data, discussed in Section 4.2.2.2. The synthetic conditioning data were used not only in conditioning but in the forward and inverse ratio and total transformations. (At least a few of them must be used as bounds in the transformations in order to reproduce the conditioning data.) These anomalies may have had some influence on semivariogram reproduction as well, because they were not included in the data sets used for semivariogram modeling but were used in conditioning.

Semivariograms of PCTDOL. Figure 97 compares two sets of east-west sample semivariograms of PCTDOL. The semivariogram of the original data (marked by the letter "B") is the same in both plots. It was calculated only from data within the domain of the conditional simulation. In plot (a), the simulated data in Subset 1, with horizontal grid origin (11,11), were used to plot semivariogram "A". This grid coincides with the drilling grid, so many of the data used in semivariograms "A" and "B" were identical, and so the sample plots are similar. In plot (b), the simulated data in Subset 2, shifted 50 feet (five grid spacings) south and east of the drilling grid, were used to plot semivariogram "A". The greater departure between the two curves is striking.

For most directions the sample semivariograms of the simulated PCTDOL data (even those of Subset 1) are systematically lower than those of the real data, as though the nugget effect had been smoothed out. This was not the case in the semivariograms of the "gaussian" PGT(PCTDOL)

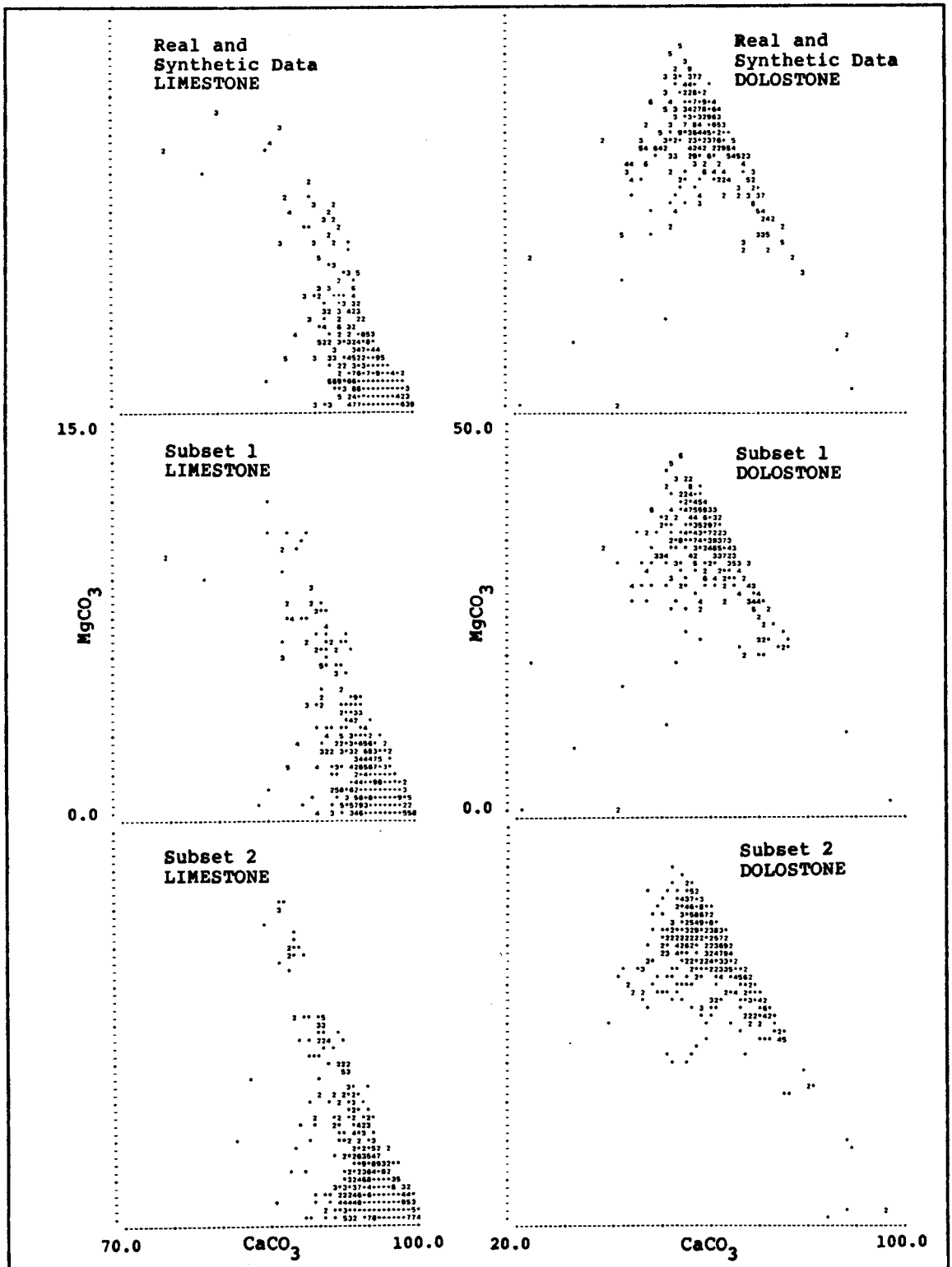
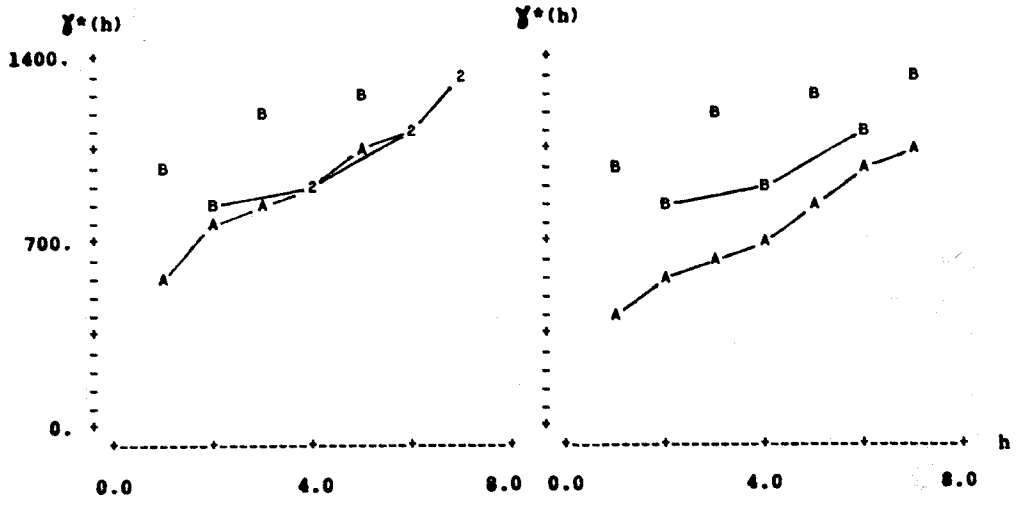


Figure 96: Scatterplots of $CaCO_3$ and $MgCO_3$ for real and conditionally simulated limestone and dolostone.



(a) Subset 1 (b) Subset 2

Figure 97: Sample semivariograms of the percentage of dolostone in two subsets of the conditional simulation. Plot "A" represents the simulation, plot "B" (which is the same in both diagrams) the real data. Only the even-numbered lags of plot B (connected by lines) are reliable because of the "two-sample" effect discussed in Section 4.2.2.5.

conditioning data and of the resulting simulations, before application of the inverse transformation. Those data showed no nugget effect. It appears that the forward-transformed "gaussian" conditioning data did not resemble a true realization of a stationary gaussian random function very closely, as in that case the inverse transformation would be forced, in theory, to reproduce the spatial distribution, and thus the semivariogram, of the original data. The likely culprit is the despiking formula on page 402, which tended to produce extremely smooth-looking "gaussian" data in large areas of pure limestone or dolostone,

whereas data near the contacts were little affected. Apparently one effect of this smoothing was obliteration of the nugget effect in the sample semivariograms of the PGT(PCTDOL) data. In the back-transformed conditional simulation of PCTDOL, the nugget effect is partly restored in Subset 1, which contains many conditioning data; but because conditioning can do little to influence short-scale semivariogram structure, the sample semivariograms of Subset 2 show a short-scale bias in the simulation.

Another practical effect of the overly smooth PCTDOL simulation is that areas completely surrounded by dolostone conditioning data are typically devoid of simulated limestone lenses, and areas of pure limestone conditioning data contain no simulated dolostone. A rougher unconditional simulation and rougher conditioning data would be expected to insert a few lenses of dolostone in places where core drilling had encountered none. In the mining simulations (Section 4.2.5), this effect would, on the average, make very dolostone-contaminated mining blocks look even worse, and make nearly pure limestone blocks look even better. This might make selective mining look a bit more effective than it should.

Fortunately, the simulation of PCTDOL is so tightly conditioned that these small-scale effects cannot persist far enough to be of great practical importance, and large-scale deviations from the truth cannot occur at all. Because the mining simulations involve blocks containing five holes each, which subsequently are lumped into homogenization piles of up to 24 blocks, the bias in short-scale structure should be of little consequence. This can be viewed in terms of a regularization:

two phenomena differing only in the sizes of their short-scale structures will become indistinguishable when regularized over a larger support, as regularization subtracts the short-scale structure from each of them.

If this simulation were to be used without being averaged into blocks[117], the short-scale inaccuracy of the PCTDOL semivariograms would be impermissible. In that case, a different despiking approach would have to be devised that would not produce such smooth "gaussian" data. The smoothness of the method used here arises mainly because the method is deterministic given the data (like a kriging), whereas what is really needed is a simulation. Because the PCTDOL data resemble indicator data so closely, the methods suggested in Section 3.8.3.4 for generating gaussian conditioning data for indicator simulations are applicable. In particular, method (5) proposed on page 297 is very similar to the approach used here, except that the suggested step of "roughening up" the smooth conditioning data was not, but should have been, taken. The appropriate amount of roughening (probably accomplished by addition of some nugget noise to the despiked values) would have to be determined by trial and error. The effects of the roughening on the appearance of the PCTDOL cross sections (Figure 101) might have to be considered in selecting the structure of the roughening data.

---

[117] Recall, however, that the data are on "core" support, so for most mining applications it would be inappropriate to use them without first averaging the data into blocks. Fortuitously, the overly smooth simulated data probably look much like the data that would be obtained by a change of support from three-foot core samples to their associated 3x10x10-foot simulation blocks.

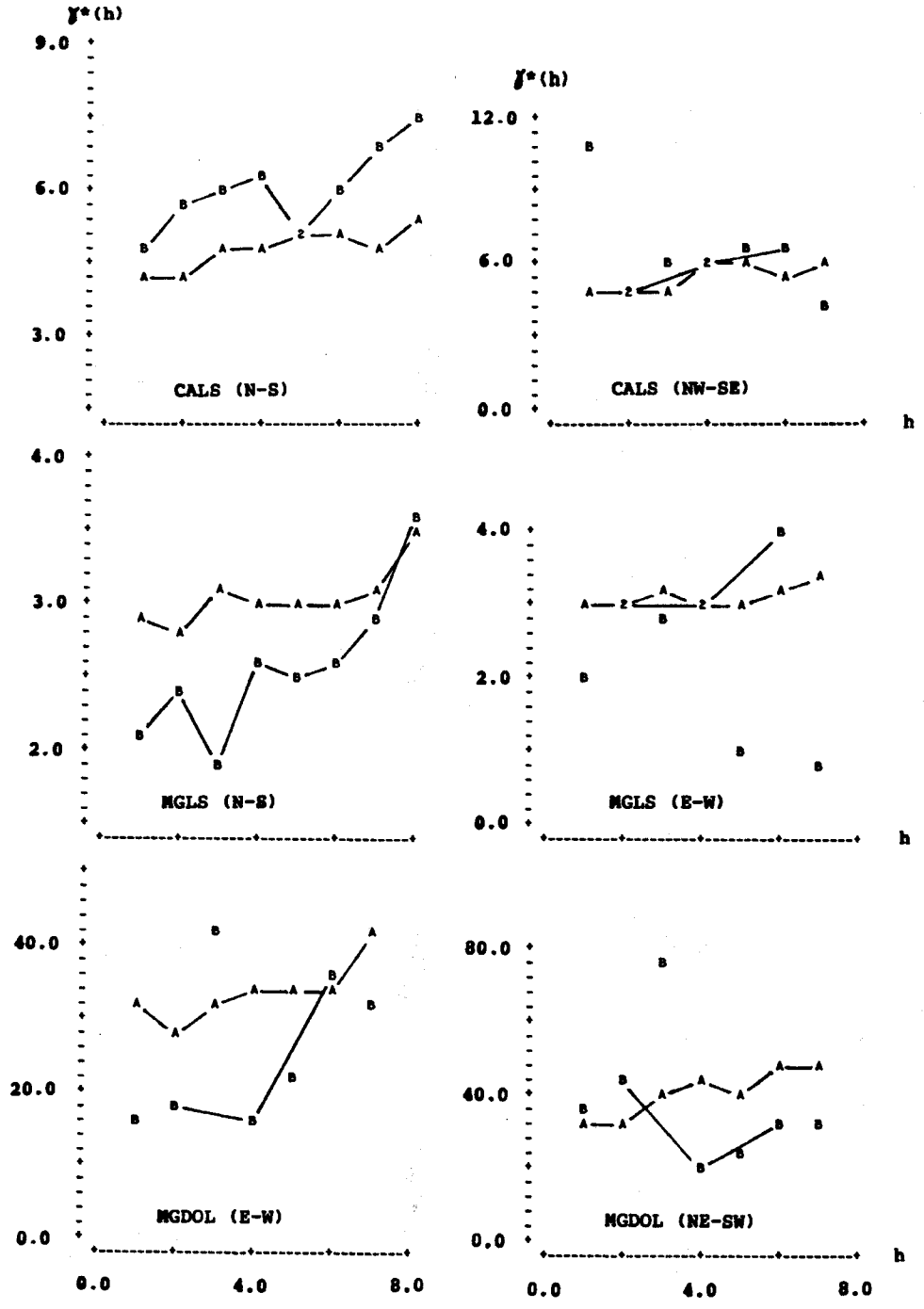
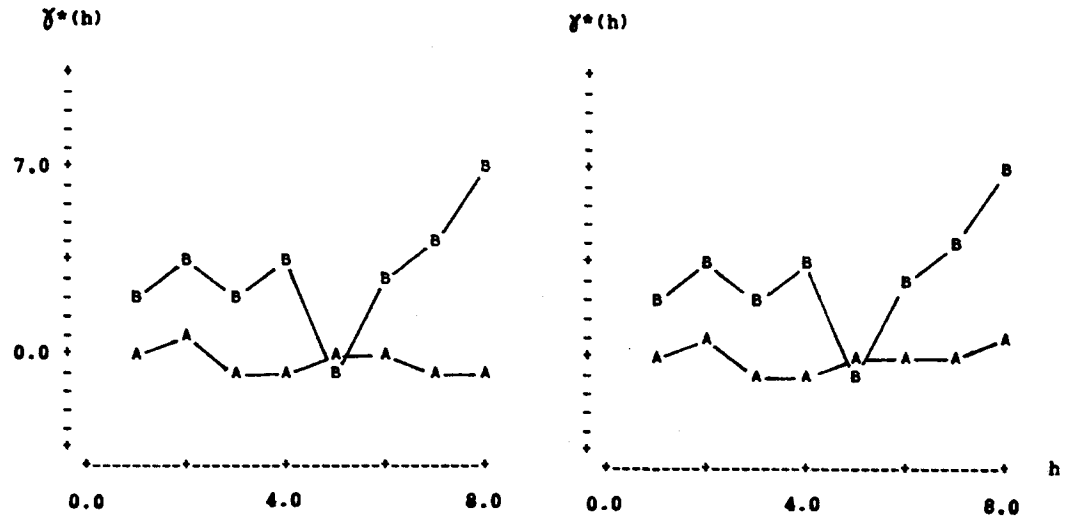


Figure 98: Selected sample semivariograms of the chemical variables. Plots "A" represent the simulation, plots "B" the real data. Odd-numbered lags of plots "B" are unreliable (Section 4.2.2.5), except north-south.

Semivariograms of chemical variables. Some typical sample direct semivariograms of the chemical variables are illustrated in Figure 98. All of the semivariograms of the simulated data were calculated from Subset 2. The discrepancies that appear occur primarily in the nugget structure and vary greatly with direction. These discrepancies merely reflect the fact that the semivariograms of the real chemical variables are poorly estimated by the available data, which tend to suggest different models in different directions. Models of the gaussian data were based upon the average behavior of the models in all directions (allowing for anisotropy). Because the variability among directions was pronounced in the nugget structures, conditioning was not able to bring the semivariograms of the simulated data entirely back into line, although the large-scale trends in curves "A" and "B" tend to be parallel as the result of conditioning.

In some cases poor reproduction reflects the paucity of data used to calculate the original semivariograms, as illustrated by the north-south plots for CALS and MGDOL shown in Figure 99. The gaussian transforms of these variables were modeled as independent, even though the very ragged sample plots in some directions suggested various types of cross structure, especially at long lags. (Apparent cross structure at long lags can usually be attributed either to drifts in the two variables or simply to a small number of pairs.) These cross structures would normally be imposed by conditioning, but it is obvious from Figure 99(a) that the cross structure "A" of the simulated variables is still just what the model of the unconditional simulation said it should be: zero. The "A" plot for Subset 1 in Figure 99(b) hardly looks better, even

though several of the same data were used for both "A" and "B". Improvement in (b) is particularly negligible here because only 255 of the original 2048 data, and even fewer of the data inside the conditional-simulation domain, contained both CALS and MGDOL values.



(a) Subset 2 (b) Subset 1

Figure 99: Sample cross semivariogram between variables CALS and MGDOL. Plots "A" represent the simulation, plots "B" the real data.

Semivariograms of total rock compositions. The variography of the simulated chemical compositions of the limestone and dolostone combined is most important in practice, as the mining simulations will be carried out on these data. Figure 100 shows plots for total $CaCO_3$ (variable CA) and total $MgCO_3$ (variable MG) in the raw material. All plots are based upon simulated Subset 2. Not surprisingly, these plots appear to be

dominated by the smooth structure of the PCTDOL simulation, which determines how much simulated limestone and how much simulated dolostone contribute to each total analysis. In general, plots "A" and "B" are similar but differ by a fairly constant value owing mostly to the smoothness of the PCTDOL data: the simulation plot "A" is systematically closer to zero than the data plot "B". Again, the effects of these discrepancies are largely filtered out in the mining and homogenization simulations, because data from several simulated "holes" (at least five holes, with usually ten to twenty simulated analyses per hole) contribute to the analysis of each mined-out and homogenized batch of stone.

Cross sections. To see that the percentage of dolostone in the samples appears geologically reasonable, eight north-south cross sections were plotted through the PCTDOL simulation at intervals of 100 feet, starting with column 11 (of the 101 columns in Region 2, the unconditional domain) and continuing to column 81. Four of these -- at columns 21, 41, 61, and 81 -- are illustrated in Figure 101. They show a realistic-looking lenticularly bedded deposit, much like the sketches in Figure 88. Dolostone is concentrated in the upper part of the section in the southwestern corner of the domain, and in the bottom of the section (just below the level of the present quarry floor) everywhere. The great thickness of the bottom dolostone in the cross sections is a figment of extrapolation by kriging, abetted by the smoothness of the PCTDOL despiking procedure. Most core-drill holes actually extended only a few feet into the top of this unit, so that its true thickness is unknown. In the simulation, the only reason that the

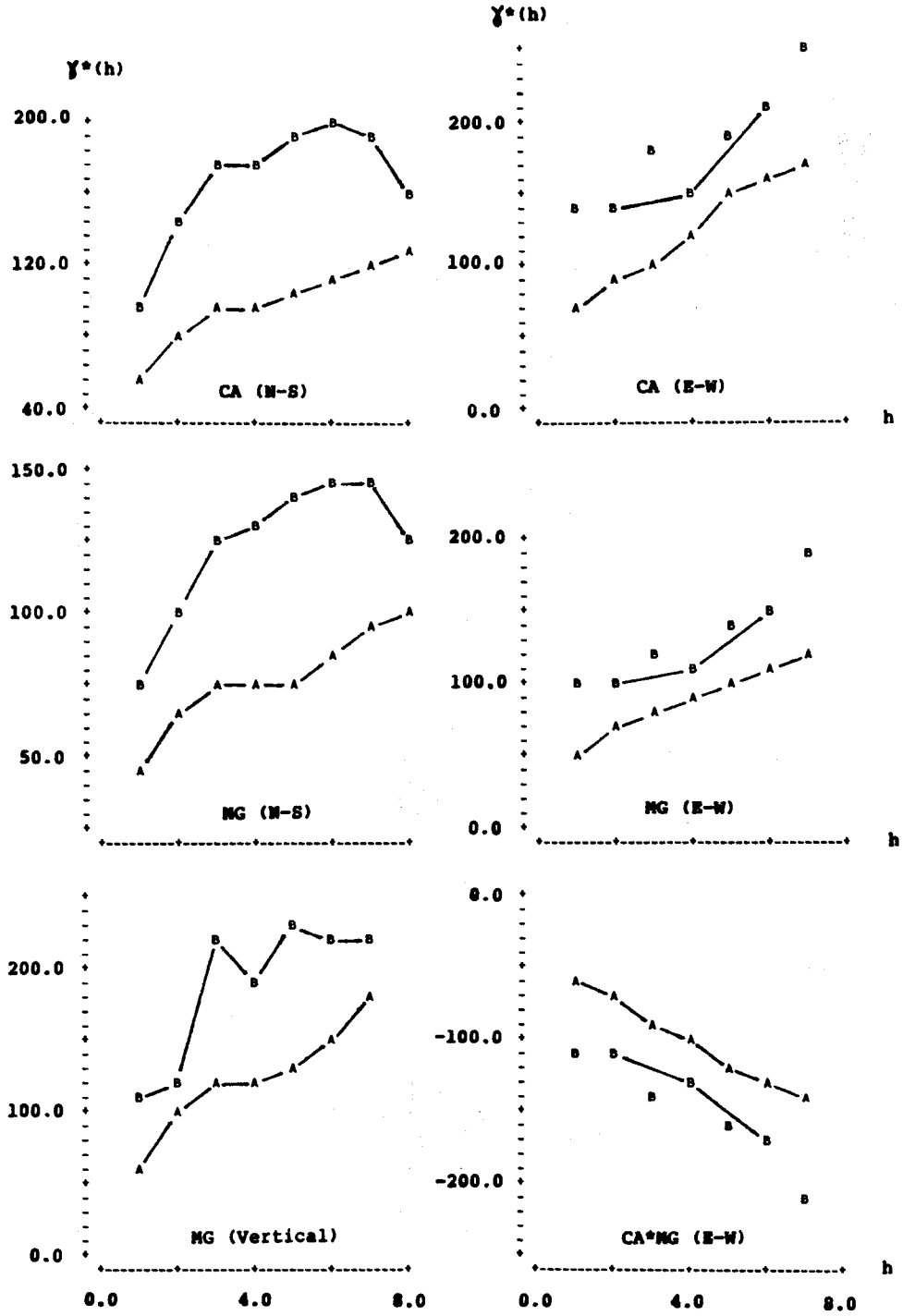


Figure 100: Sample semivariograms for the total $CaCO_3$ and $MgCO_3$ contents of limestone and dolostone combined. Plots "A" denote the simulation, plots "B" the real data. Odd-numbered lags of east-west plots "B" are unreliable (Section 4.2.2.5).

unit cuts off at all at the bottom of each section is that the downward extent of the neighborhood used in the first stage of kriging was exceeded. The artificial thickness of this unit in the simulation is of no practical importance, as the mining simulations attempt to avoid it entirely.

4.2.5 Mining and Homogenization Simulations

Mining and homogenization simulations can be made rather sophisticated, as in the mining simulations described by Deraisme and de Fouquet (1984) and the detailed simulation of a linear stacker-reclaimer system undertaken by Schofield (1980, Chapter 4). Only very elementary simulations have been performed for this case study. Nevertheless the results clearly suggest some approaches that would be most helpful in dealing with the variations in raw-material quality that will be encountered during the development of this part of the property.

4.2.5.1 Mining Alternatives

Four mining alternatives and four homogenization alternatives -- sixteen combinations in all -- were tried out on this simulated deposit. In all mining alternatives, the deposit was mined out in eighty ten-foot-wide slices, whose compositions were represented by single columns of 120 simulated "drill holes". The cross sections of Figure 101 represent four such columns of holes. Each slice of 120 holes was divided into 24 groups of five consecutive holes (Figure 102), and the average composition of the simulated values within each group was considered to be a fair representation of the average composition of the associated

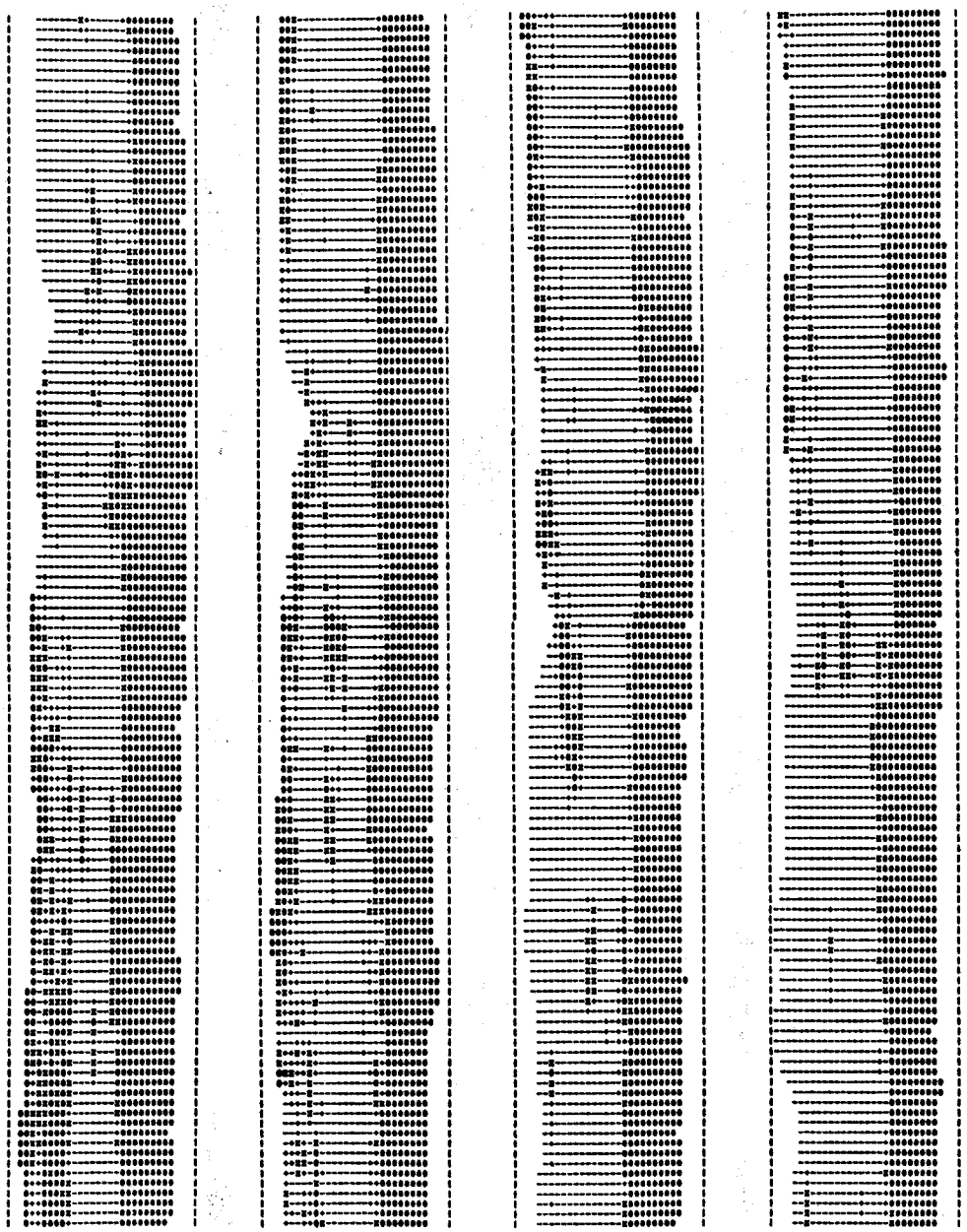


Figure 101: Four north-south cross sections depicting the simulated percentage of dolostone at each grid location. The figure should be viewed from the right side of the page. The symbols are: "-" = pure limestone, "+" = mostly limestone, "X" = mostly dolostone, "#" = pure dolostone.

block. As the simulated holes are arranged on a horizontal grid with 10x10-foot spacings, each simulated block of five holes represents the stone mined from within a ten-foot (east-west) by fifty-foot (north-south) area of the quarry property. Most of the actual mining blocks (individual "shots") currently being recovered in the quarry measure about 25 x 400 feet horizontally, so the smaller simulated blocks allow more selectivity in the horizontal directions. These smaller blocks do not necessarily represent individual shots: several adjacent blocks might be shot simultaneously, but different parts of the resulting pile of broken stone could then be recovered separately.

Selectivity alternatives. Two alternative levels of selectivity were allowed in the vertical direction. One alternative represents current mining practice: the simulated analyses (each representing a three-foot-thick sample of "drill core") from each hole are recovered down to about three feet above the thick lowermost dolostone that underlies the present quarry floor (clearly visible in the sections of Figure 101). Any dolostone lying in the upper part of the section is incorporated into the mined raw material. This "nonselective" alternative works well in those parts of the quarry where dolostone is confined to the bottom of the section; in the absence of large-scale homogenization, other areas of the property must be avoided, as is now the practice. The other, "selective", alternative avoids both the upper and lower dolostones and is similar to current practice in the nearby quarry of Northwestern States Portland Cement. The two alternatives are compared in Figure 103, which shows part of a cross section similar to those in Figure 101.

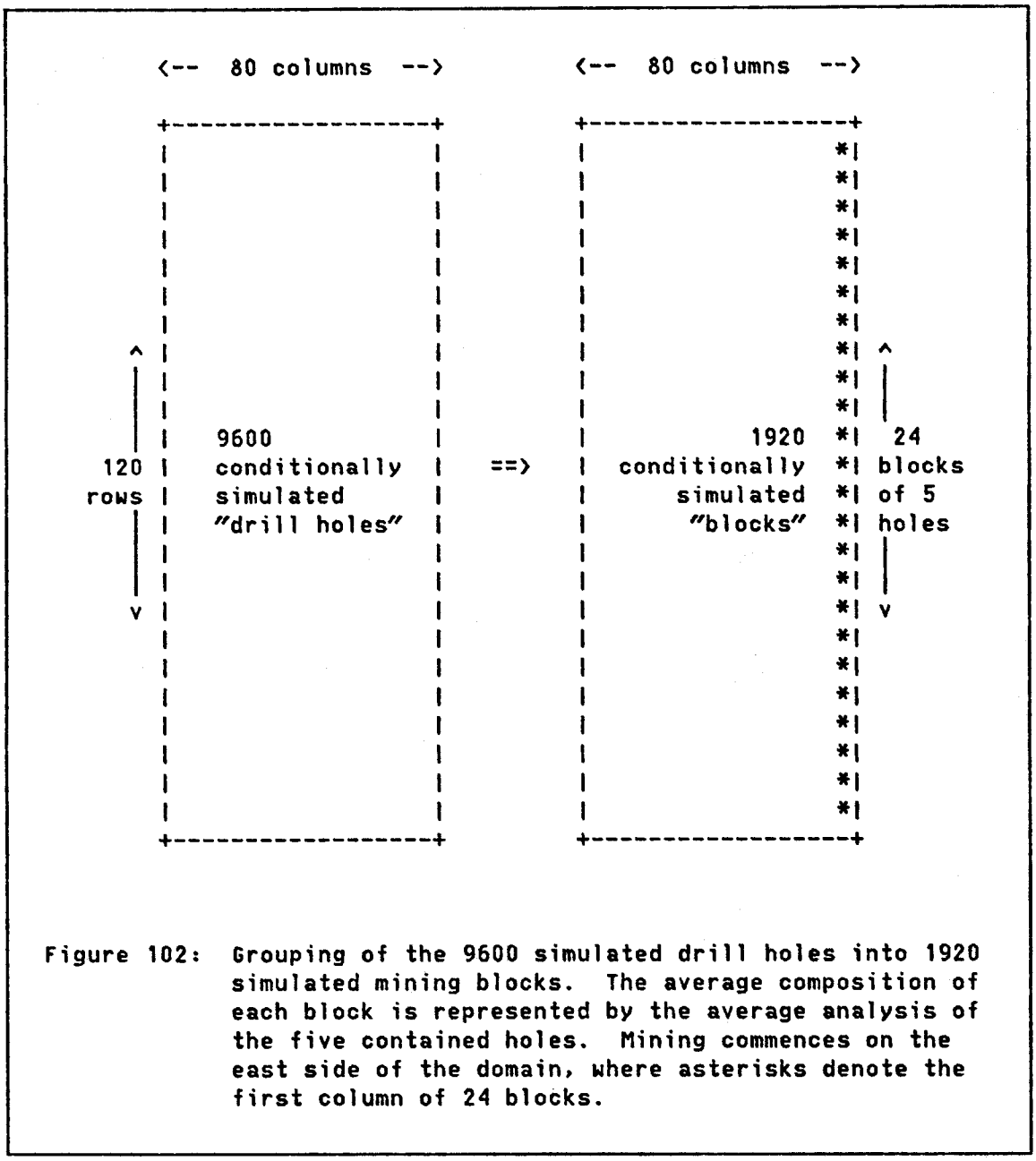


Figure 102: Grouping of the 9600 simulated drill holes into 1920 simulated mining blocks. The average composition of each block is represented by the average analysis of the five contained holes. Mining commences on the east side of the domain, where asterisks denote the first column of 24 blocks.

These selectivity alternatives were programmed very simply. The simulated data for each variable were first stored in a "holewise" sequence, top to bottom; i.e., the order of matrix storage used by the kriging program and by subroutine TB3D (Appendix B) -- "rows within

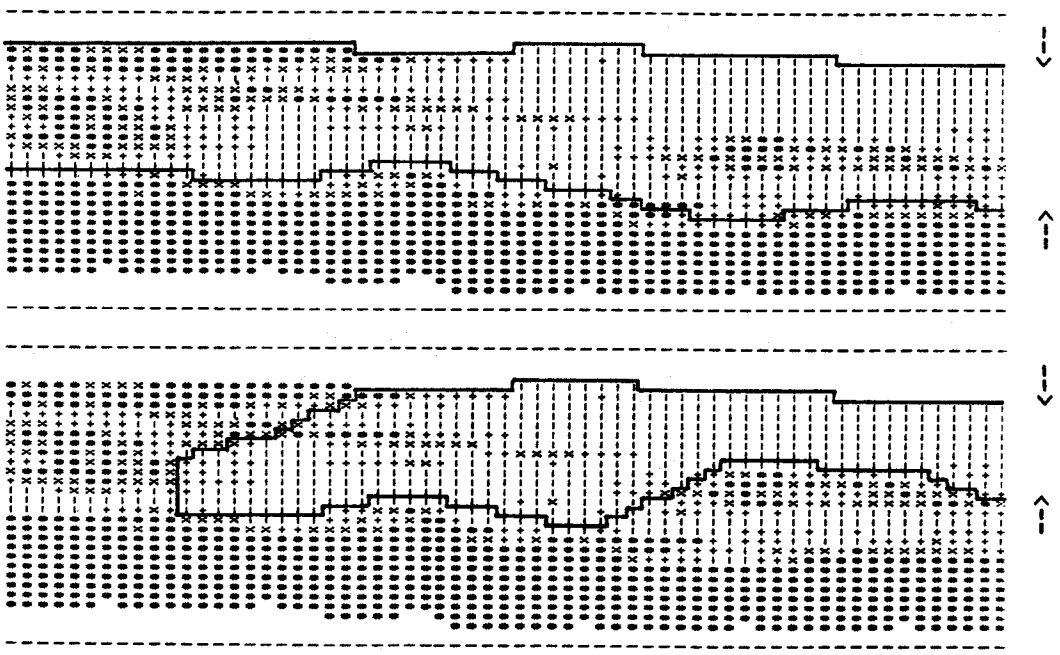


Figure 103: A cross section showing the stone recovered under each of the two selectivity alternatives. The upper diagram shows recovery (between the arrows) under the "nonselective" alternative. The lower diagram shows "selective" recovery. Symbols are the same as those in Figure 101.

columns within levels" -- was altered to "levels within rows within columns". Then a series of cross sections of the percentage of dolostone was generated along eight of the eighty columns (columns 11,21,...,81, using the numbering of the unconditional domain). Four of these sections (columns 21, 41, 61, and 81) are displayed in Figure 101. Every tenth "hole" within each section (corresponding to rows 11,21,...,121) was inspected to find the vertical coordinate of the top of the lowermost dolostone. This top was defined as the coordinate of the highest simulated core sample in the lower part of the section with

at least 50% dolostone (i.e., having an 'X' or '#' symbol on the section). Only 1% of the tops (i.e., tops from 96 of the 9600 conditionally simulated holes) were picked by eye from the cross sections in this way; tops in the other holes were estimated by inverse-distance-squared weighting and rounded to the nearest integer coordinate value.

For the "nonselective" mining alternative, the samples to be "mined out" from each hole are those at least two grid points above this estimated top, such that one sample with less than 50% dolostone is left in the quarry floor. Nevertheless, because the estimated top is very smooth whereas the actual top is locally uneven, a rare bit of lower dolostone will occasionally be included in the mined stone (as shown in Figure 103). This happens in much the same way in the actual quarry, as the floor is kept smooth and an attempt is made to stay only a few feet above the lowermost dolostone.

In the "selective" mining alternative, the bottom of the upper dolostone, where an upper dolostone is present and at least six feet thick, was estimated in the same way. In this case the limestone immediately underlying the estimated dolostone bottom was assumed to be good material. No three-foot "safety margin" was left between "ore" and "waste" because it is more expensive to waste good ore than simply to leave it in the ground, as was done near the top of the lowermost dolostone. The bottom of the recovered limestone in the selective alternative was the same as that in the nonselective alternative (i.e., about three feet above the lower dolostone), except in a few locations where thick dolostone lenses appeared in the middle of the section

(e.g., in the right side of the lower section in Figure 103), in which cases the floor of the quarry was raised. In the selective alternative no stone was recovered when the thickness between the upper and lower contacts dropped below 18 feet (six simulated core samples); this was felt to be the thinnest limestone worth recovering (e.g., no stone is recovered on the left side of the lower section in Figure 103).

The main difference between the simulated selective operation and a real operation that might be undertaken in this area of the quarry is that the limits of the mined stone are smooth estimates (migrating in three-foot hops) in the simulation, whereas in an actual operation a constant thickness of dolostone would probably be removed over large areas of the quarry, and then cut off suddenly. This difference is illustrated in Figure 104. However, after the individual holes are grouped into blocks of five holes and the blocks are eventually stacked onto a homogenization pile, the correspondence between the simulated pile compositions obtained in this way and the actual compositions that would be recovered from the same area should be very good.

Because the simulated values were stored "holewise", the limestone was easily "mined out" by a simple computer program that read the compositions in each hole and then averaged the compositions lying between the previously determined upper and lower mining limits. The recovered tonnage within the 10x10-foot area surrounding each hole (about 23 short tons per three-foot vertical increment) was accumulated at the same time, and then the weighted-average compositions and total tonnages within each five-hole mining block were similarly calculated and stored.

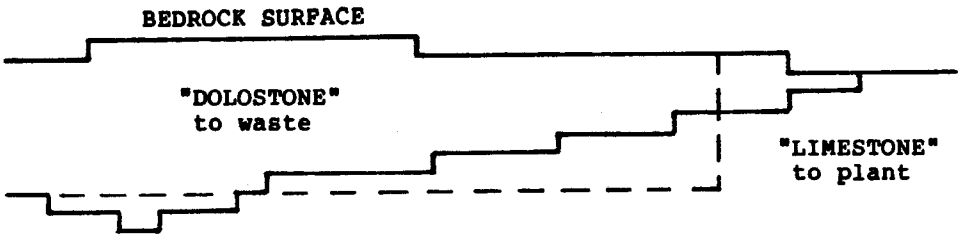


Figure 104: The likely difference between simulated and actual mining cutoff surfaces. Each simulated surface (solid line) resembles a surface that might be produced by ripping, but hard limestone and dolostone would more likely be drilled and blasted away in large rectangular chunks (broken line).

Recovery alternatives. In addition to the two selectivity alternatives, two sequences of block recovery were considered. In both sequences, blocks were recovered in ten-foot-wide, north-south slices, starting within the existing mined-out quarry area on the east and working westward toward the property boundary. Several blocks from the eastern slices correspond to stone that has actually been mined in the past. The western, much more dolomitic, stone has so far been avoided. Within each slice, blocks were recovered either "consecutively" south-to-north, or "nonconsecutively", systematically taking one block and skipping the next three, as illustrated in Figure 105. Because the amount of dolostone in most slices varies fairly regularly from south to north (see Figure 101), the consecutive alternative should produce only slight changes in the composition of the mined ore from one block to the next within a slice, but a large cumulative change over the length of

the slice and a large sudden change between the north end of one slice and the south end of the next slice westward. The nonconsecutive alternative, by contrast, will typically result in larger changes block-to-block. In the time series of mined block grades, the consecutive alternative would concentrate variability at lower frequencies (i.e., a hole-effect variogram or autocorrelation function with a long period), whereas the nonconsecutive alternative would shift much of this variability to higher frequencies.

4.2.5.2 Homogenization Alternatives

Four batch homogenization alternatives (Figure 105) were applied to each of the four mining alternatives (two selectivity alternatives, two recovery alternatives). The homogenization alternatives are: (1) homogenize each block internally; i.e., no homogenization beyond the scale of the five-hole blocks; or (2) homogenize in piles of six blocks, (3) twelve blocks, or (4) 24 blocks (an entire north-south slice).[118] It is assumed that a sufficient number of layers (say 100) will be built within each pile so that the material reclaimed from each pile will have essentially constant composition. (At least, any variations in composition should be slight and rapid enough to be filtered out by the powder homogenization system.) Within the total simulation domain there are 1920 five-hole blocks, 320 piles of six blocks, 160 piles of twelve blocks, and eighty piles of 24 blocks. The four mining alternatives and

---

[118] Notice that the simulated stockpiles will not all contain the same tonnage, particularly in the case of selective mining. In an actual operation, all piles would probably be built to the same size. This discrepancy was tolerated in the interest of much simpler programming.

Recovery sequence:	CONSECUTIVE				NONCONSECUTIVE			
Number of piles from one slice:	24	4	2	1	24	4	2	1
Number of blocks in each pile:	1	6	12	24	1	6	12	24
Pile numbers of recovered blocks:	24	4	2	1	24	4	2	1
	23	4	2	1	18	3	2	1
^	22	4	2	1	12	2	1	1
|	21	4	2	1	6	1	1	1
|	20	4	2	1	23	4	2	1
NORTH	19	4	2	1	17	3	2	1
	18	3	2	1	11	2	1	1
	17	3	2	1	5	1	1	1
	16	3	2	1	22	4	2	1
	15	3	2	1	16	3	2	1
	14	3	2	1	10	2	1	1
	13	3	2	1	4	1	1	1
	12	2	1	1	21	4	2	1
	11	2	1	1	15	3	2	1
	10	2	1	1	9	2	1	1
	9	2	1	1	--> 3	1	1	1
	8	2	1	1	20	4	2	1
	7	2	1	1	14	3	2	1
	6	1	1	1	8	2	1	1
	5	1	1	1	--> 2	1	1	1
	4	1	1	1	19	4	2	1
	--> 3	1	1	1	13	3	2	1
	--> 2	1	1	1	7	2	1	1
	--> 1	1	1	1	--> 1	1	1	1

Figure 105: Recovery and homogenization alternatives illustrated on a single north-south slice of 24 mining blocks. Each column of 24 numbers represents a north-south slice of 24 ten-foot-wide mining blocks in the quarry. The arrows show the first three blocks recovered under the consecutive and nonconsecutive alternatives. Each number in each column represents the number of the pile onto which the block at that position will be stacked. For example, the 18th block to be recovered consecutively would be the 11th block recovered nonconsecutively. This block would enter the 3rd six-block stockpile if recovered consecutively, the 2nd if recovered nonconsecutively. The objective of nonconsecutive recovery is to combine blocks from widely scattered parts of the active face into a single stockpile.

four homogenization alternatives produce a total of sixteen[119] time series of homogenized stone compositions. The characteristics of these sixteen series reveal much information about the pile-to-pile variability of the homogenized ore. In a plant with a powder (or slurry) homogenization system (i.e., almost any cement plant), the pile-to-pile (low-frequency) variability is by far the most important characteristic to be assessed for plant-design purposes. (See the quotation that opens Section 3.9.2.1.)

4.2.5.3 Results

> The completion of a piece of research is not achieved until the results are at least communicated, and often, particularly in "useful" research, the results must be implemented before the research can be considered complete. [Griffiths, 1967, p. 483]

Statistics of the sixteen simulated time series are compiled in Tables 12 and 13, for nonselective and selective mining alternatives, respectively. The upper part of each table shows statistics calculated over the whole field of 1920 mining blocks. Because this area is so heterogeneous, statistics have also been compiled for four sets of twenty north-south slices, grouped east-to-west across the area. (The four cross sections of Figure 101 run down the middles of these four subareas.)

For the total area, the statistics reported are: the total number of homogenized piles; the mean tonnage per pile; the minimum, mean, and maximum $CaCO_3$ and $MgCO_3$ contents of the piles; the standard deviations of these contents; and (very important) the minimum, mean, and maximum

---

[119] Actually only fourteen, as the data for piles of 24 blocks are identical for consecutive and nonconsecutive recovery alternatives.

TABLE 12

Statistical summary: nonselective mining.

Recovery Sequence:	CONSECUTIVE				NONCONSECUTIVE			
Pile Size (blocks):	1	6	12	24	1	6	12	24
Number of Piles:	1920	320	160	80	1920	320	160	80
TOTAL DOMAIN								
Mean Pile Tons:	1700	10198	20396	40793	1700	10198	20396	40793
min.:	67.0	76.4	77.1	83.2	67.0	82.8	83.0	83.2
$CaCO_3$ (CA) max.:	97.4	97.1	96.9	96.0	97.4	96.3	96.2	96.0
mean=91.0 s.d.:	5.3	5.0	4.8	3.6	5.3	3.7	3.7	3.6
min.:	0.1	0.2	0.6	1.4	0.1	1.1	1.2	1.4
$MgCO_3$ (MG) max.:	23.5	16.4	16.1	11.0	23.5	11.3	11.1	11.0
mean=5.2 s.d.:	4.1	3.9	3.8	2.6	4.1	2.6	2.6	2.6
Absolute differences between successive piles:								
min.:	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
$CaCO_3$ (CA) mean:	1.7	3.4	4.6	0.5	3.3	0.7	0.7	0.5
max.:	15.2	16.4	13.7	1.7	19.8	3.2	2.5	1.7
min.:	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$MgCO_3$ (MG) mean:	1.4	2.9	4.0	0.4	2.8	0.6	0.6	0.4
max.:	12.7	13.0	11.3	1.5	15.9	2.4	2.0	1.5
SUBDIVISIONS OF THE TOTAL DOMAIN								
Eastern Quarter								
Mean Pile Tons:	1769	10612	21224	42449	1769	10612	21224	42449
CA(m=94.9) s.d.:	1.8	1.3	1.2	0.8	1.8	0.8	0.8	0.8
MG(m= 2.4) s.d.:	1.6	1.3	1.2	0.7	1.6	0.7	0.7	0.7
CA abs.dif., mean:	0.8	1.3	1.9	0.3	1.5	0.4	0.4	0.3
MG abs.dif., mean:	0.7	1.3	1.8	0.3	1.4	0.4	0.4	0.3
Mideast Quarter								
Mean Pile Tons:	1726	10358	20717	41433	1726	10358	20717	41433
CA(m=93.6) s.d.:	2.1	1.4	1.1	0.9	2.1	1.1	1.0	0.9
MG(m= 3.5) s.d.:	1.9	1.3	1.0	0.8	1.9	1.0	0.9	0.8
CA abs.dif., mean:	1.3	1.8	0.9	0.4	2.3	0.7	0.8	0.4
MG abs.dif., mean:	1.2	1.6	1.1	0.3	2.0	0.6	0.7	0.3
Midwest Quarter								
Mean Pile Tons:	1683	10099	20197	40395	1683	10099	20197	40395
CA(m=89.1) s.d.:	4.1	3.7	3.4	1.8	4.1	1.9	1.8	1.8
MG(m= 6.7) s.d.:	3.4	3.0	2.7	1.2	3.4	1.3	1.2	1.2
CA abs.dif., mean:	2.0	3.5	4.2	0.5	3.7	0.8	0.6	0.5
MG abs.dif., mean:	1.6	3.0	3.6	0.4	3.1	0.7	0.6	0.4
Western Quarter								
Mean Pile Tons:	1621	9724	19447	38894	1621	9724	19447	38894
CA(m=86.5) s.d.:	6.5	6.1	5.8	1.4	6.5	1.5	1.4	1.4
MG(m= 8.3) s.d.:	5.3	5.0	4.9	1.1	5.3	1.2	1.1	1.1
CA abs.dif., mean:	2.6	7.0	11.3	0.8	5.9	1.0	0.9	0.8
MG abs.dif., mean:	2.1	5.7	9.5	0.6	4.7	0.8	0.8	0.6

TABLE 13

Statistical summary: selective mining.

Recovery Sequence:	CONSECUTIVE				NONCONSECUTIVE			
Pile Size (blocks):	1	6	12	24	1	6	12	24
Number of Piles:	1920	320	160	80	1920	320	160	80
TOTAL DOMAIN								
Mean Pile Tons:	1570	9422	18844	37687	1570	9422	18844	37687
min.:	76.7	83.6	84.9	88.1	76.7	87.8	88.1	88.1
$CaCO_3$ (CA) max.:	97.4	97.1	96.9	96.0	97.4	96.3	96.2	96.0
mean=92.3 s.d.:	3.4	2.9	2.7	2.2	3.4	2.3	2.2	2.2
min.:	0.1	0.2	0.6	1.4	0.1	1.1	1.2	1.4
$MgCO_3$ (MG) max.:	15.2	10.3	9.3	6.7	15.2	7.3	6.8	6.7
mean=4.1 s.d.:	2.5	2.1	1.9	1.4	2.5	1.5	1.5	1.4
Absolute differences between successive piles:								
min.:	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$CaCO_3$ (CA) mean:	1.6	2.1	2.4	0.5	2.6	0.7	0.6	0.5
max.:	11.2	8.2	6.9	1.9	15.8	4.3	2.2	1.9
min.:	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$MgCO_3$ (MG) mean:	1.3	1.8	2.0	0.4	2.1	0.6	0.5	0.4
max.:	9.0	6.3	5.6	1.5	11.8	3.3	1.8	1.5
SUBDIVISIONS OF THE TOTAL DOMAIN								
Eastern Quarter								
Mean Pile Tons:	1763	10578	21155	42311	1763	10578	21155	42311
CA(m=94.9) s.d.:	1.8	1.3	1.3	0.8	1.8	0.8	0.8	0.8
MG(m= 2.4) s.d.:	1.6	1.3	1.2	0.7	1.6	0.7	0.7	0.7
CA abs.dif., mean:	0.8	1.3	1.9	0.3	1.5	0.4	0.4	0.3
MG abs.dif., mean:	0.7	1.3	1.8	0.3	1.4	0.4	0.4	0.3
Mideast Quarter								
Mean Pile Tons:	1689	10134	20267	40534	1689	10134	20267	40534
CA(m=93.8) s.d.:	2.0	1.3	0.9	0.8	2.0	0.9	0.8	0.8
MG(m= 3.3) s.d.:	1.8	1.2	0.8	0.7	1.8	0.8	0.7	0.7
CA abs.dif., mean:	1.2	1.5	0.7	0.4	2.2	0.6	0.6	0.4
MG abs.dif., mean:	1.1	1.4	0.8	0.3	2.0	0.6	0.6	0.3
Midwest Quarter								
Mean Pile Tons:	1584	9502	19004	38008	1584	9502	19004	38008
CA(m=90.4) s.d.:	2.9	2.2	1.8	1.1	2.9	1.2	1.1	1.1
MG(m= 5.5) s.d.:	2.3	1.7	1.3	0.7	2.3	0.8	0.7	0.7
CA abs.dif., mean:	2.0	2.4	2.2	0.6	3.1	0.7	0.5	0.6
MG abs.dif., mean:	1.6	1.8	1.8	0.5	2.5	0.6	0.4	0.5
Western Quarter								
Mean Pile Tons:	1246	7474	14948	29897	1246	7474	14948	29897
CA(m=90.2) s.d.:	3.6	2.8	2.6	1.0	3.6	1.3	1.0	1.0
MG(m= 5.0) s.d.:	2.8	2.2	2.1	0.8	2.8	1.0	0.8	0.8
CA abs.dif., mean:	2.3	3.2	4.7	0.7	3.4	1.1	0.7	0.7
MG abs.dif., mean:	1.9	2.4	3.7	0.5	2.6	0.9	0.5	0.5

of the absolute values of the differences in $CaCO_3$ and $MgCO_3$ between successive piles. The differences are especially important because sudden big jumps in composition can wreak process- and quality-control havoc in a cement plant. The mean tonnages, compositional means and standard deviations, and mean absolute differences in composition are also reported separately for the four east-to-west sets of twenty slices. A discussion of the most important results follows. Emphasis is placed on the sensitive $MgCO_3$ grade.

Selective versus nonselective mining. These alternatives are nearly equivalent in the eastern half of the domain where there is little dolostone in the section. For nonselective mining, recovered tonnages decrease slightly westward because of a thinning of the total section; for selective mining, the westernmost tonnage is only 2/3 of that in the east. For nonselective mining, $MgCO_3$ grades increase from 2.4% in the east to 8.3% in the west. For selective mining the corresponding increase is 2.4% to 5.1%. Further selectivity would probably be uneconomical, so the western stone would probably be blended with better stone from another area to decrease the $MgCO_3$ content. Selectivity decreases the pile-to-pile variability of the western stone somewhat, particularly when consecutive block recovery is employed. However, the primary benefit of selectivity is in the reduction of the mean $MgCO_3$ grade, paid for by a reduction in tonnage and an increase in mining costs.

Consecutive versus nonconsecutive recovery. The general effects of these alternatives can be clearly visualized in the autocorrelation

functions (ACFs) of the time series.[120] Figure 106 shows ACFs of $MgCO_3$ for the western quarter, using nonselective mining and single- and twelve-block piles. Both alternatives yield a standard deviation of 5.3% $MgCO_3$ for single-block piles. However, for blocks consecutively recovered (Figure 106(a)), this variability is concentrated in a peak at lag 24, corresponding to the 24 blocks recovered consecutively from one slice through the simulation domain. If recovery shifts to every fourth block (b), the peak occurs at every sixth lag, corresponding to the six blocks recovered from each nonconsecutive pass through the slice. Because successive blocks in the nonconsecutive time series come from different parts of the deposit, the mean block-to-block absolute difference in $MgCO_3$ is 4.7%, versus only 2.1% for consecutive mining, so clearly consecutive mining is preferable in this case.

This comparison reverses dramatically when the mining output is blended into twelve-block piles. The ACF for consecutive mining (c) then shows a rapid alternation between high- and low-$MgCO_3$ piles, whereas the ACF for nonconsecutive piles (d) shows only mild alternations superimposed on a drift, which represents a gradual east-west change in stone compositions. The mean absolute difference correspondingly decreases from an astounding 11.3% for consecutive mining to 0.9% for nonconsecutive mining. And the standard deviations are no longer equal: 4.9% for consecutive, 1.1% for nonconsecutive mining. The advantage of blending stone from different parts of the quarry is realized only when the stone is really physically blended; otherwise control problems are made worse rather than better.

---

[120] The autocorrelation function $\rho(h)$ is related to the semivariogram function $\gamma(h)$ by the formula $\rho(h)=(\gamma(\infty)-\gamma(h))/\gamma(\infty)$.

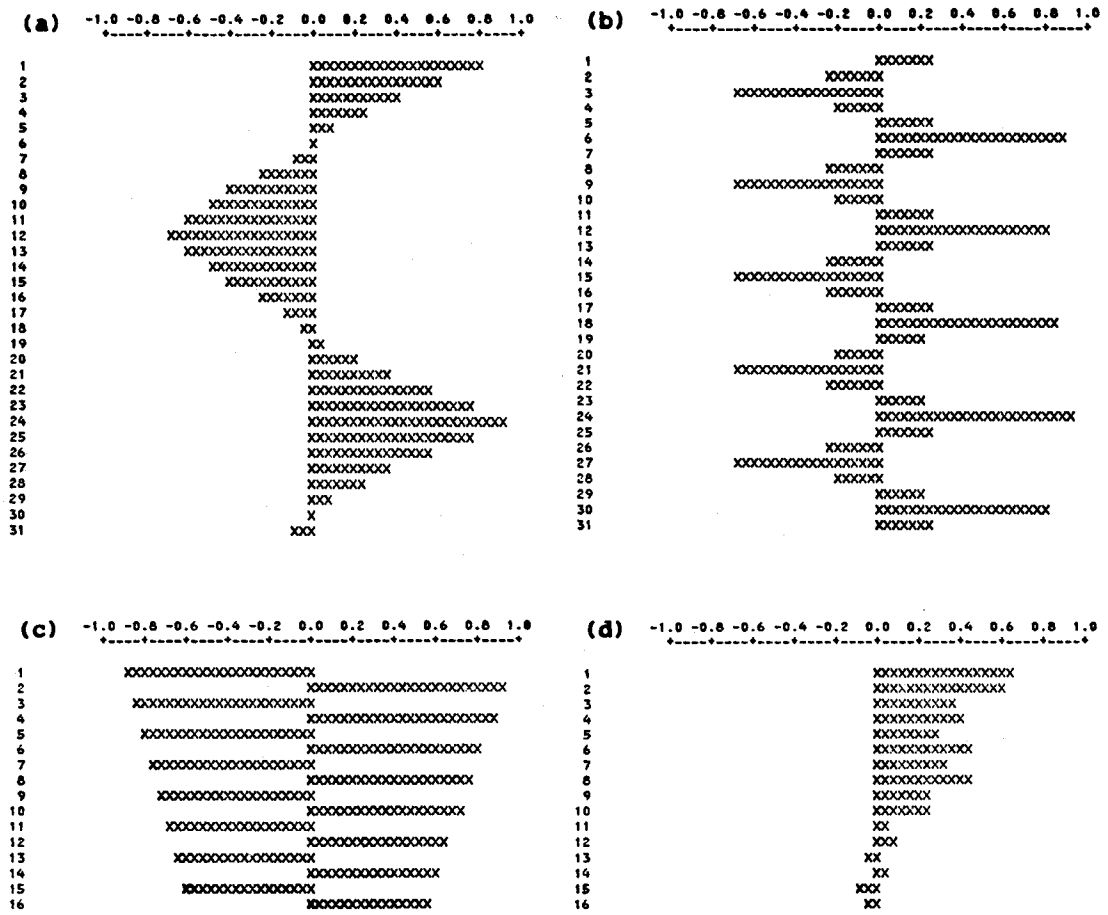


Figure 106: Autocorrelation functions of $MgCO_3$ from the western quarter, using two recovery and two homogenization alternatives. (a) single-block piles, consecutive; (b) single-block piles, nonconsecutive; (c) twelve-block piles, consecutive; (d) twelve-block piles, nonconsecutive.

Size of homogenization piles. With nonconsecutive mining, the influence of pile size is straightforward: as soon as the pile size increases from one to six blocks, both the standard deviations and mean absolute differences decrease dramatically. Further increases in the pile size cause only very modest further decreases in these values.

With consecutive mining, the standard deviations decrease modestly with increasing pile size, showing a dramatic decrease only for piles of 24 blocks. The absolute differences generally increase modestly (in the east) to severely (in the west) as the pile sizes increase to twelve blocks, then decrease dramatically (especially in the west) at twenty-four blocks. The large buildup in the west reflects the fact that the southern part of this area is very dolomitic and the northern part is not.

In general, the effect of pile size in any deposit being mined consecutively should depend essentially on the comparison between the distance within the mine over which material in the pile has been collected and the range of the variogram of ore grades. If the distance greatly exceeds the range, such that a variety of ore grades is being averaged, then successive piles should be very much alike; if the distance is less than the range, the pile-to-pile variability will increase with the size of the pile, with undesirable results. These effects can be quantified in terms of regularizations of the original point-support ore grades over volumes of different sizes, representing the pile sizes. The clear advantage of nonconsecutive mining in conjunction with pile homogenization is that the individual blocks contributing to each pile can effectively cover a large volume of the deposit while contributing only a small tonnage to the pile. In this way, large compact volumes of raw material in situ are essentially discretized into sets of a few scattered small volumes (the blocks), the averages of which are little more variable than the total compositions of the large volumes.

The simulation domain used in this case study was deliberately chosen to encompass a variety of conditions, which are reflected in the disparate results of the four subareas represented in Tables 12 and 13. Thus the results obtained here can be generalized with some confidence to the rest of the property. Furthermore, the gross statistics of the five large regions of the property summarized in Table 8 suggest that the simulation domain -- essentially Region 2 -- is statistically similar to other regions. The results clearly suggest that contributions from six or so widely scattered mining blocks, when averaged into a pile of about 10,000 tons (i.e., nonconsecutive recovery, piles of six blocks), would normally achieve very effective reduction of the variability of the incoming limestone. Modest selective mining (principally local stripping of the dolomitic caprock) in some areas, combined with deliberate blending of the output of these areas with stone from low-$MgCO_3$ areas, would allow maximization of the remaining reserves while keeping MgO levels and fluctuations in the clinker under control. The wealth of available core-drill data, supplemented by blast-hole analyses in problem areas, should allow accurate predictions (by geostatistics or less formal means) of stone quality in all areas of the property. These estimates could then be used to proportion stone from different active faces into each blending pile, further reducing the pile-to-pile variability of the stone beneath the levels suggested by these simulation results. Average pile-to-pile differences in the neighborhood of 0.1-0.3% $MgCO_3$ should be obtainable with these methods. The existing conditional simulation -- coupled with

detailed simulations of blast-hole sampling,[121] grade estimation, mining, proportioning, and homogenization -- could be used to refine these values.

Pile homogenization without blending from the quarry, and quarry blending without some type of in-plant homogenization, would be much less efficient and possibly counterproductive. If a stockpile prehomogenizer is not installed, the basic requirement is that the mean residence time of the existing powder homogenization system should greatly exceed the wavelengths at which most of the mine output variability is concentrated. For example, if stone is delivered successively from six different faces in successive thirty-ton trucks, the homogenization system should have a mixing capacity of well over 180 tons and preferably over 900 tons to ensure that the homogenized output does not vary unduly because of large temporal differences in the composition of the input. The capacity would not need to be so large if the stone delivered from the six faces were dumped onto six piles and then fed to the crusher in successive bucketloads by a front-end loader, thus shortening the wavelengths of the input series. The effectiveness of a powder homogenization system can be determined analytically or by simulation, provided that the statistical characteristics of the input time series are known and the capacity (for a batch system) or

---

[121] Because the conditional simulation was performed on "core" support, simulated blast-hole data can be drawn directly from the simulation file. Some noise might be added to the data to reflect the fact that blast-hole data are typically of lesser quality than core data. The properties of the noise could be inferred by comparing the statistics of analyses of core-drill samples and blast-hole samples taken from the same area.

residence-time distribution (for a cascade system) of the homogenizer are also available.

4.2.6 Simpler Alternatives

> Remember that complexity and accuracy are not synonymous: there is no merit in using enormously detailed sub-models if the global result desired is insensitive to such details. [Bratley et al., 1983, p. 33]

Usually there is no single "correct" way to perform a conditional simulation. In the Mason City case study, for example, choices among alternative approaches were available at every step. Because this case study was performed primarily to illustrate a variety of techniques, simplicity was not a major consideration when these choices were made. But in a "production" simulation directed toward solving a practical problem quickly and cheaply, simplicity is important. In this short concluding section, some alternatives are suggested that would have simplified the study and probably would have yielded practical results at least as acceptable as those actually obtained. These alternatives might have introduced problems of their own, but without actually trying them it is hard to say what, or how bad, the problems would be. Unless theoretical flaws can be pointed out, it can be difficult to judge the relative merits of different procedures until they have been tried out and the results compared.

One way to simplify a study such as this is to make a few small compromises with reality. Some aspects of the phenomenon under study may not greatly influence the results and so may not need to be modeled exactly, if at all. For instance, the major source of difficulty in the Mason City study was that five variables were involved, and that the

relationships among these variables were not straightforward (e.g., see Figure 91). Yet for practical purposes only one variable -- the total $MgCO_3$ content of the mined stone (limestone and dolostone combined) -- is vitally important. Thus the easiest simplification to make would be to ignore all other variables and focus on this one.

There are two ways in which the total $MgCO_3$ content could be quantified: either directly through the $MgCO_3$ analyses of the stone, starting with the analyses of the original irregular-support core data; or indirectly through the distribution of dolostone in the deposit, starting with the geologist's logs, effectively at point support, of the occurrences of limestone and dolostone in the core.

There are good reasons for thinking that the indirect approach might actually represent less of a compromise. Rock type is measured on arbitrarily small support, so no support corrections, with their attendant smoothing, need to be made. Also, the chemical compositions of the rocks have very high nugget components (especially the limestones), so that any averaging of simulated values during mining simulations rapidly reduces their variability. In such a case, little information is lost by simply replacing the limestone and dolostone analyses by their average values over the whole deposit, attributing all variability in composition to variability in rock type. Finally, rock type accounts for 95.1% of the total $MgCO_3$ variability in the original constant-support data. (The correlation coefficient between variables MG and PCTDOL is 0.975.) This is true because the average compositions of the limestone and dolostone are so different, whereas their internal variability (particularly in the limestone) is rather slight.

During mining, variability within rock types is rapidly smoothed out, such that the relationship between MG and PCTDOL should be strengthened -- particularly in the case of nonselective mining where more dolostone is included in the mined stone. This is supported by the mining simulation results. Table 14 shows the correlation coefficients between PCTDOL and MG for the real and simulated data, and for the sixteen mining and homogenization alternatives discussed in Section 4.2.5. For nonconsecutive selective mining with six-block piles -- a likely strategy for practical homogenization at Mason City -- PCTDOL accounts for 97.4% of the variability in MG. Furthermore, the correlation coefficients between MG and CA are consistently greater than 0.99 for nonselective mining and around 0.97 for selective mining, making it clear that a simulation of MG (perhaps via PCTDOL) would provide satisfactory information on CA as well.

With only one variable to simulate, more effort could be devoted to accurate reproduction of sample semivariograms and other characteristics of the data. In the case study reported here, the semivariograms of the simulated PCTDOL data, and of the simulated CA and MG data that depend so heavily on PCTDOL, poorly reproduced the semivariograms of the conditioning data because their very short-scale structures were biased by the despiking formula. In a simpler simulation there would be more time to try out several alternative despikings (e.g., one of the simulation approaches in Section 3.8.3.4) if the short-scale structure were critical to the study. Simulation of only PCTDOL would rid us of the problems of support corrections (Section 4.2.2.4), assignment of synthetic analyses to conditioning data (Section 4.2.2.2), reproduction

TABLE 14

Correlation coefficients between percentage of dolostone (PCTDOL) and total $MgCO_3$ content (MG) in Region 2.

Original data:	0.975
Simulation Subset 1:	0.971
Simulation Subset 2:	0.967

Mining and homogenization simulations, Region 2:

Recovery sequence:	CONSECUTIVE				NONCONSECUTIVE			
Pile size (blocks):	1	6	12	24	1	6	12	24
Nonselective mining:	0.992	0.994	0.996	0.997	0.992	0.997	0.997	0.997
Selective mining:	0.976	0.976	0.982	0.988	0.976	0.987	0.987	0.988

of chemical constraints by means of transformations (Section 4.2.2.6), and a host of modeling chores, plus the significant processing costs for simulating and kriging five variables. So in a "production" study of the Mason City situation, this would certainly be the approach to take.

Another simplification that might be undertaken in this study would be to skip the adjustments for geologic structure. The structural trends estimated by kriging in Section 4.2.2.3 are broad enough to be imposed by conditioning. Short-scale structural features that might influence short-scale semivariogram modeling were not removed by this adjustment anyway. Some mining-simulation methods might still require the imposition of a reasonably flat quarry floor, but the methods

actually used in Section 4.2.5.1 did not, because recovery was performed between two estimated surfaces which could just as well have incorporated the large-scale structural trends.

Chapter V

CONCLUDING REMARKS

5.1 GEOSTATISTICS APPLIED TO CEMENT PROBLEMS

Portland cement is remarkable among mineral commodities because it is nearly always manufactured from a mixture of raw materials with diverse compositions. A single raw material can rarely be declared intrinsically suitable for cement manufacture. Only suites of raw materials can be declared jointly suitable. The sole exception is a "cement rock" (a shaly limestone) that happens to possess an ignited-basis analysis comparable to that of a cement clinker. The concepts of "purity" or "grade" used to describe most mineral resources have limited value in discussions of individual cement raw materials, except in a negative sense. For example, a low-magnesium limestone or low-alkali clay might be regarded as "high-grade". A limestone averaging 99% $CaCO_3$, though certainly a high-grade source of calcium, actually would rarely be more valuable for cement making than one averaging 95%. Usage of the 99% limestone would simply require that more clay be added to the mix. The concepts of "cutoff grade" and "recoverable reserves" that figure so prominently in many geostatistical studies also have limited value in a cement context, except again in a negative sense; and even in that sense the cutoff grade and recoverable reserves can be determined only when the characteristics of all materials being used in the mix are taken into account.

In most mineral industries, knowledge of the variability of the raw materials is important for two reasons: first, local excursions of raw material quality above or below an established cutoff grade can affect recoverable reserves and influence the choice of mining procedures; second, highly variable mined materials may present control problems in the mill and require expensive redesigning of the process. For the reasons cited above, the first consideration is rarely important for cement raw materials. Selective mining is rarely practiced except on a gross scale, e.g., by wasting of thick dolostone beds or avoidance of large areas of a property where quality problems or other difficulties are known to exist. (The low unit value of the product and the low profit margins experienced by most cement plants also make expensive selective mining undesirable.) Even in those cement operations where a cutoff grade has been established, occasional violations can usually be quickly corrected by blending. Hence the most common applications of mining geostatistics -- local estimation of ore grades and global estimation of recoverable reserves -- are of little use in a typical cement operation.

The one important contribution that geostatistics can make to the study of cement raw materials is in the assessment of raw-material variability for design and control applications. Adaptations and extensions of standard geostatistical techniques, particularly the techniques of conditional simulation, to deal with these applications have been the subject of this dissertation. Plant designs, process-control methods, and quality-control procedures should be specifically tailored to the time-series characteristics of the raw materials to be

delivered to the plant. These time series are the products of complex interactions between the spatial realizations of the individual raw materials in situ and the mining procedures used to extract the materials. Before mining begins, the best way we can hope to forecast the characteristics of these time series is by conditional simulation of the in-situ realizations and simulation of the mining procedures operating on these realizations. The only alternatives are to rely on experience from other "similar" operations, or simply to guess at the characteristics of the time series and incorporate liberal safety factors into the design to allow for bad guesses.

One major obstacle to the use of geostatistics in cement applications is the nearly universal paucity of data. Because the recoverable-reserves and local-estimation problems faced by metals operations are unimportant in most cement projects, and also for several historical reasons discussed in Section 2.4.3, the cement industry has established a tradition of collecting only enough data on raw-material compositions to assess the average compositions and total tonnages of available raw materials, but rarely enough data to characterize the compositional variability of the materials. There is a particularly critical shortage of closely spaced data, which are needed for the assessment of variability at short scales -- either spatially, in the ground, or temporally, in the plant. Even the quality of the data that are available is commonly poor, because it is felt that accuracy and precision in sampling and analysis is important only for finished products, not for raw materials.

An obvious, but far from easy, solution to this problem is to collect more and better data. Unfortunately data are expensive, cement companies are commonly short of cash, and old traditions die hard. One contribution that this dissertation might make is to demonstrate that the expense of gathering data adequate for a geostatistical study can actually be worthwhile in some situations.

5.2 COSTS AND BENEFITS OF SIMULATION STUDIES

The Mason City quarry property described in Section 4.2 has been explored by over 40,000 feet of core drilling, allocated among over 700 holes and over 5000 analyzed core samples. Closely spaced holes are in short supply, but in all other respects this data set is perfectly adequate for a geostatistical study. Adequacy is achieved when the following conditions are fulfilled: (1) the spatial coverage of drilling data and the number of drilling data are sufficient for accurate modeling of the semivariogram structure at all distances and directions h of interest to the study (Section 3.1.3); (2) the data are sufficient to infer the multivariate distribution at least at h=0 (Section 3.5.2); and (3) the density of data in all parts of the property is such that several data at distances within the range of the semivariogram are available to calculate a kriged value at any point of interest. The Mason City data set is vastly larger than most data sets available for cement properties, because the critical $MgCO_3$ content of the raw materials at that site is unusually variable, and because the geological staff apparently did a good job of convincing management of the value of drilling information in that kind of situation. For

estimation or conditioning purposes, a smaller property or a property with better-behaved materials could be adequately covered with many fewer data, provided that the data were still adequate for inference of the local semivariogram and distribution at h=0. For instance, apparently quite good simulation results (for only two dimensions and a relatively coarse discretization) were obtained in Section 4.1 from a set of only 129 well placed five-variate data.

In principle, the minimum number and density of data needed for data analysis, structural analysis, and kriging (and therefore for simulation) cannot be specified a priori, because the estimation variances and fluctuation variances of the semivariogram and distribution at h=0 depend on the semivariogram and distribution to be estimated. However, the practical advice of Journel and Huijbregts (1978, p. 194) that at least thirty to fifty data pairs should be available at each lag of the sample semivariogram can be used as a guide to the number of data needed for semivariogram estimation. Unless the data set contains many variables and a complicated transformation to normality is needed, the number of data required to estimate a semivariogram should ordinarily be adequate to estimate the distribution at h=0 and to derive a reasonable transformation.

Whatever the case, the data are apt to cost many times more than the geostatistical study itself. For instance, the combined Mason City drilling programs would have cost over half a million dollars at today's prices, and the total costs of sampling and analysis might have been comparable. The case study of Section 4.2 (which did not extensively use most of the data), including the actual writing of Section 4.2,

required only about six months of work, of which about three months were spent in typing, proofreading, and otherwise verifying over 8,000 lines of raw data. About two months were consumed in the complicated data analysis and structural analysis described in Section 4.2.2. The conditional simulation required about one month, and the mining and processing simulations (the interesting part) two days. In Section 4.2.6, it is pointed out that the whole simulation project could easily have been reduced to a much simpler simulation of one variable with little loss of accuracy, probably reducing the time required for the structural analysis and simulations to a total of one month. If the data were already typed in and verified, the total cost of the study would have been one month's salary and expenses for one professional, plus computing and word-processing expenses (a few thousand dollars at the most, and perhaps free if done in-house). These costs compare to the cost of the data set as a cent compares to a dollar, and perhaps only 10% of that single cent is actually spent on simulation. About 90% of the human and computer time involved in a conditional-simulation project is spent on data analysis, structural analysis, and kriging, all of which would be performed in virtually any geostatistical study. The additional effort involved in simulation consists mainly of selecting a few parameters for the simulation routine, generating the unconditional simulations, and then performing a few simple manipulations of the simulated and kriged data. If very detailed simulations of mining, processing, and control alternatives are performed instead of the simple simulations described in Section 4.2.5, the cost of the study might increase from one cent per dollar of data to a few cents.

Now what are the benefits of a simulation study? In mining and mineral-processing applications, the benefits arrive in the form of better forecasts of raw-material variability for use in the design of mining, processing, and control systems. These systems together cost several million dollars to install, so it pays to design them correctly. It is particularly important to avoid "overdesigning" or "underdesigning" the capabilities of a system simply because the characteristics of the raw materials, including their variability, are poorly known. An example of an overdesigned system is a stacker-reclaimer homogenization system with a 40,000-ton capacity that actually uses materials that could have been adequately homogenized with a less expensive 10,000-ton system. A 10,000-ton system would be underdesigned if it had to be enlarged to 40,000 tons after the plant disastrously failed to produce within specifications because the raw feed was inadequately homogenized.

One can imagine situations in which a conditional simulation, or even an entire simulation study as described in this dissertation, would be of little practical use. For instance, a conditional simulation would be pointless if the drilling data were so closely spaced, and the mining blocks so large, that several holes fell inside each mining block. In such a case (exceedingly rare, even for metallic deposits), an estimated block composition obtained by averaging the drilling data within a block (or by kriging) would scarcely differ from a conditionally simulated grade, so mining simulations could just as well be performed on estimated block grades. Conditional simulation (or any other geostatistical study) would also be of little use if the raw materials

to be used by a new plant were so nearly constant in composition that the variability caused by contamination or segregation during handling clearly would overwhelm the effects of the in-situ variability of the materials. This is in fact the typical situation in sedimentary deposits, including probably a majority of deposits currently being mined as cement raw materials. Only in the case of a deposit with marked intraformational changes in composition is a conditional simulation potentially worthwhile; but as more and more of the best deposits in favorable market areas are mined out or made unavailable by land development, more and more of these variable deposits will be pressed into service. The greater sensitivity of new manufacturing processes to raw-material variability makes it particularly likely that the in-situ variability of cement raw materials will become an increasingly important issue to be considered in plant design and cement manufacturing operations in the future. Situations in which conditional-simulation studies will be of value will therefore become more common.

5.3 RESEARCH NEEDS

> If I know what I know and I know what I don't know, then I do know it all. [Larry Klein, in Stereo Review, July 1979, p. 16.]

5.3.1 Methods

The field of indicator simulation is in its infancy, and much work remains to be done. The methods described in Section 3.8.3 all depend on simulations of gaussian processes, which are transformed to indicator

simulations by application of a cutoff. But not all indicator phenomena can be reproduced in this way (e.g., see Figure 57). Thus new methods not based on gaussian simulations are needed. The matrix method of Davis (1985b) is a good start, but methods capable of simulating hundreds of thousands of indicator data quickly and cheaply and capable of reproducing specified nongaussian patterns are apparently unavailable. Nongaussian simulations of continuous phenomena such as those in Figure 53 also are needed. Conditional simulations of stream networks or fracture patterns, possibly integrated with simulations of the surrounding topography or rock characteristics, could have valuable applications in hydrology and rock mechanics.

Even in the case of gaussian-based indicator simulations, further research is need on the simulation of gaussian conditioning data to correspond to the available indicator conditioning data. Several untested alternatives are proposed in Section 3.8.3.4. Practical comparisons of these methods, and perhaps derivations of better ones, remain to be performed.

Univariate gaussian transformations of nongaussian data (Section 3.5.2.2) may be good enough to transform most nongaussian coregionalized data into acceptably gaussian-looking data sets for simulation purposes, but so few practical simulations of coregionalized data have been done that this preliminary conclusion may turn out to be false. In any case a fast and easy multivariate (at h=0) transformation (Section 3.5.2.3) would certainly be preferable to independent transformations of the marginal distributions. Development of software to implement such transformations and experimentation on real coregionalized data sets are needed.

5.3.2 Applications

Simulations of proportioning and control operations in cement plants using conditional simulations of several raw materials were not performed for this dissertation, because no suitable multivariate data sets for multiple raw materials were found. Some ways in which several conditional simulations could be used to design proportioning facilities are suggested in Section 3.9.2.2, but these suggestions need to be tried out in practice. These simulations could also be used to check the sensitivity of deterministic mix-design solutions (Sections 2.3.2 and 2.3.3) to variability in the compositions of the raw materials.

Constrained cluster analysis (Section 3.10.1) and methods of multivariate graphical data analysis (e.g., McDonald, 1982) appear to have numerous applications in the earth sciences. These methods may be useful for isolating members of different populations included in a data set to be used for conditional simulation, and could be of further use in "dissecting" continuous single populations with long ranges or drifts into quasi-stationary regions. In a cement operation, for instance, it may be desirable to dissect a deposit having vaguely bounded patches with different compositions into two or more subpopulations to be mined out and possibly homogenized separately.

Conditional simulation has now been widely used to produce simulated realizations of raw materials to be mined and of reservoir parameters for applications in groundwater hydrology. Petroleum applications are nearly identical to those in hydrology but have received less attention. Applications in agriculture, forestry, soil science, rock and soil mechanics, oceanography, meteorology, and other areas where spatially

distributed data are collected can easily be imagined. Other geostatistical techniques (primarily various types of kriging) have been employed in several of these areas (some described in Verly et al., 1984), but conditional simulation remains underutilized. The shortage of data that commonly afflicts geostatistical studies of cement raw materials is a problem in many of these fields as well, so the usage of subjective information for conditioning and model inference is a final area where additional research, as well as testing and refinement of existing methods, is warranted.

Appendix A

SUBROUTINE CS2D: "CIRCULAR SIMULATIONS, 2 DIMENSIONS"

Subroutine CS2D simulates up to ten independent realizations from a stationary gaussian random function with mean=0.0, semivariogram sill=1.0, using a moving-average method (Section 3.3.1.2). Each simulated realization consists of data located on a rectangular grid within a two-dimensional rectangular domain. Depending on the value of input parameter "IW", the model random function can have either a (possibly anisotropic) circular variogram (IW≠1), or some other variogram predetermined by the user (IW=1) through the choice of a moving-average weight function (stored in discretized form in matrix "W"). This subroutine is very simple, and the listing provided in this appendix is thoroughly commented, so no extended explanations or flow charts should be needed to understand it.

A circular simulation can be produced by smoothing a field of independent random numbers, averaging the numbers within a moving circular window whose diameter is the required variogram range. This subroutine accomplishes the same thing when option IW≠1 is selected, but in a slightly different fashion. The actual method is explained below, in the discussion of the first demonstration run.

If IW≠1, the moving-average weight matrix "W" is calculated by the subroutine, which assigns weights of "1" to points on or inside a circle of diameter "DIAM", "0" to points outside the circle. In the case of

isotropic (square) grid spacings (parameters DL=DC), the resulting moving-average simulations have an isotropic circular semivariogram model, derived in Section 3.3.1.2, with range=DIAM. A geometric anisotropy parallel to the grid directions can be imposed by simply manipulating the grid spacings, as illustrated in Figure 47 of Section 3.7.4.1. For example, to produce a circular simulation with range=20 in the east-west direction (along grid rows, or lines) and range=10 in the north-south direction (along grid columns) and a final grid spacing of 1x1, one can specify DIAM=10.0, DL=1.0, DC=0.5; then twenty grid spacings (actually twenty-one points) will be included within the circular moving window in the east-west direction, ten in the north-south direction.

General two-dimensional moving-average simulations (IW=1) can be performed by supplying a predetermined weight matrix "W", which may be derived to correspond to some other semivariogram model (using the relation on page 114). In practice, this option should be necessary only rarely, as the sample semivariograms of most geologic data sets can be nicely modeled using one or more nested circular models and a nugget constant. However, the option might also be used to specify uniform elliptical weight functions for circular simulations having anisotropies oblique to the grid directions.

Subroutine CS2D automatically adjusts the simulated values in the data matrix "Y" to an <u>expected</u> mean=0.0, sill=1.0. Exact standardization of the simulated data to mean=0.0, variance=1.0 is <u>not</u> performed by this subroutine, although the user is certainly free to standardize the data matrix "Y" after it is returned to the main program

by the subroutine, if this is appropriate. Standardization is avoided here because all characteristics of natural realizations of a random function should exhibit fluctuation variances if the realizations are observed within a finite spatial domain (Section 3.7.2). Standardization of the simulated realizations would artificially drive the fluctuation variances of the mean and dispersion variance within the simulation domain to zero, and also distort the fluctuation variances of other characteristics, bias the semivariogram sill, and possibly bias the short-scale semivariogram structure of the individual simulations (Section 3.7.2).

This subroutine employs a widely used multiplicative congruential pseudorandom-number generator, of the following form (constants in parentheses are REAL*8 numbers):

1. Initialize DSEED, an integer on the interval (1,2**31-1), stored as a REAL*8 constant.
2. Then, for each new pseudorandom number X to be generated,

   DSEED=DMOD((7**5)*DSEED,(2**31-1))

   X=DSEED/(2**31)

   where DMOD(A,B) is a Fortran built-in function defined by

   DMOD(A,B) = A - K * B

   where K = integer part of (A/B).

This generator is described in the textbooks by Kennedy and Gentle (1980, p. 147), Rubinstein (1981, p. 25), Law and Kelton (1982, p. 226), and Bratley et al. (1983, p. 198). Notice that the value of DSEED is reset after the generation of each new random number. If subsequent simulated values are to be assured of independence from previous values,

the parameter DSEED must not be otherwise reset during execution of the program. However, if a string of random numbers is to be repeated exactly for some reason, all that is necessary is to reset DSEED to its original value. Thus, the sequence of random numbers is actually deterministic, hence the name "pseudorandom". To ensure that simulations from different runs are independent, one can start a subsequent run using the final value of DSEED printed out by subroutine CS2D at the end of the previous run. Alternatively, one can use a table of seeds for this generator provided by Bratley et al. (1983, p. 213). Their table lists thirty seeds spaced 131,072 apart. Bratley et al. (1983, pp. 200-202) point out that the product (7**5)*DSEED, calculated as a REAL*8 number, can require as much as 46-bit accuracy, and thus will work fine on an IBM-370 type of computer but not on some others. They provide FORTRAN code (on their p. 202) for a generator using only 32 bits, including the sign bit, which might be substituted for the generator above if this subroutine is to be executed by a small computer.

The subroutine will write the simulated data directly to the output device designated by logical unit number IOUT if parameter IWRT=1. Matrix "Y" and the associated output vector are illustrated in Figure 112. This output is convenient if unit IOUT is a sequential disk file that can be edited later for input to other programs. Otherwise, the data returned by the subroutine should be stored in some form by the main program.

Two demonstration runs of subroutine CS2D are provided in this appendix. The first run is executed using the main program listed below:

```
$WATFIV TIME=300
$ASSIGN 8 TO FILE CS2DOUT OUTPUT
C*****
C*****  INTERACTIVE (ORVYL) PROGRAM -- FOR BATCH RUNS, DELETE
C*****  ALL MESSAGES TO TERMINAL MARKED BY "C*****" BELOW.
C*****
C
C     DEMONSTRATION RUN OF SUBROUTINE CS2D -- MAIN PROGRAM
C
C     DIMENSIONS FOR SUBROUTINE CS2D
C
      DIMENSION FMT(8),YMEAN(2),YVAR(2),Y(50,50,2),W(9,9)
      DIMENSION LWLF(52),LWLL(52),LWCF(56),LWCL(56),X(2)
C
C     DIMENSIONS FOR SUBROUTINES HIST, GAM2, AND SIMAP
C
      DIMENSION VR(5000),WT(5000),XFR(20),T(2)
      DIMENSION ID(4),JD(4),ND(2),UD(2),VD(2),NP(160),GAM(160),UG(160)
      DIMENSION CHAR(50)
C
      DOUBLEPRECISION NAM(10),DSEED
C
C     OUTPUT OF CS2D IS IN THE PROPER ORDER FOR INPUT TO HIST AND GAM2
C
      EQUIVALENCE (VR(1),Y(1,1,1))
C
      COMMON INP,IOUT,DUMMY,NV,NAM
C
C     DATA FOR OUTPUT FORMAT, VARIABLE NAMES, R. N. GENERATOR SEED
C
      DATA FMT/'(1H ',',5F1','0.4)',',    ',',    ',',    ',',    ',',    '/
      DATA NAM(1),NAM(2)/'TEST ONE','TEST TWO'/,DSEED/1702.D0/
C
C     DATA FOR VARIO. DIRECTIONS, VARIO. TESTS, HIST. WTS.
C
      DATA ID/1,0,1,-1/,JD/0,1,1,1/,T/2*-100.0/,WT/5000*1.0/
C
C     NUMBERS IN COMMON
C
      INP=5
      IOUT=8
      NV=2
C
C     RUN 2 SIMULATIONS
C
      CALL CS2D(50,50,2,1.0,0.5,3.1,6,1,FMT,DSEED,9,9,0,YMEAN,YVAR,Y,W,
     *LWLF,LWLL,LWCF,LWCL,X)
C
C     RUN HISTOGRAMS, VARIOGRAMS, AND MAPS
C
      DO 200 LV=1,NV
  200 CALL HIST(VR,WT,2500,LV,0,1,0.0,0.0,0.0,-100.0,XFR,U,V,N,0)
C*****
```

```
C     COMMENT TO TERMINAL
      WRITE(6,8001)
 8001 FORMAT(' STARTING GAM2')
C*****
      CALL GAM2(20,T,50,50,4,ID,JD,VR,ND,UD,VD,NP,GAM,UG,1)
      CALL SIMAP(50,50,2,Y,CHAR)
C*****
C     COMMENT TO TERMINAL
      WRITE(6,8006)
 8006 FORMAT(' ALL DONE')
C*****
      STOP
      END
```

The subroutines HIST, GAM2, and SIMAP, which are not listed here, merely calculate sample histograms, semivariograms, and line-printer maps of the simulated data. The first page of output from CS2D (i.e., all output except the data listing) appears in Figure 107, sample semivariograms of the two realizations in Figure 108, and maps of the two realizations in Figure 109. The weight matrix "W" is calculated by the subroutine (option IW≠1) using the parameters DL=1.0, DC=0.5, DIAM=3.1, resulting in an "elliptical" moving window of the form

```
0 1 1 1 1 1 0
1 1 1 1 1 1 1
0 1 1 1 1 1 0
```

as illustrated in Figure 110. The subroutine simulates NV=2 independent 50x50 arrays of data by generating two 52x56 sets of uniform pseudorandom numbers, NV=2 numbers at a time (Figure 111). Each random number is added to all elements of the corresponding 50x50 simulation array that lie within the elliptical window (described above) when it is centered on the location of the random number being generated, as shown in the upper-left corner of Figure 111. The random numbers are generated in a columnwise sequence, and the center of the window moves accordingly.

The expected value of random numbers independently and uniformly distributed on the interval (0,1) is 1/2, and their variance should be 1/12. Because each grid location in the simulation array is overlapped by the circular window seventeen times (corresponding to the seventeen positive elements of matrix "W" in this example), each element of the data matrix "Y" accumulates the contributions of seventeen independent uniform random numbers and thus is approximately normally distributed (central limit theorem) with an expected value of 17/2. The theoretical variance of the simulated values, if they were distributed over an infinite-sized domain, would be 17/12. This corresponds to the expected sill of the semivariogram. The object of subroutine CS2D is to simulate realizations from a random function with mean=0, sill=1, allowing for fluctuation variances. The "unadjusted" data, whose sample means and dispersion variances are summarized in Figure 107, are therefore "adjusted" (not standardized) to have expected means of 0.0 and sills of 1.0 by transforming each element "y" of matrix "Y" to:

$$y' = (y - 17/2) / \sqrt{(17/12)}$$

The adjusted sample means and variances are also reported in Figure 107.

The second demonstration run is executed using the main program listed below:

```
$WATFIV TIME=300
$ASSIGN 8 TO FILE CS2DWIN OUTPUT
C*****
C*****  INTERACTIVE (ORVYL) PROGRAM -- FOR BATCH RUNS, DELETE
C*****  ALL MESSAGES TO TERMINAL MARKED BY "C*****" BELOW.
C*****
C
C     DEMONSTRATION RUN OF SUBROUTINE CS2D -- MAIN PROGRAM
C
C     DIMENSIONS FOR SUBROUTINE CS2D
C
      DIMENSION FMT(8),YMEAN(2),YVAR(2),Y(50,50,2),W(9,9)
      DIMENSION LWLF(58),LWLL(58),LWCF(58),LWCL(58),X(2)
```

```
C
C     DIMENSIONS FOR SUBROUTINES HIST, GAM2, AND SIMAP
C
      DIMENSION VR(5000),WT(5000),XFR(20),T(2)
      DIMENSION ID(4),JD(4),ND(2),UD(2),VD(2),NP(160),GAM(160),UG(160)
      DIMENSION CHAR(50)
C
      DOUBLEPRECISION NAM(10),DSEED
C
C     OUTPUT OF CS2D IS IN THE PROPER ORDER FOR INPUT TO HIST AND GAM2
C
      EQUIVALENCE (VR(1),Y(1,1,1))
C
      COMMON INP,IOUT,DUMMY,NV,NAM
C
C     DATA FOR OUTPUT FORMAT, VARIABLE NAMES, R. N. GENERATOR SEED
C
      DATA FMT/'(1H ',',5F1','0.4)',','    ',','    ',','    ',','    ',','    '/
      DATA NAM(1),NAM(2)/'TEST ONE','TEST TWO'/,DSEED/1702.D0/
C
C     DATA FOR VARIO. DIRECTIONS, VARIO. TESTS, HIST. WTS.
C
      DATA ID/1,0,1,-1/,JD/0,1,1,1/,T/2*-100.0/,WT/5000*1.0/
C
C     DATA FOR SPECIAL WINDOW MATRIX "W"
C
      DATA W/1.,1.,0.,0.,0.,0.,0.,1.,1.,
     *       1.,1.,1.,0.,0.,0.,1.,1.,1.,
     *       0.,1.,2.,2.,0.,2.,2.,1.,0.,
     *       0.,0.,2.,2.,2.,2.,2.,0.,0.,
     *       0.,0.,0.,2.,3.,2.,0.,0.,0.,
     *       0.,0.,2.,2.,2.,2.,2.,0.,0.,
     *       0.,1.,2.,2.,0.,2.,2.,1.,0.,
     *       1.,1.,1.,0.,0.,0.,1.,1.,1.,
     *       1.,1.,0.,0.,0.,0.,0.,1.,1./
C
C     NUMBERS IN COMMON
C
      INP=5
      IOUT=8
      NV=2
C
C     RUN 2 SIMULATIONS
C
      CALL CS2D(50,50,2,1.0,0.5,3.1,6,1,FMT,DSEED,9,9,1,YMEAN,YVAR,Y,W,
     *LWLF,LWLL,LWCF,LWCL,X)
C
C     RUN HISTOGRAMS, VARIOGRAMS, AND MAPS
C
      DO 200 LV=1,NV
  200 CALL HIST(VR,WT,2500,LV,0,1,0.0,0.0,0.0,-100.0,XFR,U,V,N,0)
C*****
C     COMMENT TO TERMINAL
      WRITE(6,8001)
```

```
 8001 FORMAT(' STARTING GAM2')
C*****
      CALL GAM2(20,T,50,50,4,ID,JD,VR,ND,UD,VD,NP,GAM,UG,1)
      CALL SIMAP(50,50,2,Y,CHAR)
C*****
C     COMMENT TO TERMINAL
      WRITE(6,8006)
 8006 FORMAT(' ALL DONE')
C*****
      STOP
      END
```

The first page of the output appears in Figure 113, sample semivariograms in Figure 114, and maps in Figure 115. For this run, an arbitrary anisotropic, though symmetrical, weight matrix "W" is defined by a DATA statement in the main program and passed directly to the subroutine, using option IW=1. The four-fold symmetrical weight matrix dreamed up for this run is depicted below:

```
1 1 0 0 0 0 0 1 1
1 1 1 0 0 0 1 1 1
0 1 2 2 0 2 2 1 0
0 0 2 2 2 2 2 0 0
0 0 0 2 3 2 0 0 0
0 0 2 2 2 2 2 0 0
0 1 2 2 0 2 2 1 0
1 1 1 0 0 0 1 1 1
1 1 0 0 0 0 0 1 1
```

The model semivariogram is unknown, although it is clear from inspection of this matrix that the range should be about $9\sqrt{(2)}$ in the NE-SW and NW-SE directions, about 9 in the N-S and E-W directions, and generally more continuous (lower semivariogram values at short lags) in the NE-SW and NW-SE directions. This is borne out in the sample semivariograms and maps of Figures 114 and 115. The "crossed" anisotropy of the weight matrix is evidently responsible for the cross-hachured appearance of Figure 115. The concoction of unusual weight functions by trial and error might allow the production of simulations that would better

reproduce the types of phenomena illustrated in Figure 53. Before conditioning such a simulation, one should make sure that the sample semivariograms of the simulations reasonably match those of the real data.

As discussed in Section 3.3.1.2, this subroutine executes very slowly if the dimensions of matrices W and Y are large. Furthermore the entire simulated data set Y is kept in memory during execution of the program. For these reasons, subroutine CS2D is not recommended for very large simulations.

Listing of Subroutine CS2D:

```
      SUBROUTINE CS2D(NLI,NCO,NVX,DL,DC,DIAM,ITERM,IWRT,FMT,DSEED,N1,N2,
     *IW,YMEAN,YVAR,Y,W,LWLF,LWLL,LWCF,LWCL,X)
C
C
C     *****  "CIRCULAR SIMULATIONS, 2 DIMENSIONS"  *****
C
C     (WITH AN OPTION FOR GENERAL MOVING-AVERAGE SIMULATIONS)
C
C
C     SUBROUTINE TO GENERATE RECTANGULAR ARRAYS OF DATA DRAWN FROM
C     A 2-DIMENSIONAL STATIONARY GAUSSIAN RANDOM FUNCTION WITH AN
C     ISOTROPIC CIRCULAR SEMIVARIOGRAM MODEL
C
C     AUTOCORRELATIONS WITHIN THE DATA ARE ACHIEVED BY EFFECTIVELY
C     SMOOTHING AN ARRAY OF INDEPENDENT UNIFORM(0,1) RANDOM NUMBERS
C     BY AVERAGING INSIDE A DISCRETIZED CIRCULAR MOVING WINDOW.
C
C     A GEOMETRIC ANISOTROPY PARALLEL TO THE GRID DIRECTIONS CAN BE
C     IMPOSED BY APPROPRIATE ADJUSTMENTS OF THE GRID SPACINGS.
C
C     A GEOMETRIC ANISOTROPY OBLIQUE TO THE GRID DIRECTIONS, OR A
C     NONCIRCULAR SEMIVARIOGRAM MODEL, CAN BE IMPOSED BY PASSING A
C     PREDETERMINED MATRIX "W" FROM THE MAIN PROGRAM.  THIS ALLOWS THE
C     SUBROUTINE TO BE USED FOR GENERAL MOVING-AVERAGE SIMULATIONS IN
C     TWO DIMENSIONS.
C
C     UP TO 10 INDEPENDENT REALIZATIONS OF THE SAME RANDOM FUNCTION
C     CAN BE SIMULATED IN ONE CALL OF THE SUBROUTINE.
C
C     VERSION OF DECEMBER, 1984, BY G. R. LUSTER
C
C
C     *****  INPUT PARAMETERS (SEE ALSO "COMMON VARIABLES" BELOW)  *****
```

```
C
C
C     NLI       NUMBER OF LINES OR ROWS (E.G., TRENDING E-W) IN THE
C               SIMULATION DOMAIN
C
C     NCO       NUMBER OF COLUMNS (E.G., TRENDING N-S)
C
C     NVX       EQUAL TO COMMON VARIABLE NV -- MUST BE SPECIFIED IN THE
C               SUBROUTINE STATEMENT AS A DIMENSION OF ARRAY "Y"
C
C     DL        SPACING BETWEEN LINES (ALONG COLUMNS)
C               -- NOT USED IF IW=1
C
C     DC        SPACING BETWEEN COLUMNS (ALONG LINES)
C               -- NOT USED IF IW=1
C
C     DIAM      DIAMETER OF MOVING WINDOW (RANGE OF CIRCULAR
C               SEMIVARIOGRAM) -- NOT USED IF IW=1
C
C     ITERM     LOGICAL UNIT NUMBER FOR WRITING TERMINAL MESSAGES
C               TO SHOW HOW FAR EXECUTION HAS PROGRESSED
C               -- ****************  WARNING  ******************
C                  MESSAGES MARKED "C*****" SHOULD BE REMOVED FOR
C                  BATCH RUNS.
C
C     IWRT      THIS SUBROUTINE WRITES OUT THE SIMULATED VALUES IN
C               MATRIX "Y" (ON UNIT IOUT) ONLY IF WRT=1.
C               OUTPUT IS AS FOLLOWS:  LINE NUMBERS (LL=1,NLI) ARE
C               NESTED WITHIN COLUMN NUMBERS (LC=1,NCO), WHICH ARE
C               NESTED WITHIN VARIABLE (REALIZATION) NUMBERS (LV=1,NV).
C
C     FMT(8)    VARIABLE FORMAT FOR WRITING OUT THE SIMULATED VALUES
C               -- THE FORMAT IS PASSED FROM THE MAIN PROGRAM
C                  IN A VECTOR OF 8 4-BYTE CHARACTER STRINGS
C                  (TOTAL OF 32 SPACES AVAILABLE).
C               -- THESE CHARACTERS MUST INCLUDE BLANKS FOR UNUSED
C                  SPACES, E.G., '(1H ,5F10.4)                    '.
C
C     DSEED     INTEGER INITIALIZING THE RANDOM-NUMBER GENERATOR
C               -- TREATED AS A REAL*8 NUMBER
C               -- MUST LIE BETWEEN 1 AND (2**31)-1
C               -- UPDATED EACH TIME A NEW NUMBER IS GENERATED
C               -- ****************  WARNING  ****************
C                  THIS GENERATOR NEEDS AS MUCH AS 46-BIT ACCURACY
C                  AND MAY FAIL ON SOME MACHINES; REFER TO THE
C                  USER'S MANUAL.
C
C     N1,N2     DIMENSIONS OF WINDOW INDICATOR (WEIGHT) MATRIX "W"
C               (SEE "WORKING ARRAYS", BELOW)
C
C     IW        OPTION FOR GENERAL MOVING-AVERAGE SIMULATIONS:
C               IF IW=1, A PRESPECIFIED MATRIX "W" (SEE "WORKING
C               ARRAYS", BELOW) MUST BE PASSED FROM THE MAIN PROGRAM;
C               IN THIS CASE THE VALUES OF DL, DC, AND DIAM ARE
```

```
C                IRRELEVANT.
C
C
C     *****  OUTPUT DATA AND STATISTICS  *****
C
C
C     ANPT       THE SUM OF ALL ELEMENTS OF ARRAY "W" (EQUAL TO THE
C                NUMBER OF POINTS INSIDE THE CIRCULAR WINDOW, UNLESS IW=1)
C
C     BNPT       THE SUM OF SQUARES OF ALL ELEMENTS OF ARRAY "W" (EQUAL
C                TO ANPT UNLESS IW=1)
C
C     RAT        WINDOW DISCRETIZATION RATIO -- THE RATIO OF "ANPT"
C                TO THE PRODUCT OF THE DIMENSIONS OF ARRAY "W" BELOW
C                -- THE ACCURACY OF THE DISCRETIZATION GENERALLY
C                   IMPROVES AS THIS RATIO APPROACHES PI/4 = 0.7854,
C                   WHICH IS THE RATIO OF THE AREAS OF AN INSCRIBED
C                   CIRCLE AND ITS SQUARE.
C                -- NOT USED IF IW=1
C
C     YMEAN(NV)    MEAN OF THE SIMULATED VALUES IN EACH REALIZATION
C                  BEFORE ADJUSTMENT -- SHOULD BE CLOSE TO ANPT/2.0.
C                  AFTER ADJUSTMENT, THIS VECTOR CONTAINS THE ADJUSTED
C                  MEANS (SEE "Y" BELOW), WHICH ARE RETURNED BY THE
C                  SUBROUTINE.
C
C     YVAR(NV)     VARIANCE OF THE SIMULATED VALUES IN EACH REALIZATION
C                  BEFORE ADJUSTMENT -- IF RANGE << DOMAIN DIMENSIONS,
C                  THIS SHOULD BE CLOSE TO BNPT/12.0.  AFTER ADJUSTMENT,
C                  THIS VECTOR CONTAINS THE ADJUSTED VARIANCES (SEE "Y"
C                  BELOW), WHICH ARE RETURNED BY THE SUBROUTINE.
C
C     Y(NLI,NCO,NV)     3-D MATRIX OF SIMULATED VALUES, WRITTEN COLUMN-
C                       WISE, NLI VALUES AT A TIME; TOTAL OF NCO COLUMNS
C                       PER REALIZATION; NV REALIZATIONS
C                       -- VALUES FOR EACH REALIZATION ARE ADJUSTED TO:
C                          MEAN=(YMEAN(LV)-ANPT/2.0)/SQRT(BNPT/12.0)
C                          VARIANCE=YVAR(LV)/(BNPT/12.0)
C                       -- THESE ADJUSTMENTS RESULT APPROXIMATELY IN
C                          MEAN=0, SILL=1.  EXACT STANDARDIZATION IS
C                          AVOIDED TO PRESERVE FLUCTUATION VARIANCES.
C                       -- THE MATRIX AND FILE ARE STORED IN THE CORRECT
C                          ORDER FOR USE IN STANFORD SUBROUTINES HIST,
C                          GAM2, ETC.
C
C
C     *****  WORKING ARRAYS AND DIMENSIONS  *****
C
C
C     NOTE ON DIMENSIONS:  THE DIMENSIONS OF MATRIX W(N1,N2)
C                          MUST BE PASSED FROM THE MAIN PROGRAM
C                          IN THE CALL STATEMENT.
C                          IF IW=1, THESE VALUES MUST BE EQUAL TO THE
C                          VALUES SPECIFIED BELOW.  OTHERWISE THEY MAY
```

```
C                          BE GREATER.
C
C     W(2*NLH+1,2*NCH+1)   ARRAY OF INDICATORS DETERMINING WHICH
C                          GRID POINTS IN A SQUARE MOVING WINDOW
C                          LIE ON OR INSIDE THE INSCRIBED CIRCLE
C
C     LWLF(NLI+2*NLH)      LOOP RANGES USED TO DETERMINE WHICH
C     LWLL(NLI+2*NLH)      PARTS OF THE SQUARE MATRIX SURROUNDING
C     LWCF(NCO+2*NCH)      THE CIRCULAR MOVING WINDOW OVERLAP THE
C     LWCL(NCO+2*NCH)      SIMULATION DOMAIN AS THE WINDOW MOVES
C
C     X(NV)      NV UNIFORM(0,1) RANDOM NUMBERS CONTRIBUTING TO THE
C                NV SIMULATED REALIZATIONS
C                -- GENERATED INDEPENDENTLY AT EACH GRID LOCATION
C
C     WHERE      NLH=INT(0.5*DIAM/DL)    INTEGER HALF-SPANS (HALF-AXES)
C                NCH=INT(0.5*DIAM/DC)    OF THE CIRCULAR WINDOW
C
C                OR, IF IW=1,
C
C                NLH=(N1-1)/2            WHERE N1 AND N2 ARE THE
C                NCH=(N2-1)/2            DIMENSIONS OF THE PRESPECIFIED
C                                        MATRIX W(N1,N2).  NOTICE THAT
C                                        N1 AND N2 MUST BE ODD.
C
C
C     *****  COMMON VARIABLES  *****
C
C
C     INP        INPUT UNIT
C
C     IOUT       OUTPUT UNIT
C
C     DUMMY      NOT USED
C
C     NV         NUMBER OF INDEPENDENT REALIZATIONS (UP TO 10)
C                -- EACH REALIZATION MUST HAVE THE SAME PARAMETERS.
C
C     NAM(10)    8-CHARACTER VARIABLE NAMES (UP TO 10 OF THEM)
C                CORRESPONDING TO THE NV INDEPENDENT REALIZATIONS
C
C
C     *****  INSPECTION OF INPUT PARAMETERS  *****
C
C
      DIMENSION Y(NLI,NCO,NVX),FMT(8),W(N1,N2),X(1),YMEAN(1),YVAR(1)
      DIMENSION LWLF(1),LWLL(1),LWCF(1),LWCL(1)
C
      DOUBLEPRECISION DSEED,NAM(10)
      COMMON INP,IOUT,DUMMY,NV,NAM
C
C     WRITE OUT THE INPUT PARAMETERS
C
      WRITE(IOUT,1000) NLI,NCO,NV,(NAM(I),I=1,NV)
```

```
      WRITE(IOUT,1001) IWRT,FMT,DSEED,NLI,NCO,NV,N1,N2
      IF(IW.EQ.1) WRITE(IOUT,2001)
C
C     LOOP PARAMETERS USED BELOW
C
      NL=N1
      NC=N2
C
C     THE FOLLOWING PARAMETERS, AND THE DISCRETIZATION SECTION BELOW,
C     ARE NOT USED IF IW=1.
C
      IF(IW.EQ.1) GO TO 150
      WRITE(IOUT,3001) DL,DC,DIAM
C
C     TEST GRID FINENESS WITH RESPECT TO THE SEMIVARIOGRAM RANGE
C     -- WE WANT AT LEAST 3 LAGS INSIDE THE RANGE FOR THE RANGE
C        TO BE VISIBLE IN SAMPLE SEMIVARIOGRAMS IN ANY DIRECTION
C
      IL=INT(DIAM/DL)
      IC=INT(DIAM/DC)
      I2=INT(DIAM/SQRT(DL*DL+DC*DC))
      IF(IL.LT.4) WRITE(IOUT,1002) IL
      IF(IC.LT.4) WRITE(IOUT,1003) IC
      IF(I2.LT.4) WRITE(IOUT,1004)
C
C
C     *****  DISCRETIZATION OF CIRCULAR MOVING WINDOW  *****
C
C
C     INTEGER HALF-SPANS (HALF-AXES) OF THE CIRCULAR MOVING WINDOW
C     -- CALCULATED HERE IN CASE LARGER VALUES ARE GIVEN FOR N1 AND N2
C
      NLH=INT(0.5*DIAM/DL)
      NCH=INT(0.5*DIAM/DC)
C
C     CALCULATE INDICATORS FOR POINTS INSIDE THE CIRCULAR MOVING WINDOW
C     -- THE WINDOW HAS SPANS (AXES) OF LENGTHS (IN TERMS OF GRID
C        POINTS) NL = 2*NLH+1, NC = 2*NCH+1
C     -- TO DISCRETIZE THE WINDOW, WE ASSIGN A VALUE OF "1" TO THE
C        INDICATOR "W" IF THE POINT IN QUESTION IS WITHIN THE RADIUS
C        (DIAM/2) OF THE CIRCLE
C
C     RAD2 = RADIUS**2 OF CIRCLE, (NL,NC) = DIMENSIONS OF WINDOW MATRIX
C
      RAD2=DIAM*DIAM/4.0
      NL=2*NLH+1
      NC=2*NCH+1
C
C     INITIALIZE INDICATOR ARRAY "W", OCCUPYING A SQUARE
C     CIRCUMSCRIBING THE CIRCULAR WINDOW
C
      DO 90 I = 1,NL
      DO 90 J = 1,NC
C
```

```
   90 W(I,J)=1.0
C
C     IF EITHER NLH OR NCH IS ZERO, WE HAVE A 1-D WINDOW, GENERATING
C     NO CORRELATIONS IN THE OTHER DIRECTION, AND HAVING NO OFF-AXIS
C     POINTS.  IF BOTH ARE ZERO, WE ARE JUST INEFFICIENTLY FILLING
C     THE ARRAY "Y" WITH UNIFORM RANDOM NUMBERS.
C
      IF(NLH.EQ.0.OR.NCH.EQ.0) GO TO 150
C
C     CHECK OFF-AXIS ELEMENTS OF "W" TO SEE WHETHER THEY LIE ON OR
C     INSIDE THE CIRCLE -- IF OUTSIDE, SET TO ZERO
C
      DO 100 LIC=1,NCH
      DO 100 LIL=1,NLH
C
      WIND=1.0
      HYPOT=LIL*LIL*DL*DL + LIC*LIC*DC*DC
      IF(HYPOT.GT.RAD2) WIND=0.0
C
C     SYMMETRICAL OFF-AXIS ASSIGNMENTS IN 4 QUADRANTS OF "W"
C
      W(NLH+1-LIL,NCH+1-LIC)=WIND
      W(NLH+1+LIL,NCH+1-LIC)=WIND
      W(NLH+1-LIL,NCH+1+LIC)=WIND
  100 W(NLH+1+LIL,NCH+1+LIC)=WIND
C
C     CALCULATE SUM AND SUM OF SQUARES OF WINDOW ELEMENTS
C
  150 ANPT=0.0
      BNPT=0.0
C
      DO 170 I=1,NL
      DO 170 J=1,NC
C
      ANPT=ANPT+W(I,J)
  170 BNPT=BNPT+W(I,J)**2
C
C     DISCRETIZATION RATIO -- THE CLOSER TO PI/4, THE BETTER
C
      RAT=ANPT/FLOAT(NL*NC)
C*****
C     COMMENT TO TERMINAL
      WRITE(ITERM,8002)
 8002 FORMAT(' STARTING MOVING WINDOW SIMULATION')
C*****
C
C
C     *****  GENERATE REALIZATIONS BY MOVING-WINDOW TECHNIQUE  *****
C
C
C     ARRAY Y(NLI,NCO,NV) INITIALLY IS USED TO ACCUMULATE THE SUMS
C     OF CONTRIBUTIONS FROM THE RANDOM-NUMBER ARRAY (OF DIMENSIONS
C     NLI+2*NLH, NCO+2*NCH, NV), OF WHICH ONLY ONE SET OF NV VALUES
C     IS STORED AT ANY ONE TIME.  EACH OF THE NV REALIZATIONS WITHIN
```

```
C     Y IS THEN ADJUSTED TO APPROXIMATELY MEAN=0, SILL=1.
C     EXACT STANDARDIZATION OF THE SIMULATED REALIZATIONS TO MEAN=0,
C     VARIANCE=1 IS NOT PERFORMED.
C
C     NOTICE THAT WE ARE ADDING THE CENTRAL X VALUE TO ALL Y VALUES
C     IN THE WINDOW, INSTEAD OF ADDING ALL X VALUES IN THE WINDOW
C     TO THE CENTRAL Y VALUE -- ACHIEVES SAME RESULT.
C
C     INITIALIZE ARRAY "Y"
C
  180 DO 200 LIC=1,NCO
      DO 200 LIL=1,NLI
      DO 200 LV=1,NV
C
  200 Y(LIL,LIC,LV)=0.0
C
C     WE DO NOT WANT TO LOOP OVER THE WHOLE WINDOW WHEN PART OF IT
C     LIES OUTSIDE THE SIMULATION DOMAIN, SO WE CALCULATE AND STORE
C     WINDOW LOOP RANGES AS A FUNCTION OF POSITION OF THE WINDOW
C     CENTER IN THE RANDOM-NUMBER ARRAY (COORDS. ML,MC).
C
C     WINDOW LINE LOOP RANGES LWLF,LWLL AS FUNCTIONS OF ML
C
      MLL=NLI+NL-1
      DO 510 ML=1,MLL
      LWLF(ML)=MAXO(1,1+NL-ML)
  510 LWLL(ML)=MINO(NL,1+MLL-ML)
C
C     WINDOW COLUMN LOOP RANGES LWCF,LWCL AS FUNCTIONS OF MC
C
      MCL=NCO+NC-1
      DO 520 MC=1,MCL
      LWCF(MC)=MAXO(1,1+NC-MC)
  520 LWCL(MC)=MINO(NC,1+MCL-MC)
C
C     LOOP WINDOW CENTER OVER RANDOM-NUMBER ARRAY,
C     UPDATING WINDOW LOOP RANGES EACH TIME
C
      DO 530 MC=1,MCL
C*****
C     COMMENT TO TERMINAL
      WRITE(ITERM,8003) MC
 8003 FORMAT(' R. N. COLUMN NO.',I5)
C*****
C
      NWCF=LWCF(MC)
      NWCL=LWCL(MC)
C
      DO 530 ML=1,MLL
C
      NWLF=LWLF(ML)
      NWLL=LWLL(ML)
C
C     GENERATE NV UNIFORM(0,1) RANDOM NUMBERS AT EACH POSITION
```

```
C      IN THE RANDOM-NUMBER ARRAY
C      SEE *****WARNING***** IN DESCRIPTION OF INPUT PARAMETERS (ABOVE)
C
       DO 540 LV=1,NV
       DSEED=DMOD(16807.D0*DSEED,2147483647.D0)
  540 X(LV)=DSEED/2147483648.D0
C
C      ADD EACH OF THE NV RANDOM NUMBERS X TO THE ELEMENTS OF Y THAT
C      INTERSECT THE WINDOW ON EACH OF THE NV LEVELS OF Y
C
       DO 530 LWC=NWCF,NWCL
       DO 530 LWL=NWLF,NWLL
       DO 530 LV=1,NV
C
  530 Y(ML-NL+LWL,MC-NC+LWC,LV)=Y(ML-NL+LWL,MC-NC+LWC,LV)+X(LV)
     **W(LWL,LWC)
C*****
C      COMMENT TO TERMINAL
       WRITE(ITERM,8004)
 8004 FORMAT(' STARTING STATS & OUTPUT')
C*****
C
C
C      *****  STATISTICS OF THE SIMULATIONS  *****
C
C
C      CALCULATE UNADJUSTED MEANS AND VARIANCES
C
       DO 600 LV=1,NV
C
       YM=0.0
       YV=0.0
       TOT=0.0
C
       DO 700 LC=1,NCO
       DO 700 LL=1,NLI
C
       TOT=TOT+1.0
       YM=YM+Y(LL,LC,LV)
       IF(TOT.LT.2.0) GO TO 700
       YV=YV+(YM-TOT*Y(LL,LC,LV))**2/(TOT*(TOT-1.0))
C
  700 CONTINUE
C
       YMEAN(LV)=YM/TOT
  600 YVAR(LV)=YV/TOT
C
C      THEORETICAL UNADJUSTED MEAN AND SILL
C      -- EXPECTATIONS FOR A LINEAR COMBINATION OF INDEPENDENT
C         UNIFORM(0,1) RANDOM DATA
C
       EXMEAN=ANPT/2.0
       EXVAR=BNPT/12.0
       EXSTD=SQRT(EXVAR)
```

```
C
      IF(IW.NE.1) WRITE(IOUT,1005) EXMEAN,EXVAR,ANPT,RAT
      IF(IW.EQ.1) WRITE(IOUT,2005) EXMEAN,EXVAR
C
C     WRITE UNADJUSTED MEANS AND VARIANCES
C
      WRITE(IOUT,1006) (NAM(I),I=1,NV)
      WRITE(IOUT,2006) (YMEAN(I),I=1,NV)
      WRITE(IOUT,3006) (YVAR(I),I=1,NV)
C
C     WRITE ADJUSTED MEANS AND VARIANCES
C
      DO 750 LV=1,NV
      YMEAN(LV)=(YMEAN(LV)-EXMEAN)/EXSTD
  750 YVAR(LV)=YVAR(LV)/EXVAR
      WRITE(IOUT,2007) (YMEAN(I),I=1,NV)
      WRITE(IOUT,3007) (YVAR(I),I=1,NV)
C
C     WRITE FINAL VALUE OF DSEED
C
      WRITE(IOUT,1008) DSEED
C
C
C     *****  ADJUSTMENTS AND OUTPUT  *****
C
C
C     ADJUST THE SIMULATIONS APPROXIMATELY TO MEAN=0, SILL=1, AND
C     WRITE THE RESULTS, LINES WITHIN COLUMNS WITHIN VARIABLES
C
      IF(IWRT.EQ.1) WRITE(IOUT,1009)
C
      DO 900 LV=1,NV
C
      DO 900 LC=1,NCO
      DO 800 LL=1,NLI
C
  800 Y(LL,LC,LV)=(Y(LL,LC,LV)-EXMEAN)/EXSTD
C
      IF(IWRT.EQ.1) WRITE(IOUT,FMT) (Y(LL,LC,LV),LL=1,NLI)
  900 CONTINUE
C
C
C     *****  OUTPUT FORMATS  *****
C
C
 1000 FORMAT('1SUBROUTINE CS2D -- TWO-DIMENSIONAL MOVING-AVERAGE ',
     1'SIMULATION'//'0THE INPUT PARAMETERS ARE:'//' GRID ',
     2'DIMENSIONS OF SIMULATION DOMAIN:'/'      NUMBER OF LINES = ',
     3'NLI = ',I6/'      NUMBER OF COLUMNS = NCO = ',I4//
     4' NUMBER OF INDEPENDENT REALIZATIONS (ALL WITH THE SAME ',
     5'PARAMETERS) = NV = ',I2//' NAMES OF THE REALIZATIONS:'/' ',
     65(A8,3X)/' ',5(A8,3X))
 1001 FORMAT('0LISTING OF SIMULATED DATA BELOW?',
     1' ("1" MEANS YES) = IWRT =',I2/' FORMAT FOR ',
```

```
     2'DATA LISTING = FMT = ',8A4//' INITIAL SEED FOR RANDOM-NUMBER ',
     3'GENERATOR = DSEED = ',D20.12//' DIMENSIONS OF DATA MATRIX "Y" =',
     4' (NLI,NCO,NV) = (',I4,',',I4,',',I4,')'//' DIMENSIONS OF WINDOW',
     5' INDICATOR (WEIGHT) MATRIX "W" = (N1,N2) = (',I3,',',I3,')')
 2001 FORMAT('0A MOVING-AVERAGE WEIGHT FUNCTION (MATRIX "W") ',
     1'HAS BEEN SUPPLIED BY THE USER.'/)
 3001 FORMAT('0GRID SPACINGS:'/
     1'       BETWEEN LINES (ALONG COLUMNS) = DL = ',F10.4/
     2'       BETWEEN COLUMNS (ALONG LINES) = DC = ',F10.4//' VARIO',
     3'GRAM RANGE (DIAMETER OF CIRCULAR WINDOW) = DIAM = ',F10.4//
     4' THESE REALIZATIONS SHOULD EXHIBIT CIRCULAR SEMIVARIOGRAMS.'/)
 1002 FORMAT('0NOTICE THAT THE RANGE IS ONLY',I2,' LAGS LONG ',
     1'BETWEEN LINES (ALONG COLUMNS).')
 1003 FORMAT('0NOTICE THAT THE RANGE IS ONLY',I2,' LAGS LONG ',
     1'BETWEEN COLUMNS (ALONG LINES).')
 1004 FORMAT('0NOTICE THAT THE RANGE IS SHORT (IN TERMS OF GRID ',
     1'SPACINGS) IN DIAGONAL DIRECTIONS.')
 1005 FORMAT('0'//'0STATISTICS FOR SIMULATED REALIZATIONS: '/
     1' (DATA ARE APPROXIMATELY NORMAL(0,1) AFTER ADJUSTMENT)'//
     2'       THEORETICAL UNADJUSTED MEAN    =   ANPT/2.0  = ',F12.4/
     3'       THEORETICAL UNADJUSTED SILL    =   ANPT/12.0 = ',F12.4/
     4'       NUMBER OF POINTS WITHIN CIRCULAR WINDOW = ANPT = ',F12.4/
     5'       DISCRETIZATION RATIO = RAT = ',F7.4,'   (APPROACHES PI/4 ',
     6'= 0.7854)')
 2005 FORMAT('0'//'0STATISTICS FOR SIMULATED REALIZATIONS: '/
     1' (DATA ARE APPROXIMATELY NORMAL(0,1) AFTER ADJUSTMENT)'//
     2'       THEORETICAL UNADJUSTED MEAN = ',F12.4/
     3'       THEORETICAL UNADJUSTED SILL = ',F12.4)
 1006 FORMAT('0VARIABLE NAME:  ',10(2X,A8))
 2006 FORMAT('0UNADJ. MEAN:    ',10F10.4)
 3006 FORMAT(' UNADJ. VARIANCE:',10F10.4)
 2007 FORMAT('0ADJ. MEAN:      ',10F10.4)
 3007 FORMAT(' ADJ. VARIANCE:  ',10F10.4)
 1008 FORMAT('0FINAL VALUE OF RANDOM-NUMBER SEED = DSEED = ',D20.12)
 1009 FORMAT('1LISTING OF SIMULATED DATA IN FORMAT "FMT"'//)
C
      RETURN
      END
```

```
SUBROUTINE CS2D -- TWO-DIMENSIONAL MOVING-AVERAGE SIMULATION


THE INPUT PARAMETERS ARE:

GRID DIMENSIONS OF SIMULATION DOMAIN:
     NUMBER OF LINES = NLI =      50
     NUMBER OF COLUMNS = NCO =    50

NUMBER OF INDEPENDENT REALIZATIONS (ALL WITH THE SAME PARAMETERS) = NV =  2

NAMES OF THE REALIZATIONS:
TEST ONE    TEST TWO

LISTING OF SIMULATED DATA BELOW? ("1" MEANS YES) = IWRT = 1
FORMAT FOR DATA LISTING = FMT = (1H ,5F10.4)

INITIAL SEED FOR RANDOM-NUMBER GENERATOR = DSEED =   0.170200000000D 04

DIMENSIONS OF DATA MATRIX "Y" = (NLI,NCO,NV) = (  50,  50,   2)
DIMENSIONS OF WINDOW INDICATOR (WEIGHT) MATRIX "W" = (N1,N2) = (  9,  9)

GRID SPACINGS:
     BETWEEN LINES (ALONG COLUMNS) = DL =      1.0000
     BETWEEN COLUMNS (ALONG LINES) = DC =      0.5000

VARIOGRAM RANGE (DIAMETER OF CIRCULAR WINDOW) = DIAM =      3.1000

THESE REALIZATIONS SHOULD EXHIBIT CIRCULAR SEMIVARIOGRAMS.


NOTICE THAT THE RANGE IS ONLY 3 LAGS LONG BETWEEN LINES (ALONG COLUMNS).

NOTICE THAT THE RANGE IS SHORT (IN TERMS OF GRID SPACINGS) IN DIAGONAL DIRECTIONS.


STATISTICS FOR SIMULATED REALIZATIONS:
(DATA ARE APPROXIMATELY NORMAL(0,1) AFTER ADJUSTMENT)

     THEORETICAL UNADJUSTED MEAN    =   ANPT/2.0  =       8.5000
     THEORETICAL UNADJUSTED SILL    =   ANPT/12.0 =       1.4167
     NUMBER OF POINTS WITHIN CIRCULAR WINDOW = ANPT =      17.0000
     DISCRETIZATION RATIO = RAT =  0.8095   (APPROACHES PI/4 = 0.7854)

VARIABLE NAME:     TEST ONE  TEST TWO

UNADJ. MEAN:         8.4112    8.4501
UNADJ. VARIANCE:     1.4105    1.5417

ADJ. MEAN:          -0.0746   -0.0419
ADJ. VARIANCE:       0.9957    1.0883

FINAL VALUE OF RANDOM-NUMBER SEED = DSEED =   0.790862172000D 09
```

Figure 107: Output of the first demonstration run of subroutine CS2D.

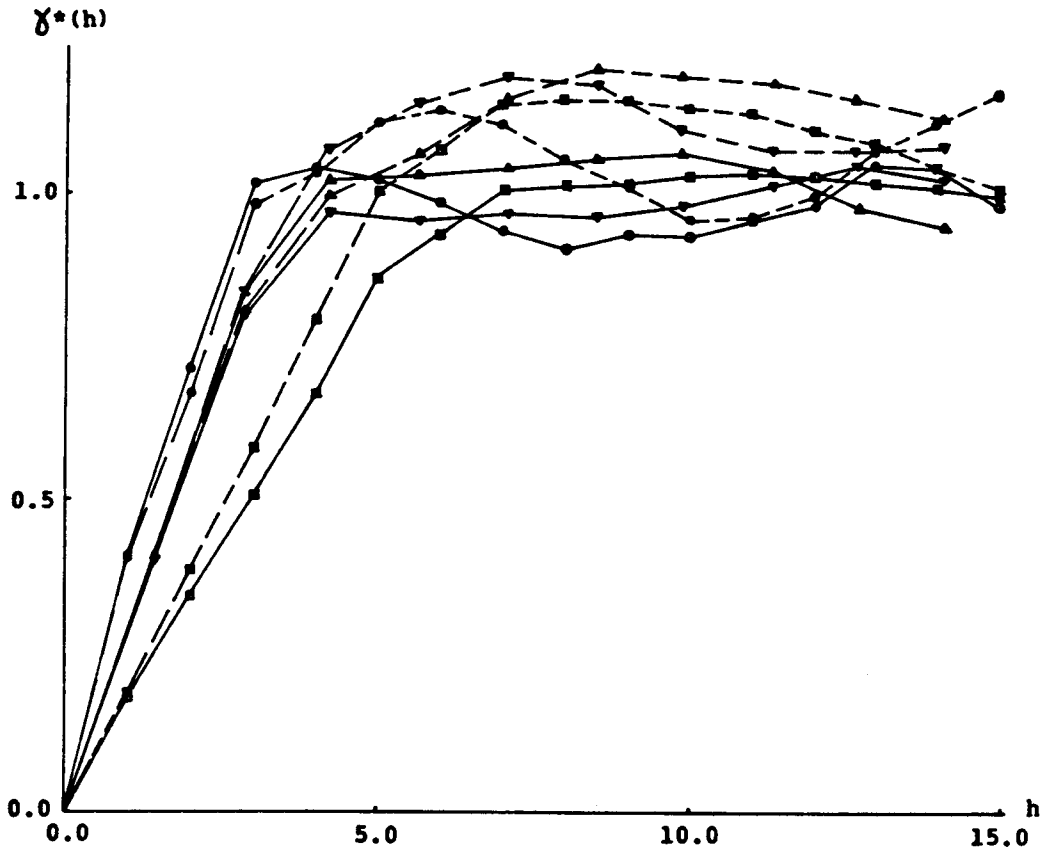


Figure 108: Sample semivariograms from the first demonstration run of subroutine CS2D. Solid lines denote semivariograms from the first simulated realization, dashed lines the second realization. The symbol "□" denotes the east-west direction, "○" denotes the north-south direction, and "△" and "▽" denote the 45-degree diagonal directions.

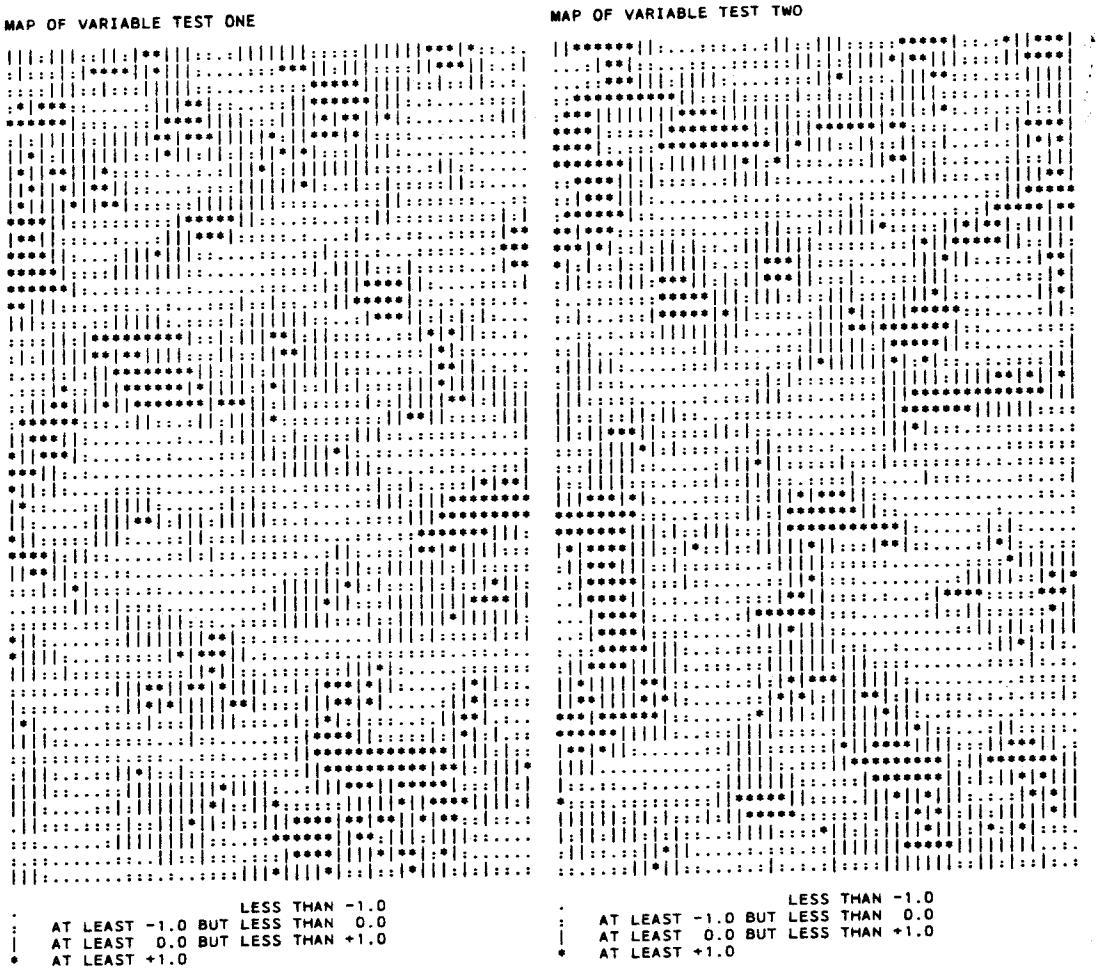
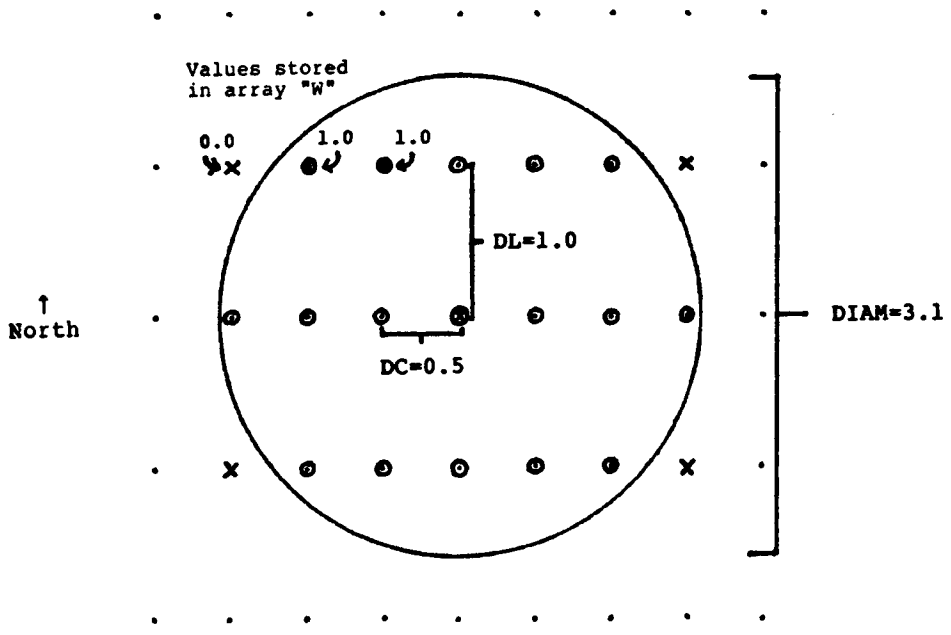


Figure 109: Maps of the realizations produced in the first demonstration run of subroutine CS2D.

ANPT = 17.0 points (circled below) inside circular window



NLH = integer part of (0.5*DIAM/DL)
= INT(0.5*3.1/1.0) = 1

NCH = integer part of (0.5*DIAM/DC)
= INT(0.5*3.1/0.5) = 3

Array W(2*NLH+1,2*NCH+1) = W(3,7) contains the indicators:
1.0 = on or inside the circle, 0.0 = outside the circle

Model semivariogram ranges, in grid spacings:
North-South Range = DIAM/1.0 = 3.1
East - West Range = DIAM/0.5 = 6.2
Diagonal Range = DIAM / $\sqrt{[(0.5)^2 + (1.0)^2]}$ = 2.77

Figure 110: Example of a discretized circular moving window, using parameters from the first demonstration run.

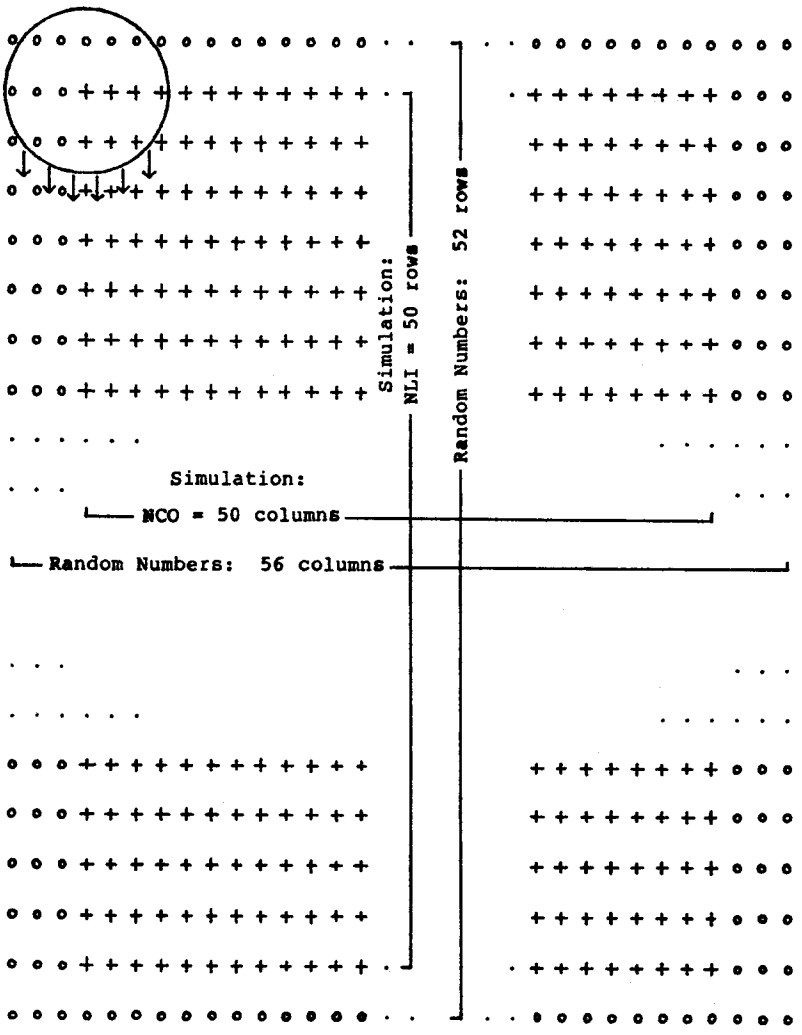


Figure 111: Random-number and simulation grids used by Subroutine CS2D. The inner grid (marked by +'s) represents one of the NV=2 levels of the simulation array "Y" with dimensions 50x50x2 in the first demonstration run. The outer grid (marked by o's) represents a level of the random-number array "X" with dimensions 52x56x2 in the first demonstration run. Actually, only NV=2 values of X are stored at any one time. The moving window is shown in its initial position in the upper-left corner of array X. The direction of movement is columnwise top to bottom (indicated by arrows), then left to right. In the first demonstration run, two independent realizations are generated simultaneously; only one set of corresponding grid locations is shown here.

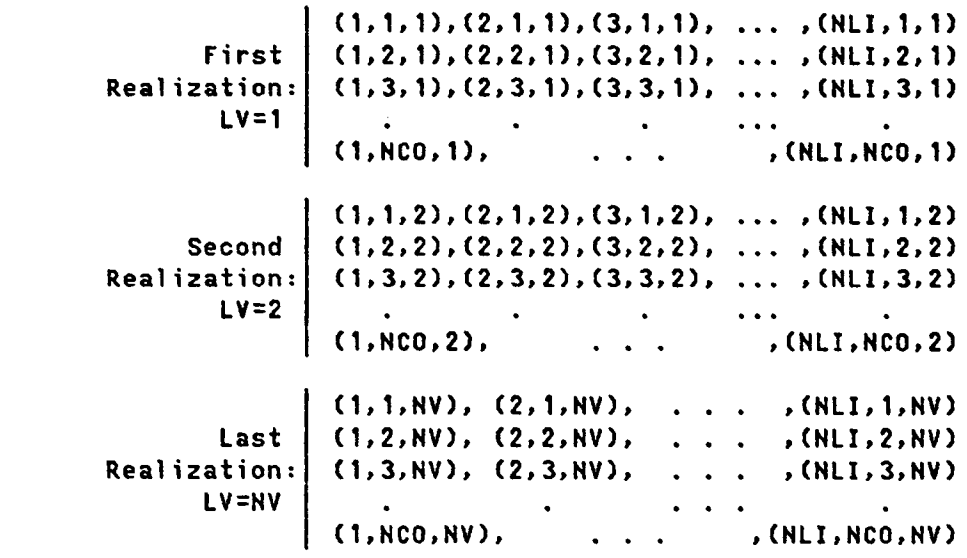


Figure 112: Arrangement of the elements of data matrix "Y" in the data output vector written by subroutine CS2D. The vector is written out one column of NLI values at a time, as represented by the individual rows above. Each row above corresponds to a single execution of the WRITE statement. This peculiar order is the same order in which the matrix Y is stored, as a vector, in the machine. In the two demonstration runs reported in this appendix, there are NLI=50 lines, NCO=50 columns, and NV=2 realizations. The output array, as a vector of length 5000, thus begins with Y(1,1,1) and ends with Y(50,50,2). An EQUIVALENCE statement in the main program allows this array to be passed directly to subroutines HIST and GAM2 as vector VR(5000).

```
SUBROUTINE CS2D -- TWO-DIMENSIONAL MOVING-AVERAGE SIMULATION


THE INPUT PARAMETERS ARE:

GRID DIMENSIONS OF SIMULATION DOMAIN:
     NUMBER OF LINES = NLI =     50
     NUMBER OF COLUMNS = NCO =   50

NUMBER OF INDEPENDENT REALIZATIONS (ALL WITH THE SAME PARAMETERS) = NV =  2

NAMES OF THE REALIZATIONS:
TEST ONE    TEST TWO

LISTING OF SIMULATED DATA BELOW? ("1" MEANS YES) = IWRT = 1
FORMAT FOR DATA LISTING = FMT = (1H ,5F10.4)

INITIAL SEED FOR RANDOM-NUMBER GENERATOR = DSEED =   0.170200000000D 04

DIMENSIONS OF DATA MATRIX "Y" = (NLI,NCO,NV) = (  50,  50,   2)
DIMENSIONS OF WINDOW INDICATOR (WEIGHT) MATRIX "W" = (N1,N2) = (  9,  9)

A MOVING-AVERAGE WEIGHT FUNCTION (MATRIX "W") HAS BEEN SUPPLIED BY THE USER.



STATISTICS FOR SIMULATED REALIZATIONS:
(DATA ARE APPROXIMATELY NORMAL(0,1) AFTER ADJUSTMENT)

     THEORETICAL UNADJUSTED MEAN =      33.5000
     THEORETICAL UNADJUSTED SILL =       9.4167

VARIABLE NAME:     TEST ONE  TEST TWO

UNADJ. MEAN:        32.8712   33.3247
UNADJ. VARIANCE:     8.9063   10.3527

ADJ. MEAN:          -0.2049   -0.0571
ADJ. VARIANCE:       0.9458    1.0994

FINAL VALUE OF RANDOM-NUMBER SEED = DSEED =   0.2018887375000 10
```

Figure 113: Output of the second demonstration run of subroutine CS2D.

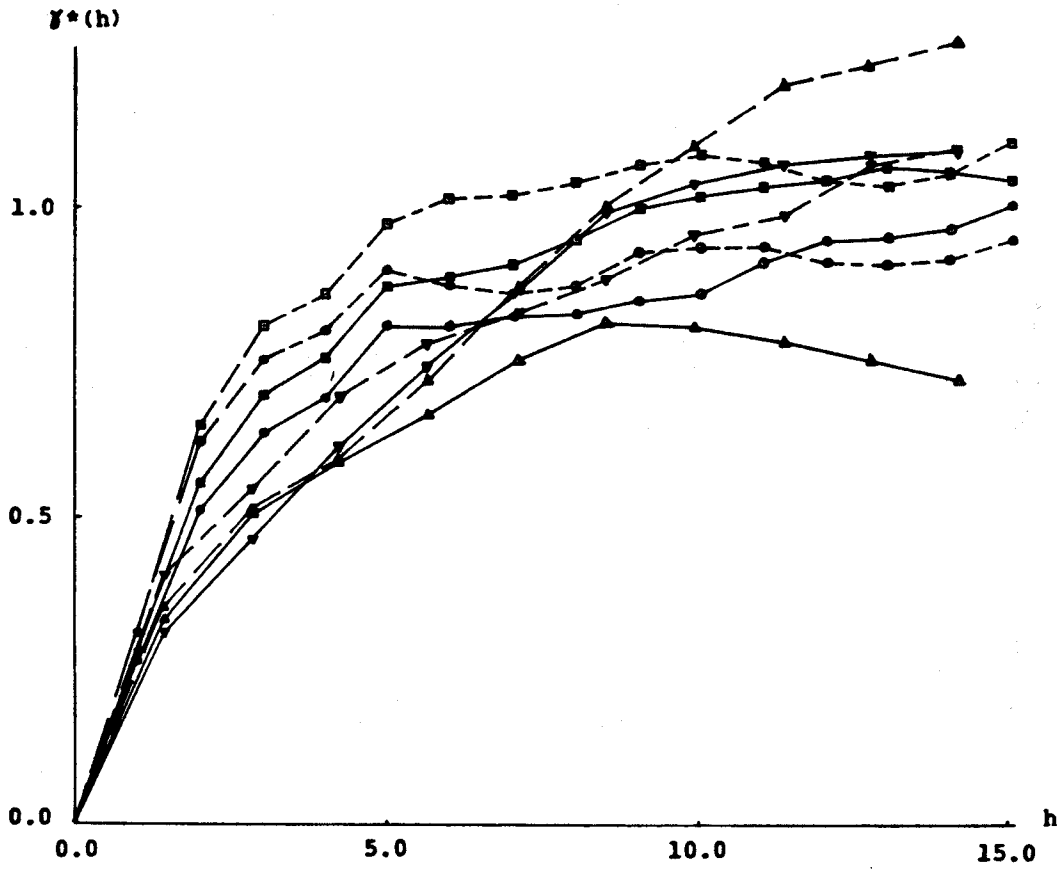


Figure 114: Sample semivariograms from the second demonstration run of subroutine CS2D. Solid lines denote semivariograms from the first simulated realization, dashed lines the second realization. The symbol "□" denotes the east-west direction, "○" denotes the north-south direction, and "△" and "▽" denote the 45-degree diagonal directions.

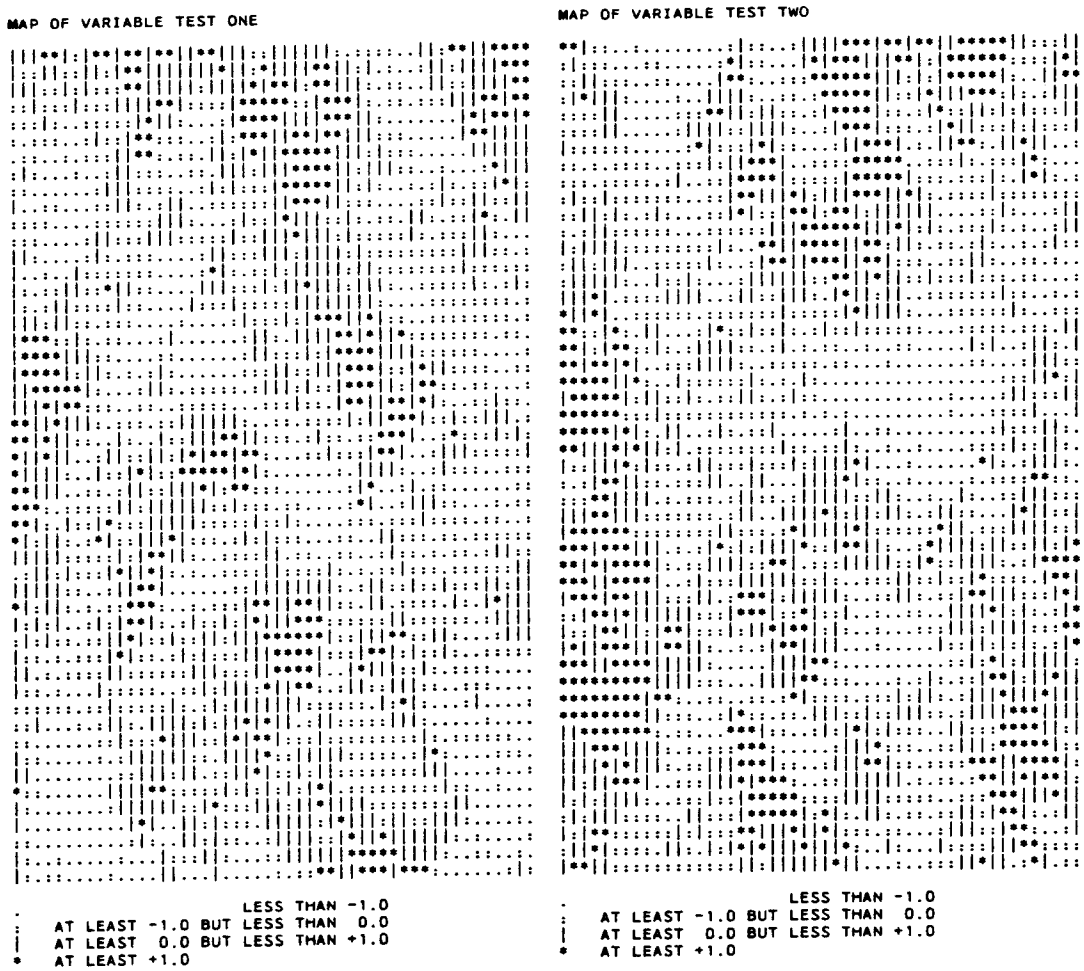


Figure 115: Maps of the realizations produced in the second demonstration run of subroutine CS2D.

Appendix B

SUBROUTINE TB3D: "TURNING BANDS, 3 DIMENSIONS"

Subroutine TB3D simulates up to ten independent realizations from a stationary gaussian random function with mean=0.0, semivariogram sill=1.0, using the turning-bands method (Section 3.3.1.4). Each simulated realization consists of data located on a three-dimensional rectangular grid filling a rectangular-prism-shaped simulation domain. The semivariogram model may be either spherical (parameter A > 0.0) or exponential (A < 0.0), the absolute value of A being the horizontal scale parameter (range, in the spherical case) of the semivariogram (Journel and Huijbregts, 1978, p. 164). A geometric anisotropy parallel to the grid directions can be imposed by manipulating the grid spacings (DL,DC,DN) as illustrated in Figure 47, Section 3.7.4.1, and in Appendix A. This subroutine is an extensively revised version of subroutine SIMUL, listed in Journel and Huijbregts (1978, p. 538).

Unlike Subroutine CS2D (Appendix A), Subroutine TB3D can be used for very large simulations, for two reasons: first, the turning-bands algorithm executes rapidly, and second, the whole array of simulated data is not kept in memory simultaneously. The subroutine calculates only NCO simulated values at a time and then immediately writes them onto sequential output file IDAT, according to a format FMT specified by the user. The order in which the simulated values are written is indicated in Figure 116. A report of important parameters and

statistics is written to a different output file, IOUT. The value of parameter IWRT determines what optional information will be included in this report, including statistics of the 15 line simulations and a printout of the simulated values (as in file IDAT).

Subroutine TB3D automatically adjusts the simulated values to an expected mean=0.0, sill=1.0, but exact standardization is not performed, in order to preserve natural fluctuation variances and to avoid biasing the short-scale structure of the semivariogram. (See Section 3.7.2 and Appendix A for more complete discussions.) The method used for adjustment is described by Journel and Huijbregts (1978, p. 536).

The same multiplicative congruential pseudorandom-number generator used in Subroutine CS2D is employed here, and the caveat expressed in Appendix A applies here as well.

Results of a single demonstration run of Subroutine TB3D are included below. In this run, subroutine TB3D is called twice. The first call simulates two independent realizations from a spherical model with north-south (across rows) range 5.0, east-west (across columns) range 2.0, and vertical (across levels) range 5.0, using parameters A=1.0, (DL,DC,DN) = (0.2, 0.5, 0.2). (Ranges are expressed in cubic grid spacings.) The dimensions of the discretized simulation domain are (rows, columns, levels) = (NLI,NCO,NLV) = (20,10,5). This choice of spacings and dimensions will not result in very good reproduction of the model semivariogram, for two reasons: the spacing across columns is very coarse in comparison to the semivariogram range, yielding poor short-scale reproduction in that direction; and the size of the domain is very small in comparison to the range, especially in the vertical

direction, yielding poor large-scale reproduction (wild fluctuations about the sill). The essential "printed" output (excluding line statistics and data printout) directed to file IOUT during the first call of the subroutine is displayed in Figure 117, and the sample semivariograms in Figure 118. The data file IDAT=9 is rewound by the main program and the 2000 data simulated by the first call of TB3D are read into vector VR(2000) for processing by subroutines HIST and GAM3 (not listed). This might not be practical in a big production run of TB3D, as vector VR might be too large and execution of GAM3 to produce exhaustive sample semivariograms might take too much time.

The second call of TB3D simulates two independent realizations of an exponential model with a "practical" range equal to the spherical range used in the first call. All other parameters are the same. The corresponding printed output and sample semivariograms are displayed in Figures 119 and 120. The simulated data in this case are routed to a different output file, IDAT=10, but file IOUT=8 is used for all "printed" information from both the spherical and the exponential simulations. A listing of the main program for the demonstration run is provided below:

```
$WATFIV TIME=300
$ASSIGN  8 TO FILE TB3DOUT OUTPUT
$ASSIGN  9 TO FILE TB3DSPH INOUT
$ASSIGN 10 TO FILE TB3DEXP INOUT
C*****
C*****  INTERACTIVE (ORVYL) PROGRAM -- FOR BATCH RUNS, DELETE
C*****  ALL MESSAGES TO TERMINAL MARKED BY "C*****" BELOW.
C*****
C
C     DEMONSTRATION RUN OF SUBROUTINE TB3D -- MAIN PROGRAM
C
C     DIMENSIONS FOR SUBROUTINE TB3D
C
      DIMENSION Y(161),DX(600),YMEAN(2),YVAR(2),W(161),FMT(8)
C
```

```
C      DIMENSIONS FOR SUBROUTINES HIST AND GAM3
C
       DIMENSION VR(2000), NP(80),GAM(80),UG(80),ND(2),UD(2),VD(2)
       DIMENSION T(2),ID(4),JD(4),KD(4),WT(2000),XFR(50)
C
       DOUBLEPRECISION NAM(10), N1S,N2S,N1E,N2E, DSEED
C
       COMMON INP,IOUT,DUMMY,NV,NAM
C
C      DATA FOR OUTPUT FORMAT, VARIABLE NAMES, R. N. GENERATOR SEED
C
       DATA FMT/'(1H ',',5F1','0.4)',','    ',','    ',','    ',','    ',','    '/
       DATA N1S,N2S,N1E,N2E/'SPH.   1','SPH.   2','EXP.   1','EXP.   2'/
       DATA DSEED/52086.D0/
C
C      DATA FOR GRID DIMENSIONS, GRID SPACINGS
C
       DATA NLI,NCO,NLV,DL,DC,DN/20,10,5,0.2,0.5,0.2/
C
C      DATA FOR VARIO. DIRECTIONS, VARIO. TESTS, HIST. WTS.
C
       DATA ID/1,0,0,1/,JD/0,1,0,1/,KD/0,0,1,1/
       DATA T/2*-100.0/,WT/2000*1.0/
C
C      NUMBERS IN COMMON
C
       INP=5
       IOUT=8
       NV=2
C
C      RUN 2 SPHERICAL AND 2 EXPONENTIAL SIMULATIONS
C
C      FIRST, THE SPHERICAL SIMULATIONS, RANGE=1, WRITTEN TO FILE IDAT=9
C
       A=1.0
       NAM(1)=N1S
       NAM(2)=N2S
       IDAT=9
C*****
C      COMMENT TO TERMINAL
       WRITE(6,8001)
 8001 FORMAT(' STARTING SPHERICAL SIMULATIONS')
C*****
       CALL TB3D(NLI,NCO,NLV,DL,DC,DN,A,6,1,IDAT,FMT,DSEED,YMEAN,YVAR,Y,
      *DX,W)
C
C      REWIND FILE IDAT AND READ DATA INTO VECTOR VR
C
C      OUTPUT OF TB3D IS IN THE PROPER ORDER FOR INPUT TO HIST AND GAM3
C
C      LOOP THROUGH NCO*NLV*NV COLUMNS OF NLI DATA WRITTEN IN FORMAT FMT
C
       REWIND IDAT
       NNN=NCO*NLV*NV
```

```
      DO 4 I=1,NNN
      K1=NLI*(I-1)+1
      K2=NLI*I
    4 READ(IDAT,FMT) (VR(J),J=K1,K2)
C
C     RUN HISTOGRAMS AND VARIOGRAMS
C
      NDD=NLI*NCO*NLV
      DO 5 IV=1,NV
    5 CALL HIST(VR,WT,NDD,IV,0,1,1.0,1.0,0.0,-100.0,XFR,U,V,N,0)
C*****
C     COMMENT TO TERMINAL
      WRITE(6,8003)
 8003 FORMAT(' STARTING GAM3')
C*****
      CALL GAM3(10,T,NLI,NCO,NLV,4,ID,JD,KD,VR,ND,UD,VD,NP,GAM,UG,1)
C
C     NOW THE EXPONENTIAL SIMULATIONS, PARAM.=1/3, WRITTEN TO IDAT=10
C
      A=-0.3333
      NAM(1)=N1E
      NAM(2)=N2E
      IDAT=10
C*****
C     COMMENT TO TERMINAL
      WRITE(6,8002)
 8002 FORMAT(' STARTING EXPONENTIAL SIMULATIONS')
C*****
      CALL TB3D(NLI,NCO,NLV,DL,DC,DN,A,6,1,IDAT,FMT,DSEED,YMEAN,YVAR,Y,
     *DX,W)
      REWIND IDAT
      DO 14 I=1,NNN
      K1=NLI*(I-1)+1
      K2=NLI*I
   14 READ(IDAT,FMT) (VR(J),J=K1,K2)
      DO 15 IV=1,NV
   15 CALL HIST(VR,WT,NDD,IV,0,1,1.0,1.0,0.0,-100.0,XFR,U,V,N,0)
C*****
C     COMMENT TO TERMINAL
      WRITE(6,8003)
C*****
      CALL GAM3(10,T,NLI,NCO,NLV,4,ID,JD,KD,VR,ND,UD,VD,NP,GAM,UG,1)
      STOP
      END
```

Although subroutine TB3D could be used for two-dimensional simulations by setting one of the parameters (NLI,NCO,NLV) to 1 or by drawing a slice from a three-dimensional simulation file, the

distribution of the 15 lines in space is not well suited for this purpose, so a two-dimensional turning-bands routine such as that used by Mantoglou and Wilson (1981, 1982) would be greatly preferable. No two-dimensional routine has been developed for this dissertation. Details on the distinction between two- and three-dimensional turning-bands simulations are provided in Section 3.3.1.4.

Listing of Subroutine TB3D:

```
      SUBROUTINE TB3D(NLI,NCO,NLV,DL,DC,DN,A,ITERM,IWRT,IDAT,FMT,DSEED,
     *YMEAN,YVAR,Y,DX,W)
C
C
C     *****  "TURNING BANDS, 3 DIMENSIONS"  *****
C
C
C     SUBROUTINE TO GENERATE 3-D RECTANGULAR ARRAYS OF DATA DRAWN
C     FROM A 3-D STATIONARY GAUSSIAN RANDOM FUNCTION WITH EITHER
C     A SPHERICAL OR AN EXPONENTIAL 3-D ISOTROPIC SEMIVARIOGRAM MODEL
C
C     A GEOMETRIC ANISOTROPY PARALLEL TO THE GRID DIRECTIONS CAN BE
C     IMPOSED BY APPROPRIATE ADJUSTMENTS OF THE GRID SPACINGS.
C
C     UP TO 10 INDEPENDENT REALIZATIONS OF THE SAME RANDOM FUNCTION
C     CAN BE SIMULATED IN ONE CALL OF THE SUBROUTINE.
C
C     THIS SUBROUTINE USES THE TURNING-BANDS METHOD, DESCRIBED BY
C     A. G. JOURNEL AND CH. J. HUIJBREGTS (1978), MINING GEOSTATISTICS,
C     PP. 498-508 AND 534-545, AND IS AN EXTENSIVE REVISION OF THEIR
C     SUBROUTINE SIMUL, PP. 538-545.
C
C     VERSION OF AUGUST, 1985, BY G. R. LUSTER
C
C
C     *****  INPUT PARAMETERS (SEE ALSO "COMMON VARIABLES" BELOW)  *****
C
C
C     NLI,NCO,NLV  NUMBER OF LINES (ROWS), COLUMNS, AND LEVELS
C                  DEFINING THE SIMULATION DOMAIN
C
C     DL,DC,DN     SPACINGS OF THE SIMULATION GRID BETWEEN LINES,
C                  COLUMNS, AND LEVELS
C
C     A            RANGE PARAMETER:
C                  > 0.  RANGE OF THE SPHERICAL MODEL
C                  < 0.  PARAMETER "A" OF THE EXPONENTIAL MODEL
C                        (PRACTICAL RANGE IS 3A)
C                  = 0.  NOT PERMITTED -- RUN TERMINATES
```

```
C
C     ITERM         LOGICAL UNIT NUMBER FOR WRITING TERMINAL MESSAGES
C                   TO SHOW HOW FAR EXECUTION HAS PROGRESSED
C                   -- ***************** WARNING *****************
C                      MESSAGES MARKED "C*****" SHOULD BE REMOVED FOR
C                      BATCH RUNS.
C
C     IWRT          THE SIMULATED VALUES ARE WRITTEN TO UNIT IOUT
C                   ONLY IF IABS(IWRT)=1.  THE ORDER IN WHICH THEY
C                   ARE WRITTEN IS THE SAME AS THAT IN UNIT IDAT.
C                   STATISTICS OF THE 15 LINE SIMULATIONS ARE REPORTED
C                   TO UNIT IOUT FOR EACH OF THE NV REALIZATIONS ONLY
C                   IF IWRT IS POSITIVE.
C
C     IDAT          UNIT NUMBER OF THE DATA FILE (SIMULATED VALUES);
C                   THE OUTPUT IS AS FOLLOWS:  LINE NUMBERS (I=1,NLI)
C                   ARE NESTED WITHIN COLUMN NUMBERS (J=1,NCO), WHICH
C                   ARE NESTED WITHIN LEVEL NUMBERS (M=1,NLV), WHICH
C                   ARE NESTED WITHIN VARIABLE (FIELD) NUMBERS
C                   (IS=1,NV).
C
C     FMT(8)        VARIABLE FORMAT FOR WRITING OUT THE SIMULATED VALUES
C                   -- THE FORMAT IS PASSED FROM THE MAIN PROGRAM
C                      IN A VECTOR OF 8 4-BYTE CHARACTER STRINGS
C                      (TOTAL OF 32 SPACES AVAILABLE).
C                   -- THESE CHARACTERS MUST INCLUDE BLANKS FOR UNUSED
C                      SPACES, E.G., '(1H ,5F10.4)                    '.
C
C     DSEED         INTEGER INITIALIZING THE RANDOM-NUMBER GENERATOR
C                   -- TREATED AS A REAL*8 NUMBER
C                   -- MUST LIE BETWEEN 1 AND (2**31)-1
C                   -- UPDATED EACH TIME A NEW NUMBER IS GENERATED
C                   -- ***************** WARNING *****************
C                      THIS GENERATOR NEEDS AS MUCH AS 46-BIT ACCURACY
C                      AND MAY FAIL ON SOME MACHINES; REFER TO THE
C                      USER'S MANUAL.
C
C
C     *****  OUTPUT STATISTICS  *****
C
C
C     YMEAN(NV)     SAMPLE MEAN OF THE NLI*NCO*NLV SIMULATED VALUES
C                   FOR EACH REALIZATION (IS=1,NV)
C
C     YVAR(NV)      SAMPLE VARIANCE OF THE NLI*NCO*NLV SIMULATED VALUES
C                   FOR EACH REALIZATION (IS=1,NV)
C
C
C     *****  WORKING ARRAYS AND DIMENSIONS  *****
C
C
C     DIMENSION PARAMETERS
C
C     UN            = AMIN1(DL,DC,DN)
```

```
C                    BAND (POINT) SPACING FOR LINE SIMULATIONS
C
C      NX            = SQRT((NLI*DL)**2+(NCO*DC)**2+(NLV*DN)**2)/UN + 6
C                    NUMBER OF POINTS SIMULATED ON EACH LINE,
C                    CORRESPONDING TO THE DIAGONAL OF A RECTANGULAR
C                    PRISM (THE SIMULATION DOMAIN), PLUS 6
C
C      KD            SPAN OF THE LINE WEIGHT FUNCTION W; I.E.,
C                    THE NUMBER OF RANDOM NUMBERS CONTRIBUTING TO EACH
C                    SIMULATED POINT ON A LINE;
C                    IN THE CASE OF A SPHERICAL MODEL:      KD = 2*NR+1
C                    IN THE CASE OF AN EXPONENTIAL MODEL:   KD = 8*NR+1
C                    WHERE NR = MAX1( ABS(A)/(2*UN) , 38. )
C
C      WORKING ARRAYS
C
C      Y(  )         ARRAY CONTAINING:
C                    -- FIRST:  RANDOM NUMBERS FOR THE 1-D SIMULATIONS
C                    -- THEN:   THE NLI SIMULATED VALUES ON ANY SINGLE
C                       COLUMN OF THE RECTANGULAR PRISM
C                    THE DIMENSION OF Y IS:  MAX0(KD,NLI)
C
C      DX(NX*15)     WORKING ARRAY CONTAINING THE NX 1-D
C                    SIMULATED VALUES ON THE 15 LINES
C
C      UL(15)        SAMPLE MEAN AND VARIANCE OF 1-D VALUES
C      VL(15)        ON EACH OF THE 15 LINES
C
C      X(3)          COORDINATES OF A 3-D SIMULATED VALUE WITH REFERENCE
C                    TO THE FIRST TRIHEDRON, HAVING AXES PARALLEL TO THE
C                    THREE SIDES OF THE PRISM, AND ORIGIN APPROXIMATELY
C                    AT ITS CENTER
C
C      XI(3)         COORDINATES OF A 3-D SIMULATED VALUE WITH REFERENCE
C                    TO ANY OF THE FIVE TRIHEDRONS, SAME ORIGIN
C
C      S1(9)         TRANSPOSED ROTATION MATRIX "R", FROM J&H, P. 504
C
C      SX(I=1,9)     CORRESPONDS TO A UNIT DIAGONAL MATRIX
C
C      SX(I=10,45)   CORRESPONDS TO THE FOUR TRANSFORM MATRICES WHICH
C                    ALLOW PASSAGE FROM THE FIRST TRIHEDRON TO ANY OF
C                    THE FOUR OTHER TRIHEDRONS
C
C      S(3,3,4)      WORKING ARRAY USED TO DETERMINE THE VARIOUS
C                    TRANSFORM MATRICES STORED IN SX
C
C      W(KD)         WEIGHT FUNCTION USED FOR THE 1-D SIMULATIONS
C
C
C      *****  COMMON VARIABLES  *****
C
C
C      INP            INPUT UNIT
```

```
C
C     IOUT          OUTPUT UNIT (NOT THE DATA FILE, IDAT)
C
C     DUMMY         NOT USED
C
C     NV            NUMBER OF INDEPENDENT REALIZATIONS (UP TO 10)
C                   -- EACH REALIZATION MUST HAVE THE SAME PARAMETERS.
C
C     NAM(10)       8-CHARACTER VARIABLE NAMES (UP TO 10 OF THEM)
C                   CORRESPONDING TO THE NV INDEPENDENT REALIZATIONS
C
C
C     *****  INSPECTION OF INPUT PARAMETERS  *****
C
C
      DIMENSION Y(1),DX(1),YMEAN(1),YVAR(1),G(1),W(1),FMT(1)
      DIMENSION X(3),XI(3),SX(45),S1(9),S(3,3,4)
      DIMENSION UL(15),VL(15)
C
      DOUBLEPRECISION NAM(10), DSEED
      COMMON INP,IOUT,DUMMY,NV,NAM
      EQUIVALENCE (S(1,1,1),S1(1)),(SX(10),S(1,1,1))
C
C     DATA FOR ROTATION MATRIX
C
      DATA S1/0.5,-0.809017,0.309017,0.809017,0.309017,-0.5,
     *0.309017,0.5,0.809017/
C
C     WRITE OUT INPUT PARAMETERS
C
      IWRTA=IABS(IWRT)
      WRITE(IOUT,1000) NLI,NCO,NLV,NV,(NAM(IS),IS=1,NV)
      WRITE(IOUT,1001) IWRTA,FMT,IWRT,DSEED
C
C     TEST GRID FINENESS WITH RESPECT TO THE SEMIVARIOGRAM RANGE
C     (PRACTICAL RANGE FOR AN EXPONENTIAL SEMIVARIOGRAM)
C     -- WE WANT AT LEAST 3 LAGS INSIDE THE RANGE FOR THE RANGE
C        TO BE VISIBLE IN SAMPLE SEMIVARIOGRAMS IN ANY DIRECTION
C
      AA=ABS(A)
      A3=AA
      IF(A.LT.0.0) A3=3.0*A3
      IF(A.GT.0.0) WRITE(IOUT,1002) DL,DC,DN, A3
      IF(A.LT.0.0) WRITE(IOUT,1003) DL,DC,DN, A3
      IL=INT(A3/DL)
      IC=INT(A3/DC)
      IN=INT(A3/DN)
      I2=INT(A3/SQRT(DL*DL+DC*DC+DN*DN))
      IF(IL.LT.4) WRITE(IOUT,1004) IL
      IF(IC.LT.4) WRITE(IOUT,1005) IC
      IF(IN.LT.4) WRITE(IOUT,1006) IN
      IF(I2.LT.4) WRITE(IOUT,1007)
C
C
```

```
508

C      *****  CALCULATION OF FURTHER PARAMETERS  *****
C
C
C      COMPUTE ROTATION MATRIX SX OF THE 15 LINES
C      FIRST TRANSFORM = UNIT DIAGONAL MATRIX
C      FURTHER TRANSFORMS ARE SUCCESSIVE ROTATIONS
C
       DO 1 K=2,8
     1 SX(K)=0.
       SX(1)=1.
       SX(5)=1.
       SX(9)=1.
       DO 10 K=2,4
       K0=K-1
       DO 11 I=1,3
       DO 11 J=1,3
       S(I,J,K)=0.
       DO 12 I1=1,3
    12 S(I,J,K)=S(I,J,K)+S(I,I1,K0)*S(I1,J,1)
    11 CONTINUE
    10 CONTINUE
C
C      GET BAND SPACING, LINE LENGTH, TOTAL GRID POINTS
C
       UN=AMIN1(DL,DC,DN)
       NX=SQRT(NLI*NLI*DL*DL+NCO*NCO*DC*DC+NLV*NLV*DN*DN)/UN + 6
       NP=NLI*NCO*NLV
C
C      FIND CENTER COORDINATES OF PRISM AND SIMULATED LINES
C
       N1=(NLI+1)/2
       N2=(NCO+1)/2
       N3=(NLV+1)/2
       NG=NX/2
C
C      VARIANCE CORRECTION FOR SIMULATED VALUES SUMMED FROM 15 LINES
C
       SQ15=SQRT(15.)
C
C      CALCULATE SIMULATION PARAMETERS
C
C      NUGGET-EFFECT SIMULATIONS SHOULD NOT BE GENERATED BY TURNING BANDS
C
       IF(A.NE.0.) GO TO 30
       WRITE(IOUT,1008)
       STOP
C
C      PARAMETERS FOR SPHERICAL OR EXPONENTIAL SIMULATION
C
C      USED TO GET SPAN FOR CONVOLUTION:  NR
C      RANDOM NUMBERS PER LINE SPACING:  ILAG
C      SPACING ON LINE:  ALAG
C      RELATIVE SPACING:  EPS
C
```

```
   30 NR=AA/(2*UN)
      ILAG=1
      IF(NR.GE.20) GO TO 301
      NR=MAX0(1,NR)
      DO 300 ILAG=2,20
      IF(NR*ILAG.GE.20) GO TO 301
  300 CONTINUE
  301 ALAG=UN/ILAG
      EPS=ALAG/AA
      NR=NR*ILAG
      IF(A.LT.0.) GO TO 31
C
C     PARAMETERS FOR SPHERICAL SIMULATION
C
C     SPAN OF WEIGHT FUNCTION:  KD
C     WEIGHT FUNCTION:  W
C     VARIANCE CORRECTION FACTOR:  CK
C
      KD=2*NR+1
      DO 303 K=1,KD
  303 W(K)=K-NR-1
      CK=SQRT(36./NR/(NR+1)/(2*NR+1))
      GO TO 33
C
C     PARAMETERS FOR EXPONENTIAL SIMULATION
C     (ANALOGOUS TO SPHERICAL PARAMETERS)
C
   31 KD=8*NR+1
      DO 32 K=1,KD
      XT=EPS*(K-1)
   32 W(K)=(1.-XT)*EXP(-XT)
      C1=1.-EXP(-2.*EPS)
      CK=SQRT(12.*C1*C1/(C1-EPS*EXP(-2.*EPS)))
C
C     LOOP OVER NV INDEPENDENT REALIZATIONS
C
   33 DO 4 IS=1,NV
C*****
C     COMMENT TO TERMINAL
      WRITE(ITERM,3001) IS, NAM(IS)
 3001 FORMAT(' BEGINNING SIMULATION OF VARIABLE ',I2,2X,'(',A8,')')
C*****
C
C
C     *****  LINE SIMULATIONS  *****
C
C
C     INITIALIZE INDEX OF VECTOR DX OF LINE VALUES; LOOP OVER 15 LINES
C
      IP=0
      DO 34 ID=1,15
C*****
C     COMMENT TO TERMINAL
      WRITE(ITERM,3002) ID
```

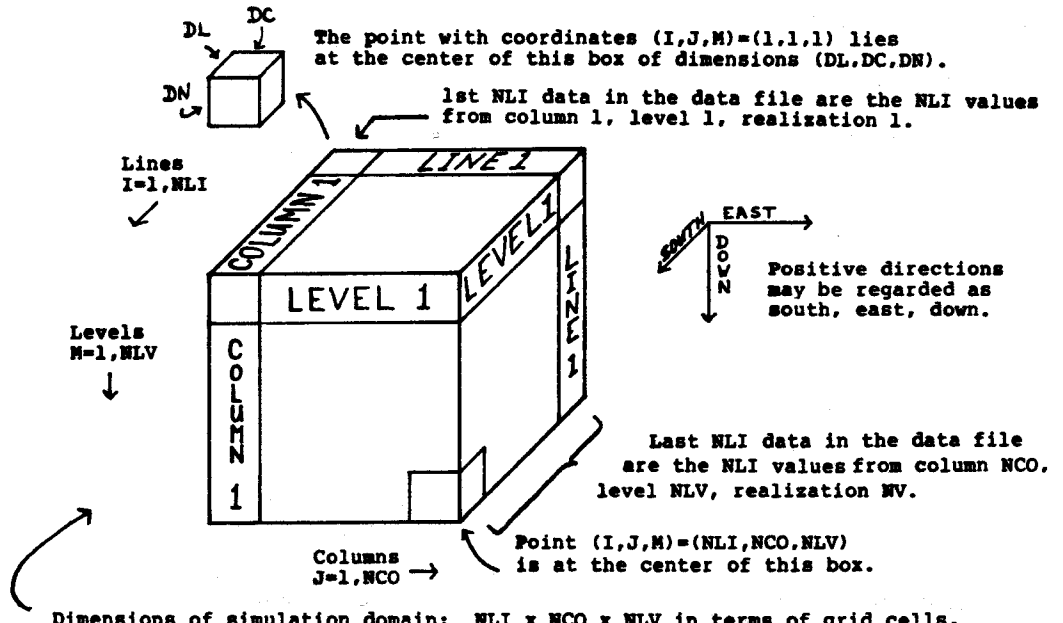
```
510

 3002 FORMAT(' STARTING LINE ',I2)
C*****
C
C     INITIALIZE MEAN AND VARIANCE ACCUMULATORS FOR LINE NO. ID
C
      UL(ID)=0.0
      VL(ID)=0.0
C
C     GET KD UNIFORM RANDOM NUMBERS FOR THE FIRST SIMULATED VALUE;
C     NUMBERS ARE ADJUSTED FOR THEIR EXPECTED VALUE OF 0.5;
C     SEE *****WARNING***** IN DESCRIPTION OF INPUT PARAMETERS (ABOVE)
C
      DO 340 K=1,KD
      DSEED=DMOD(16807.D0*DSEED,2147483647.D0)
  340 Y(K)=DSEED/2147483648.D0 - 0.5D0
C
C     POINT SIMULATION BY CONVOLUTION OF KD RANDOM NUMBERS
C     UNCORRECTED SIMULATED LINE VALUE:  AD
C
      DO 35 J=1,NX
      IP=IP+1
      AD=0.
      KK=MOD((J-1)*ILAG,KD)
      DO 350 K=1,KD
      IF(KK.EQ.KD) KK=0
      KK=KK+1
  350 AD=AD+W(K)*Y(KK)
C
C     GET ILAG RANDOM NUMBERS FOR THE NEXT SIMULATED VALUE
C     SEE *****WARNING***** IN DESCRIPTION OF INPUT PARAMETERS (ABOVE)
C
      DO 351 K=1,ILAG
      IF(KK.EQ.KD) KK=0
      KK=KK+1
      DSEED=DMOD(16807.D0*DSEED,2147483647.D0)
  351 Y(KK)=DSEED/2147483648.D0 - 0.5D0
C
C     VARIANCE CORRECTION; SIMULATED LINE VALUE DX(IP)
C
      DX(IP)=AD*CK
C
C     ACCUMULATE LINE STATISTICS
C
      UL(ID)=UL(ID)+DX(IP)
      IF(J.EQ.1) GO TO 35
      VL(ID)=VL(ID)+(UL(ID)-J*DX(IP))**2/(J*(J-1))
   35 CONTINUE
C
C     STATISTICS PER LINE
C
      UL(ID)=UL(ID)/NX
      VL(ID)=VL(ID)/NX
   34 CONTINUE
C
```

```
C     WRITE LINE STATISTICS IF IWRT IS POSITIVE
C
      IF(IWRT.LE.0) GO TO 39
      WRITE(IOUT,1009) IS, NAM(IS)
      DO 38 ID=1,15
   38 WRITE(IOUT,1010) ID, UL(ID), VL(ID)
C
C     SIMULATED DATA ARE WRITTEN ON UNIT IOUT IF IABS(IWRT)=1
C
   39 IF(IWRTA.EQ.1) WRITE(IOUT,1014)
C
C
C     *****  THREE-DIMENSIONAL SIMULATION  *****
C
C
C     EACH NEW 3-D VALUE IS SIMULATED BY ADDING
C     CONTRIBUTIONS FROM 15 LINES
C
C     INITIALIZE ACCUMULATORS OF 3-D STATISTICS
C*****
C     COMMENT TO TERMINAL
      WRITE(ITERM,3003)
 3003 FORMAT(' STARTING 3-D SIMULATION'/)
C*****
C
      TOT=0.0
      YMEAN(IS)=0.
      YVAR(IS)=0.
C
C     NESTED LOOPS OVER LEVELS M, COLUMNS J, LINES I
C
C     POINT COORDS. X, WITH ORIGIN AT PRISM CENTER:
C     COORDS. FOR LEVEL (X(3)), COLUMN (X(2)), LINE (X(1))
C
      DO 4 M=1,NLV
      X(3)=-0.5+(M-N3)*DN/UN
      DO 4 J=1,NCO
      X(2)=-0.5+(J-N2)*DC/UN
      DO 40 I=1,NLI
      X(1)=-0.5+(I-N1)*DL/UN
C
C     INITIALIZE SIMULATED VALUE Y TO ZERO
C
      Y(I)=0.
      INO=0
C
C     LOOP OVER 5 TRIHEDRA (IR), 3 AXES (JD)
C
      DO 41 IR=1,5
      KO=(IR-1)*3
      DO 41 JD=1,3
      XI(JD)=0.
C
C     COMPUTE THE COORDINATES WITH REFERENCE TO TRIHEDRON IR
```

```
C
      DO 42 IL=1,3
      INO=INO+1
   42 XI(JD)=XI(JD)+X(IL)*SX(INO)
      ARG=XI(JD)+0.5
      LK=ARG
      IF(ARG.LT.0.) LK=LK-1
      LK=LK+NG
      KO=KO+1
C
C     COMPUTE SUM OF CONTRIBUTIONS FROM THE TURNING BANDS
C
      I2=LK+NX*(KO-1)
      Y(I)=Y(I)+DX(I2)
   41 CONTINUE
C
C     ACCUMULATE 3-D STATISTICS
C
      Y(I)=Y(I)/SQ15
      TOT=TOT+1.0
      YMEAN(IS)=YMEAN(IS)+Y(I)
      IF(TOT.LT.2.0) GO TO 40
      YVAR(IS)=YVAR(IS)+(YMEAN(IS)-TOT*Y(I))**2/(TOT*(TOT-1.0))
   40 CONTINUE
C
C     WRITE COLUMN J ON FILE IDAT AND ON IOUT IF IABS(IWRT)=1
C
      WRITE(IDAT,FMT) (Y(I),I=1,NLI)
      IF(IWRTA.EQ.1) WRITE(IOUT,FMT) (Y(I),I=1,NLI)
    4 CONTINUE
C
C
C      *****  STATISTICS OF 3-D SIMULATIONS  *****
C
C
C     WRITE MEANS AND VARIANCES
C
      IF(IWRT.GT.0.OR.IWRTA.EQ.1) WRITE(IOUT,1015)
      WRITE(IOUT,1011)
      DO 5 IS=1,NV
      YMEAN(IS)=YMEAN(IS)/NP
      YVAR(IS)=YVAR(IS)/NP
    5 WRITE(IOUT,1012) IS,NAM(IS),YMEAN(IS),YVAR(IS)
C
C     WRITE FINAL VALUE OF DSEED
C
      WRITE(IOUT,1013) DSEED
C
C
C     *****  OUTPUT FORMATS  *****
C
C
```

```
 1000 FORMAT('1SUBROUTINE TB3D -- THREE-DIMENSIONAL TURNING-BANDS ',
     1'SIMULATION'//'0THE INPUT PARAMETERS ARE:'//' GRID ',
     2'DIMENSIONS OF SIMULATION DOMAIN:'/'      NUMBER OF LINES = ',
     3'NLI = ',I6/'      NUMBER OF COLUMNS = NCO = ',I4/
     4'      NUMBER OF LEVELS = NLV = ',I5//
     5' NUMBER OF INDEPENDENT REALIZATIONS (ALL WITH THE SAME ',
     6'PARAMETERS) = NV = ',I2//' NAMES OF THE REALIZATIONS:'/' ',
     75(A8,3X)/' ',5(A8,3X))
 1001 FORMAT('0LISTING OF SIMULATED DATA BELOW?',
     1' ("1" MEANS YES) = IABS(IWRT) =',I2//' FORMAT FOR ',
     2'DATA OUTPUT = FMT = ',8A4//
     3' REPORT OF LINE STATISTICS BELOW?',
     4' (POSITIVE VALUE MEANS YES) = IWRT = ',I2//
     5' INITIAL SEED FOR RANDOM-NUMBER',
     6' GENERATOR = DSEED = ',D20.12)
 1002 FORMAT('0GRID SPACINGS:'/
     1'      BETWEEN LINES = DL = ',F12.4/
     2'      BETWEEN COLUMNS = DC = ',F10.4/
     3'      BETWEEN LEVELS = DN = ',F11.4//
     4' VARIOGRAM RANGE = AA = ',F10.4//
     5' THE MODEL RANDOM FUNCTION HAS A **SPHERICAL** VARIOGRAM.')
 1003 FORMAT('0GRID SPACINGS:'/
     1'      BETWEEN LINES = DL = ',F12.4/
     2'      BETWEEN COLUMNS = DC = ',F10.4/
     3'      BETWEEN LEVELS = DN = ',F11.4//
     4' PRACTICAL VARIOGRAM RANGE = 3*AA = ',F10.4//
     5' THE MODEL RANDOM FUNCTION HAS AN **EXPONENTIAL** VARIOGRAM.')
 1004 FORMAT('0NOTICE THAT THE RANGE IS ONLY',I2,' LAGS LONG ',
     1'BETWEEN LINES.')
 1005 FORMAT('0NOTICE THAT THE RANGE IS ONLY',I2,' LAGS LONG ',
     1'BETWEEN COLUMNS.')
 1006 FORMAT('0NOTICE THAT THE RANGE IS ONLY',I2,' LAGS LONG ',
     1'BETWEEN LEVELS.')
 1007 FORMAT('0NOTICE THAT THE RANGE IS SHORT (IN TERMS OF GRID ',
     1'SPACINGS) IN DIAGONAL DIRECTIONS.')
 1008 FORMAT('0ZERO RANGE PARAMETER NOT ACCEPTABLE:  STOP')
 1009 FORMAT('1LINE STATISTICS FOR SIMULATION NUMBER ',I2,' (',A8,')'/
     1'0EXPECTED MEAN = 0.0, VARIANCE = 1.0 (FOR AN INFINITE LINE)'/
     2'0OBSERVED VALUES:  LINE   MEAN   VARIANCE'/)
 1010 FORMAT(' ',18X,I3,2X,F6.2,3X,F6.2)
 1011 FORMAT('0'/'0STATISTICS FOR SIMULATED REALIZATIONS:'/
     1'0THEORETICAL MEAN = 0.0, SILL = 1.0'/
     2'0 REALIZATION       MEAN   VARIANCE'/)
 1012 FORMAT(' ',I2,1X,'(',A8,')',3X,F7.4,3X,F8.4)
 1013 FORMAT('0FINAL VALUE OF RANDOM-NUMBER SEED = DSEED = ',D20.12)
 1014 FORMAT('1LISTING OF SIMULATED DATA IN FORMAT "FMT"'//)
 1015 FORMAT('1')
C
      RETURN
      END
```



Data file for NLI*NCO*NLV*NV 3-D simulated values:

(I,J,M) <-- grid coordinates of data stored in file

	First Level: M=1	(1,1,1),(2,1,1),(3,1,1), ... ,(NLI,1,1) (1,2,1),(2,2,1),(3,2,1), ... ,(NLI,2,1) (1,3,1),(2,3,1),(3,3,1), ... ,(NLI,3,1) . . . . (1,NCO,1), . . . ,(NLI,NCO,1)	<-- 1st NLI values for 1st column within 1st level within 1st realization
First Realization: IS = 1	Second Level: M=2	(1,1,2),(2,1,2),(3,1,2), ... ,(NLI,1,2) (1,2,2),(2,2,2),(3,2,2), ... ,(NLI,2,2) (1,3,2),(2,3,2),(3,3,2), ... ,(NLI,3,2) . . . . (1,NCO,2), . . . ,(NLI,NCO,2)	
		. . . .	
	Bottom Level: M=NLV	(1,1,NLV),(2,1,NLV), ... ,(NLI,1,NLV) (1,2,NLV),(2,2,NLV), ... ,(NLI,2,NLV) (1,3,NLV),(2,3,NLV), ... ,(NLI,3,NLV) . . . . . (1,NCO,NLV), . . . ,(NLI,NCO,NLV)	
		. . . .	
Last Realization: IS = NV		(1,1,1), ... ,(NLI,NCO,NLV)	Last NLI values for NCOth column within NLVth level within NVth realization <--

Figure 116: Arrangement of simulated data on the three-dimensional grid and in the data file IDAT.

```
SUBROUTINE TB3D -- THREE-DIMENSIONAL TURNING-BANDS SIMULATION

THE INPUT PARAMETERS ARE:

GRID DIMENSIONS OF SIMULATION DOMAIN:
     NUMBER OF LINES = NLI =      20
     NUMBER OF COLUMNS = NCO =    10
     NUMBER OF LEVELS = NLV =      5

NUMBER OF INDEPENDENT REALIZATIONS (ALL WITH THE SAME PARAMETERS) = NV =  2

NAMES OF THE REALIZATIONS:
SPH. # 1    SPH. # 2

LISTING OF SIMULATED DATA BELOW? ("1" MEANS YES) = IABS(IWRT) = 1

FORMAT FOR DATA OUTPUT = FMT = (1H ,5F10.4)

REPORT OF LINE STATISTICS BELOW? (POSITIVE VALUE MEANS YES) = IWRT =  1

INITIAL SEED FOR RANDOM-NUMBER GENERATOR = DSEED =   0.5208600000000D 05

GRID SPACINGS:
     BETWEEN LINES = DL =        0.2000
     BETWEEN COLUMNS = DC =      0.5000
     BETWEEN LEVELS = DN =       0.2000

VARIOGRAM RANGE = AA =      1.0000

THE MODEL RANDOM FUNCTION HAS A **SPHERICAL** VARIOGRAM.

NOTICE THAT THE RANGE IS ONLY 2 LAGS LONG BETWEEN COLUMNS.

NOTICE THAT THE RANGE IS SHORT (IN TERMS OF GRID SPACINGS) IN DIAGONAL DIRECTIONS.



         STATISTICS FOR SIMULATED REALIZATIONS:

         THEORETICAL MEAN = 0.0, SILL = 1.0

           REALIZATION        MEAN    VARIANCE

         1 (SPH. # 1)     -0.2772      0.9715
         2 (SPH. # 2)     -0.0614      0.9168

         FINAL VALUE OF RANDOM-NUMBER SEED = DSEED =   0.110940126200D 10
```

Figure 117: Output from the first call of Subroutine TB3D in the demonstration run.

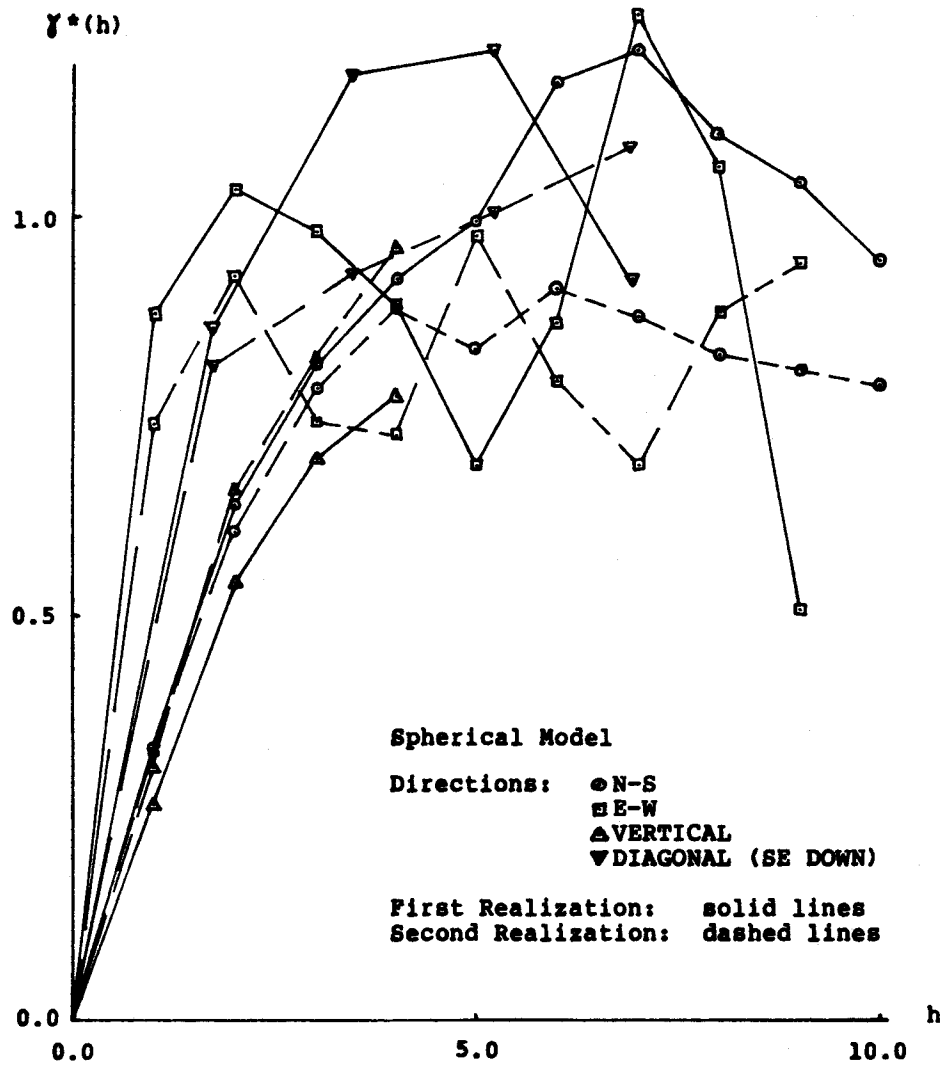


Figure 118: Sample semivariograms from the first call of Subroutine TB3D in the demonstration run. Notice the geometric anisotropy, the poor model reproduction at long lags (owing to the small size of the domain in comparison to the range), and the poor visibility of short-scale structure in the east-west (between columns) direction (owing to coarse discretization of the domain in comparison to the range).

```
SUBROUTINE TB3D -- THREE-DIMENSIONAL TURNING-BANDS SIMULATION

THE INPUT PARAMETERS ARE:

GRID DIMENSIONS OF SIMULATION DOMAIN:
     NUMBER OF LINES = NLI =      20
     NUMBER OF COLUMNS = NCO =    10
     NUMBER OF LEVELS = NLV =      5

NUMBER OF INDEPENDENT REALIZATIONS (ALL WITH THE SAME PARAMETERS) = NV =  2

NAMES OF THE REALIZATIONS:
EXP. # 1    EXP. # 2

LISTING OF SIMULATED DATA BELOW? ("1" MEANS YES) = IABS(IWRT) = 1

FORMAT FOR DATA OUTPUT = FMT = (1H ,5F10.4)

REPORT OF LINE STATISTICS BELOW? (POSITIVE VALUE MEANS YES) = IWRT =  1

INITIAL SEED FOR RANDOM-NUMBER GENERATOR = DSEED =   0.110940126200D 10

GRID SPACINGS:
     BETWEEN LINES = DL =        0.2000
     BETWEEN COLUMNS = DC =      0.5000
     BETWEEN LEVELS = DN =       0.2000

PRACTICAL VARIOGRAM RANGE = 3*AA =      0.9999

THE MODEL RANDOM FUNCTION HAS AN **EXPONENTIAL** VARIOGRAM.

NOTICE THAT THE RANGE IS ONLY 1 LAGS LONG BETWEEN COLUMNS.

NOTICE THAT THE RANGE IS SHORT (IN TERMS OF GRID SPACINGS) IN DIAGONAL DIRECTIONS.


          STATISTICS FOR SIMULATED REALIZATIONS:

          THEORETICAL MEAN = 0.0, SILL = 1.0

            REALIZATION       MEAN    VARIANCE

          1 (EXP. # 1)    -0.4992      0.9642
          2 (EXP. # 2)    -0.0467      0.8564

          FINAL VALUE OF RANDOM-NUMBER SEED = DSEED =   0.142124286300D 10
```

Figure 119: Output from the second call of Subroutine TB3D in the demonstration run.

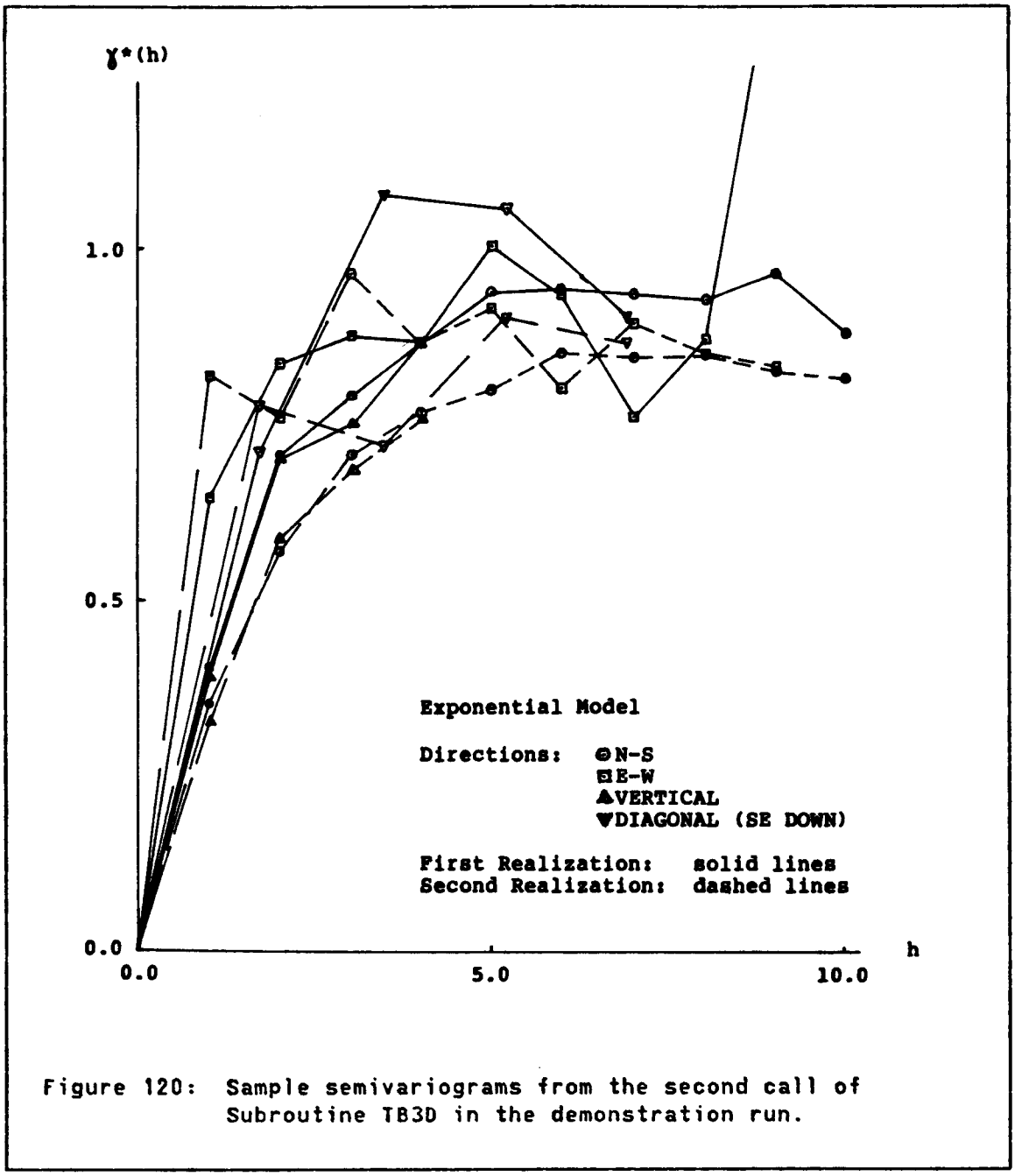


Figure 120: Sample semivariograms from the second call of Subroutine TB3D in the demonstration run.

REFERENCES CITED

Abramowitz, M., and I. A. Stegun, eds. (1972), Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables, 9th (revised) Dover printing, Dover, New York, 1046 pp.

Aitchison, J. (1981), "A New Approach to Null Correlations of Proportions", Mathematical Geology, v. 13, no. 2, pp. 175-189.

Alfaro, M. (1979), Robustesse des Simulations de Fonctions Aleatoires, doctoral thesis, L'Ecole Nationale Superieure des Mines de Paris, 161 pp.

Alfaro, M. (1980), "The Random Coin Method: Solution of the Problem of the Simulation of a Random Function in the Plane", Mathematical Geology, v. 12, no. 1, pp. 25-32.

Alfaro, M. (1984), "Statistical Inference of the Semi-Variogram and the Quadratic Model", in Verly et al. (1984), pp. 45-53.

Alfaro, M., and C. J. Huijbregts (1974), "Simulation of a Subhorizontal Sedimentary Deposit", in Proceedings of the 12th International APCOM Symposium, Golden, Colorado, April 1974, pp. F65-F77.

Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, Wiley, New York, 374 pp.

Anderson, W. I. (1984), Guidebook for the 48th Annual Tri-State Geological Field Conference: General Geology of North-Central Iowa, Department of Earth Science, University of Northern Iowa, Cedar Falls, 150 pp.

ASTM (1984), "Standard Specification for Portland Cement" (C 150-84), in 1984 Annual Book of ASTM Standards, Vol. 04.01, American Society for Testing and Materials, Philadelphia, pp. 155-161.

Baker, R. (1984), "Modeling Soil Variability as a Random Field", Mathematical Geology, v. 16, no. 5, pp. 435-448.

Barbaro, R. W., and R. V. Ramani (1983), "A Generalized Multiperiod MIP Model For Production Scheduling and Processing Facilities Selection and Location", presented at AIME Annual Meeting, Atlanta, Preprint No. 83-123, Society of Mining Engineers of AIME, Littleton, Colorado, 12 pp.

Bellman, R. (1961), Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 255 pp.

Bogue, R. H. (1955), The Chemistry of Portland Cement, Reinhold, New York, 793 pp.

Borgman, L. E. (1982), "Techniques for Computer Simulation of Ocean Waves", in A. R. Osborne and P. Malanotte Rizzoli, eds., Topics in Ocean Physics, North-Holland, Amsterdam, International School of Physics "Enrico Fermi", v. 80, pp. 387-417.

Borgman, L. E., and R. B. Frahme (1976), "A Case Study: Multivariate Properties of Bentonite in Northeastern Wyoming", in Guarascio et al. (1976), pp. 381-390.

Borgman, L. E., S. M. Taheri, and R. L. Hagan (1984), "Three-Dimensional, Frequency-Domain Simulations of Geological Variables", in Verly et al. (1984), pp. 517-541.

Box, G. E. P., and D. R. Cox (1964), "An Analysis of Transformations", Journal of the Royal Statistical Society, v. B26, pp. 211-252.

Box, G. E. P., and G. M. Jenkins (1976), Time Series Analysis: Forecasting and Control, revised ed., Holden-Day, San Francisco, 575 pp.

Bras, R. L., and I. Rodriguez-Iturbe (1985), Random Functions and Hydrology, Addison-Wesley, Reading, Massachusetts, 559 pp.

Bratley, P., B. L. Fox, and L. E. Schrage (1983), A Guide to Simulation, Springer-Verlag, New York, 383 pp.

Brooker, P. I. (1984), "Two Dimensional Simulation by Turning Bands", presented at the SME-AIME Annual Meeting, Los Angeles, Preprint No. 84-78, Society of Mining Engineers of AIME, Littleton, Colorado, 5 pp.

Brooker, P. I. (1985), "Two-Dimensional Simulations by Turning Bands", Mathematical Geology, v. 17, no. 1, pp. 81-90.

Bryan, R. C., and F. Roghani (1982), "Application of Conventional and Advanced Methods to Uranium Ore Reserve Estimation and the Development of a Method to Adjust for Disequilibrium Problems", in T. B. Johnson and R. J. Barnes, eds. (1982), 17th Application of Computers and Operations Research in the Mineral Industry, Society of Mining Engineers of AIME, New York, pp. 109-120.

Butler, J. C. (1981), "Effect of Various Transformations on the Analysis of Percentage Data", Mathematical Geology, v. 13, no. 1, pp. 53-68.

Bye, G. C. (1983), Portland Cement: Composition, Production and Properties, Pergamon, Oxford, 149 pp.

Chatfield, C. (1980), The Analysis of Time Series: An Introduction, 2nd ed., Chapman and Hall, London, 268 pp.

Chiles, J. P. (1984), "Simulation of a Nickel Deposit: Problems Encountered and Practical Solutions", in Verly et al. (1984), pp. 1015-1030.

Colijn, H. (1980), "Storage, Blending, and Reclaiming Systems", presented at AIME Annual Meeting, Las Vegas, Preprint No. 80-1, Society of Mining Engineers of AIME, Littleton, Colorado, 16 pp.

Colijn, H. (1983), Weighing and Proportioning of Bulk Solids, 2nd ed., Trans Tech, Clausthal, Germany, 398 pp.

Cooley, J. W., and J. W. Tukey (1965), "An Algorithm for the Machine Calculation of Complex Fourier Series", Mathematics of Computation, v. 19, no. 2, pp. 297-301.

Cooley, W. W., and P. R. Lohnes (1971), Multivariate Data Analysis, Wiley, New York, 364 pp.

Dagbert, M. (1981), "The Simulation of Space-Dependent Data in Geology", in R. G. Craig and M. L. Labovitz, eds. (1981), Future Trends in Geomathematics, Pion, London, 318 pp.

Dagbert, M., M. David, D. Crozel, A. Desbarats (1984), "Computing Variograms in Folded Strata-Controlled Deposits", in Verly et al. (1984), pp. 71-89.

David, M. (1977), Geostatistical Ore Reserve Estimation, Elsevier, Amsterdam, 364 pp.

Davis, B. M., and K. A. Greenes (1983), "Estimation Using Spatially Distributed Multivariate Data: An Example With Coal Quality", Mathematical Geology, v. 15, no. 2, pp. 287-300.

Davis, B. M., R. L. Hagan, and L. E. Borgman (1981), "A Program for the Finite Fourier Transform Simulation of Realizations From a One-Dimensional Random Function With Known Covariance", Computers & Geosciences, v. 7, no. 2, pp. 199-206.

Davis, J. C. (1973), Statistics and Data Analysis in Geology, Wiley, New York, 550 pp.

Davis, M. W. (1985a), "Production of Conditional Simulations Via the LU Decomposition of the Covariance Matrix", to be published in Mathematical Geology.

Davis, M. W. (1985b), "Simulation of a Spatially Correlated Bivariate Distribution", to be submitted to Mathematical Geology.

de Oliveira Leite, S. (1983), "Tests for Multinormality", unpublished note, Department of Applied Earth Sciences, Stanford University, 71 pp.

Delhomme, J. P. (1979), "Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach", Water Resources Research, v. 15, no. 2, pp. 269-280.

Deraisme, J. (1977), "Modelisation de Gisements et Choix d'une Methode d'Exploitation", Revue de l'Industrie Minerale, v. 59, pp. 483-489.

Deraisme, J., and C. de Fouquet (1984), "Recent and Future Developments of 'Down Stream' Geostatistics", in Verly et al. (1984), pp. 979-999.

Deraisme, J., and R. Dumay (1981), "Etude de la Planification d'une Exploitation Miniere par Simulation sur Modele de Gisements", Internal Report N-726, Centre de Geostatistique, Fontainebleau, France.

Deraisme, J., and J.-P. E. Marbeau (1983), "Geostatistics and Productivity in the Mining Industry, Part II: From Development to Production", presented at the SME-AIME Annual Meeting, Atlanta, Preprint No. 83-79, Society of Mining Engineers of AIME, Littleton, Colorado, 13 pp.

Diggle, P. J. (1983), Statistical Analysis of Spatial Point Patterns, Academic Press, London, 148 pp.

Dixon, W. J., and Brown, M. B., eds. (1979), BMDP-79: Biomedical Computer Programs, P-Series, University of California Press, Los Angeles, 880 pp.

Dowd, P. A. (1978), Advances in Geostatistics: Numerical Methods and Their Application, doctoral thesis, University of Leeds, 380 pp.

Dowd, P. A. (1984), "Conditional Simulation of Inter Related Beds in an Oil Deposit", in Verly et al. (1984), pp. 1031-1043.

Dubrule, O., and C. Kostov (1985), "An Interpolation Method Taking Into Account Inequality Constraints: I. Methodology", to be published in Mathematical Geology.

Duda, W. H. (1977), Cement-Data-Book, 2nd ed., Bauverlag, Wiesbaden, Germany, 539 pp.

Everitt, B. (1980), Cluster Analysis, 2nd ed., Halsted, New York, 136 pp.

Friedman, J. H., W. Stuetzle, and A. Schroeder (1981), "Projection Pursuit Density Estimation", Project Orion, Report 002, Department of Statistics, Stanford University, 37 pp.

Garrett, H. M. (1976), "The Potential Promise -- Prospects and Pitfalls in Energy Conservation by the U. S. Cement Industry", presented at the 1976 Cement Chemist's Seminar, Portland Cement Association, Skokie, Illinois, 37 pp.

Gershon, M. (1982), "A Linear Programming Approach to Mine Scheduling Optimization", in T. B. Johnson and R. J. Barnes, eds. (1982), 17th Application of Computers and Operations Research in the Mineral Industry, Society of Mining Engineers of AIME, New York, pp. 483-493.

Ghosh, S. N., ed. (1983), Advances in Cement Technology: Critical Reviews and Case Studies on Manufacturing, Quality Control, Optimization and Use, Pergamon, Oxford, 804 pp.

Gill, P. E., and W. Murray (1972), "Two Methods for the Solution of Linearly Constrained and Unconstrained Optimization Problems", National Physical Laboratory (England), Report DNAC-25.

Gnanadesikan, R. (1977), Methods for Statistical Data Analysis of Multivariate Observations, Wiley, New York, 311 pp.

Gordon, A. D. (1981), Classification, Chapman and Hall, London, 193 pp.

Griffiths, J. C. (1967), Scientific Method in Analysis of Sediments, McGraw-Hill, New York, 508 pp.

Guarascio, M., M. David, and C. Huijbregts, eds. (1976), Advanced Geostatistics in the Mining Industry, Proceedings of the NATO Advanced Study Institute, Rome, October 13-25, 1975, D. Reidel, Dordrecht, Holland, 461 pp.

Gy, P. M. (1981), "A New Theory of Bed-Blending Derived From the Theory of Sampling -- Development and Full-Scale Experimental Check", International Journal of Mineral Processing, v. 8, pp. 201-238.

Gy, P. M. (1982), Sampling of Particulate Materials: Theory and Practice, 2nd ed., Elsevier, Amsterdam, 431 pp.

Hammersley, J. M., and J. A. Nelder (1955), "Sampling From an Isotropic Gaussian Process", Proceedings of the Cambridge Philosophical Society, v. 51, pp. 652-662.

Hand, D. J. (1981), Discrimination and Classification, Wiley, Chichester, U. K., 218 pp.

Harbaugh, J. W., and G. Bonham-Carter (1970), Computer Simulation in Geology, Wiley, New York, 575 pp.

Helwick, S. J., and G. R. Luster (1984), "Fluid-Flow Modeling Using a Conditional Simulation of Porosity and Permeability", in Verly et al. (1984), pp. 635-650.

Hillier, F. S., and G. J. Lieberman (1980), Introduction to Operations Research, 3rd ed., Holden-Day, San Francisco, 829 pp.

IMSL (1982), IMSL Library Reference Manual, 9th ed., IMSL, Houston.

Isaaks, E. H. (1984a), "Indicator Simulation: Application to the Simulation of a High Grade Uranium Mineralization", in Verly et al. (1984), pp. 1057-1069.

Isaaks, E. H. (1984b), "Conditioning Indicator Simulations to Indicator Data", unpublished note, 5 pp.

Joreskog, K. G., J. E. Klovan, and R. A. Reyment (1976), Geological Factor Analysis, Elsevier, Amsterdam, 178 pp.

Journel, A. G. (1974a), Simulations Conditionnelles: Theorie et Pratique, doctoral thesis, University of Nancy.

Journel, A. G. (1974b), "Geostatistics for Conditional Simulation of Ore Bodies", Economic Geology, v. 69, pp. 673-687.

Journel, A. G. (1980), "Geostatistical Simulation: Methods For Exploration and Mine Planning", in P. Mousset-Jones, ed., Geostatistics, McGraw-Hill, New York, pp. 93-108.

Journel, A. G. (1983), "Non-Parametric Estimation of Spatial Distributions", Mathematical Geology, v. 15, no. 3, pp. 445-468.

Journel, A. G. (1984a), "Geostatistics: Simple Tools Applied to Difficult Problems", in H. A. David and H. T. David, eds., Statistics: An Appraisal, Proceedings, 50th Anniversary Conference, Iowa State University Statistical Laboratory, Iowa State University Press, Ames, pp. 237-256.

Journel, A. G. (1984b) "mAD and Conditional Quantile Estimators", in Verly et al. (1984), pp. 261-270.

Journel, A. G. (1984c), "The Place of Non-Parametric Geostatistics", in Verly et al. (1984), pp. 307-335.

Journel, A. G. (1984d), "Conditioning With Indicator Data", unpublished note, Department of Applied Earth Sciences, Stanford University, 5 pp.

Journel, A. G. (1984e), "Expert Information and Constrained Interpolation: The Soft Kriging Approach", to be published in Mathematical Geology.

Journel, A. G. (1985), "The Deterministic Side of Geostatistics", Mathematical Geology, v. 17, no. 1, pp. 1-15.

Journel, A. G., and C. J. Huijbregts (1978), Mining Geostatistics, second (revised) printing, Academic Press, London, 600 pp.

Journel, A. G., and E. H. Isaaks (1985), "Conditional Indicator Simulation: Application to a Saskatchewan Uranium Deposit", Mathematical Geology, v. 17, no. 1, pp. 1-15.

Kennedy, W. J., Jr., and J. E. Gentle (1980), Statistical Computing, Marcel Dekker, New York, 591 pp.

Koch, D. L. (1970), Stratigraphy of the Upper Devonian Shell Rock Formation of North-Central Iowa, Report of Investigations 10, Iowa Geological Survey, Iowa City, 123 pp.

Kostov, C., and O. Dubrule (1984), "An Interpolation Method Taking Into Account Inequality Constraints: II. A Practical Approach", to be published in Mathematical Geology.

Kulkarni, R. B. (1984), "Bayesian Kriging in Geotechnical Problems", in Verly et al. (1984), pp. 775-786.

Laney, John W. III (1977), "Process Control Philosophies Used To Prepare a Uniform Kiln Feed At Citadel's Roanoke Plant", 1977 IEEE Cement Industry Technical Conference, 19 pp.

Law, A. M., and W. D. Kelton (1982), Simulation Modeling and Analysis, McGraw-Hill, New York, 400 pp.

Lea, F. M. (1971), The Chemistry of Cement and Concrete, Chemical Publishing Company, New York, 727 pp.

LeMaitre, R. W. (1982), Numerical Petrology: Statistical Interpretation of Geochemical Data, Elsevier, Amsterdam, 281 pp.

Lin, C., and J. W. Harbaugh (1984), Graphic Display of Two- and Three-Dimensional Markov Computer Models in Geology, Van Nostrand Reinhold, New York, 180 pp.

Lonergan, J. E. (1984), "The Application of Linear Goal Programming to Coal-Fired Power Plant Blending Problems", presented at AIME Annual Meeting, Los Angeles, Preprint No. 84-32, Society of Mining Engineers of AIME, Littleton, Colorado, 8 pp.

Mandelbrot, B. B. (1977), Fractals: Form, Chance, and Dimension, W. H. Freeman, San Francisco, 365 pp.

Mandelbrot, B. B. (1982), The Fractal Geometry of Nature, W. H. Freeman, San Francisco, 460 pp.

Mantoglou, A., and J. L. Wilson (1981), Simulation of Random Fields With The Turning Bands Method, Report No. 264, Ralph M. Parsons Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 199 pp.

Mantoglou, A., and J. L. Wilson (1982), "The Turning Bands Method for Simulation of Random Fields Using Line Generation by a Spectral Method", Water Resources Research, v. 18, no. 5, pp. 1379-1394.

Marechal, A. (1976), "The Practice of Transfer Functions: Numerical Methods and Their Application", in Guarascio et al. (1976), pp. 253-276.

Marriott, F. H. C. (1974), The Interpretation of Multiple Observations, Academic Press, London, 117 pp.

Matern, B. (1960), Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations, Meddelanden Fran Statens Skogsforskningsinstitut, Stockholm, v. 49, no. 5., 144 pp.

Matheron, G. (1965), Les Variables Regionalisees et Leur Estimation: Une Application de la Theorie des fonctions aleatoires aux Sciences de la Nature, Masson, Paris, 305 pp.

Matheron, G. (1973), "The Intrinsic Random Functions and Their Applications", Advances in Applied Probability, v. 5, pp. 439-468.

McDonald, J. A. (1982), "Interactive Graphics for Data Analysis", Project Orion, Report 011, Department of Statistics, Stanford University, 60 pp.

Mejia, J. M., and I. Rodriguez-Iturbe (1974), "On the Synthesis of Random Field Sampling From the Spectrum: An Application to the Generation of Hydrologic Spatial Processes", Water Resources Research, v. 10, no. 4, pp. 705-711.

Miller, B. L. (1934), "Practical Value of Economic Geology in the Manufacture of Cement", Pit & Quarry, April, 1934, pp. 29-40.

Miller, S. M., and L. E. Borgman, "Spectral-Type Simulation of Spatially Correlated Fracture Set Properties", Mathematical Geology, v. 17, no. 1, pp. 41-52.

Mosteller, F., and J. W. Tukey (1977), Data Analysis and Regression: A Second Course in Statistics, Addison-Wesley, Reading, Massachusetts, 588 pp.

Niederjohn, J. A. (1969), "Matrix Algebra and the Cement Mix Design Problem", Quarterly of the Colorado School of Mines, v. 64, no. 3 (7th APCOM), pp. 143-157.

Omre, H. (1984), "The Variogram and Its Estimation", in Verly et al. (1984), pp. 107-125.

Omre, H., and L. Holden (1984), "Surface Modelling Merging Subjective and Objective Information", presented at Eurocarto 3 symposium, Graz, Austria, October 1984, 16 pp.

Parnaby, J., P. G. Battye, and G. S. Waite (1973), "Optimal Design of Homogenising Systems Incorporating Layered Stockpiles and Fluidised Silos for the Control of Raw Materials Quality", Transactions, Institution of Chemical Engineers, v. 51, pp. 323-330.

Parzen, E. (1962), Stochastic Processes, Holden-Day, San Francisco, 324 pp.

Peray, K. E. (1979), Cement Manufacturer's Handbook, Chemical Publishing Co., New York, 382 pp.

Ragan, D. M. (1973), Structural Geology: An Introduction to Geometrical Techniques, 2nd ed., Wiley, New York, 208 pp.

Rendu, J.-M., and M. David (1979), "A New Geostatistical Model for the Estimation of Coal Deposits and Other Sedimentary Deposits", in T. J. O'Neil, ed., 16th Application of Computers and Operations Research in the Minerals Industry, Society of Mining Engineers of AIME, New York, pp. 182-195.

Rice, S. O. (1954), "Mathematical Analysis of Random Noise", in N. Wax, ed., Selected Papers on Noise and Stochastic Processes, Dover, New York, pp. 133-294.

Ripley, B. D. (1981), Spatial Statistics, Wiley, New York, 252 pp.

Rosenblatt, M. (1952), "Remarks on a Multivariate Transformation", Annals of Mathematical Statistics, v. 23, pp. 470-472.

Royle, A. G., S. Khosrowshahi, S. Atmowidjojo, and J. Wijntje (1982), "Bedded Deposits and Alluvial", presented at the SME-AIME Annual Meeting, Dallas, Preprint No. 82-12, Society of Mining Engineers of AIME, Littleton, Colorado, 7 pp.

Rubinstein, R. Y. (1981), Simulation and the Monte Carlo Method, Wiley, New York, 278 pp.

Ryan, T. A., Jr., B. L. Joiner, and B. F. Ryan (1976), MINITAB Student Handbook, Duxbury, North Scituate, Massachusetts, 341 pp.

Scheuer, E. M., and D. S. Stoller (1962), "On the Generation of Normal Random Vectors", Technometrics, v. 4, pp. 278-281.

Schofield, C. G. (1980), Homogenisation/Blending Systems Design and Control for Minerals Processing (With FORTRAN Programs), Trans Tech, Clausthal, Germany, 315 pp.

Sharp, W. E., and L. A. Aroian (1985), "The Generation of Multidimensional Autoregressive Series by the Herringbone Method", Mathematical Geology, v. 17, no. 1, pp. 67-79.

Shinozuka, M., and C. M. Jan (1972), "Digital Simulation of Random Processes and Its Applications", Journal of Sound and Vibration, v. 25, no. 1, pp. 111-128.

Sinding-Larsen, R. (1975), "A Computer Method for Dividing a Regional Geochemical Survey Area Into Homogeneous Subareas Prior to Statistical Interpretation", in I. L. Elliott and W. H. Fletcher, eds., Geochemical Exploration 1974, Elsevier, Amsterdam, pp. 191-217.

Smith, L., and R. A. Freeze (1979), "Stochastic Analysis of Steady State Groundwater Flow in a Bounded Domain, 2: Two-Dimensional Simulations", Water Resources Research, v. 15, no. 6, pp. 1543-1559.

Solow, A. R. (1984), "The Analysis of Second-Order Stationary Processes: Time Series Analysis, Spectral Analysis, Harmonic Analysis, and Geostatistics", in Verly et al. (1984), pp. 573-585.

Switzer, P. (1965), "A Random Set Process in the Plane With a Markovian Property", Annals of Mathematical Statistics, v. 36, pp. 1859-1863.

Switzer, P. (1983), "Some Spatial Statistics for the Interpretation of Satellite Data", Technical Report 4, Department of Statistics, Stanford University.

Verly, G. W. (1984a), "The Block Distribution Given a Point Multivariate Normal Distribution", in Verly et al. (1984), pp. 495-515.

Verly, G. W. (1984b), Estimation of Spatial Point and Block Distributions: The Multigaussian Model, doctoral thesis, Stanford University, 416 pp.

Verly, G. W., M. David, A. G. Journel, and A. Marechal, eds. (1984), Geostatistics for Natural Resources Characterization, Proceedings of the NATO Advanced Study Institute, South Lake Tahoe, California, September 6-17, 1983, D. Reidel, Dordrecht, Holland, 1092 pp.

Webster, R., and P. A. Burrough (1972), "Computer-Based Soil Mapping of Small Areas From Sample Data", Journal of Soil Science, v. 23, no. 2, pp. 210-234.

Witt, J. C. (1966), Portland Cement Technology, 2nd ed., Chemical Publishing Co., New York, 346 pp.

Wohlbier, R. H., ed. (1977), Stacking Blending Reclaiming of Bulk Materials, Trans Tech, Clausthal, Germany, 862 pp.

Xirokostas, D. A., and C. E. Zoppas (1977), "Mathematical Programming Approach to the Problem of Cement Blending Optimization", Cement and Concrete Research, v. 7, pp. 503-514.