2003 Distinguished Instructor Short Course

Distinguished Instructor Series, No. 6

sponsored by the

Society of Exploration Geophysicists

European Association of Geoscientists & Engineers

Geostatistics for Seismic Data Integration in Earth Models

presented by

Olivier Dubrule

EAGE

EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Library of Congress Cataloging-in-Publication Data

Distinguished Instructor Short Course (2003 : Tulsa, Okla.)
Geostatistics for seismic data integration in earth models : 2003 Distinguished Instructor Short Course / presented by Olivier Dubrule ; sponsored by the Society of Exploration Geophysicists [and] European Association of Geoscientists & Engineers.
p. cm.—(Distinguished instructor series ; no. 6)
Includes bibliographical references.
ISBN 1-56080-121-2
1. Prospecting—Geophysical methods—Statistical methods—Congresses. I. Dubrule, Olivier. II. Society of Exploration Geophysicists. III. European Association of Geoscientists and Engineers. IV. Title. V. Series.

TN269.D5725 2003
622'.15'0727—dc21 2003042820

ISBN 1-56080-086-0 (Series)
ISBN 1-56080-121-2 (Volume)

Published 2003

Printed in United States of America

About the author

After graduating from the Ecole Nationale Supérieure des Mines de Paris in 1978, Olivier Dubrule obtained a PhD degree in Petroleum Geostatistics from the same school in 1981. He then worked for Sohio Petroleum Company in the USA (1982-1986), Shell International in The Netherlands (1986–1991), and, since 1991, he has been with Elf and TotalFinaElf, working in Pau (France) and at the Geoscience Research Centre in London (UK). During his career, he has held a variety of staff and management positions where he developed and applied new techniques for reservoir characterization, earth modeling, risk analysis, and uncertainty quantification. After leading the "Earth Modeling and Uncertainty Quantification" group at TotalFinaElf, he became manager of the Geoscience Training and Communication Department in 2001.

Olivier has authored more than twenty papers in the field of geostatistics, reservoir characterization, and earth modeling. In 1991, he received the President's prize of the International Association of Mathematical Geology, for his "Outstanding Contribution to Mathematical Geology by an individual 35 years or younger." In the last decade, he organized and chaired a number of events organized by SPE, EAGE, SEG, and AAPG. He is the author of AAPG Course Notes Series #38: *Geostatistics in Petroleum Geology*, and editor (with E. Damsleth) of "Petroleum Geostatistics," a Special Issue of EAGE's *Petroleum Geoscience* journal, published in 2001.

Olivier is 46 years old, is married to Anne, and has four children: Pauline, Noémie, Arthur, and Thibaud. He is a big fan of Groucho Marx, Woody Allen, Mozart, Marcel Proust, Richard Feynmann, and Steven Jay Gould. He is also a very bad golfer.

EAGE

EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Dear DISC Participant:

It is a great pleasure to welcome you to the sixth annual SEG/EAGE Distinguished Instructor Short Course (DISC), "Geostatistics for Seismic Data Integration in Earth Models," by Olivier Dubrule. The SEG, EAGE, and your local society are proud to provide this premier course in geophysics education.

With rapidly changing technologies, geoscientists around the world have an increasing need to acquire expert geophysical knowledge. SEG's Distinguished Programs, which include the joint SEG/EAGE Distinguished Instructor Short Course, the Distinguished Lecture, and the joint SEG/AAPG Distinguished Lecture, aid in the promotion of technologies that will have a significant impact on geophysics and geology. Likewise, the EAGE supports Geoscience Education through its Distinguished Lecture and Short Courses Programs.

Previous DISC programs included "Time-Lapse Seismic in Reservoir Management," by Mr. Ian Jack of BP Amoco, in 1998; "The Seismic Velocity Model as an Interpretation Asset," by Dr. Phil Schultz of Spirit Energy, a division of Unocal Corporation, in 1999; "Shear Waves from Acquisition to Interpretation," by Mr. Robert Garotta, formerly with CGG, now retired, in 2000; "Seismic Amplitude Interpretation," by Fred J. Hilterman of Geophysical Development Corporation, in 2001; and "Understanding Seismic Anisotropy in Exploration and Exploitation," by Leon Thomsen of BP Amoco, in 2002.

Geophysics education is one of the top priorities for both SEG and EAGE. Because the DISC is an annually renewed program, your participation is key to its success. The Distinguished Programs and DISC have become highly sought opportunities for financial sponsorship and endowment by companies and organizations. In partnership and collaboration with the EAGE, the DISC is sponsored and coordinated throughout Europe, Africa, and the Middle East by the EAGE. The DISC program affords important opportunities for local organizations to provide first-rate geophysical education opportunities at modest cost. The DISC is truly a cooperative effort of many people dedicated to the promotion and advancement of geophysics.

We are honored to have Olivier Dubrule for our 2003 DISC program. This is a great opportunity to learn from one of the recognized experts about the use of geostatistics to integrate seismic and reservoir data. We encourage you to take full advantage of this opportunity to broaden your perspectives through participation in the 2003 DISC.

Sincerely,

Mike Bahorich
SEG President

Paul van Riel
EAGE President

Table of Contents

1 Introduction

1.1 Historical Perspective

In this introduction, we would like to highlight what appear to be the important landmarks in the history of geostatistical applications in the petroleum industry. What do we mean by "geostatistics?" In this course, this term will cover the petroleum applications resulting from the pioneering work of Prof. Georges Matheron and his Research Group at the Centre de Géostatistique de l'Ecole des Mines de Paris. As far as this course is concerned, the main pillars of this work are the developments of variogram-based modeling applications.

Variogram-based modeling applications can be classified in two broad categories, the first of which can be called deterministic geostatistics and is essentially all the development around kriging. We will see later that this covers a very wide number of techniques, including external drift kriging, error cokriging, factorial kriging, and collocated cokriging. Although kriging is a technique based on a stochastic model, it generates one single model as a result, and it is deterministic in that sense.

The second category can be called stochastic geostatistics, and it covers the numerous techniques developed around the conditional simulation concept. Conditional simulation is stochastic in the sense that, as with the Monte-Carlo simulation, it generates a family of "realizations" of 1D, 2D, or 3D models, all compatible with the a priori model and the existing data. With regard to kriging, conditional simulation includes several techniques, such as indicator simulation, collocated cosimulation, or geostatistical inversion. This explains why this one-day course is subdivided in two half-days, the first half-day presenting the basic concepts and the deterministic family of applications, the second half-day covering the stochastic applications (Fig. 1-1). The most complete synthesis of Matheron's work can be found in Chilès and Delfiner (1999). Isaaks and Srivastava (1989), Hohn (1988), and Deutsch (2002) are also other excellent presentations of geostatistics.

Following the work of Matheron, petroleum applications went through different episodes (Fig. 1-2). The first one could be qualified as deterministic mapping. This was the first development of kriging for mapping applications; see, for instance, the papers of Haas and Viallix (1974) or Haas and Jousselin (1976). This period saw the development of commercial mapping applications, such as Bluepack (Renard, 1990). Another important step in the development of 2D mapping applications was Doyen's (1988) paper showing the potential of cokriging for mapping porosity using seismic-derived information and well data.

The mid-1980s to mid-1990s saw the explosion of 3D stochastic (simulation-based) reservoir modeling. This followed the arrival of Prof. Journel who, after closely working with Matheron, joined the faculty at Stanford University. Thanks to the development of sequential algorithms, simulation proved its value for generating heterogeneous 3D reservoir models, a task made even easier by the development of the public-domain GSLIB software library (Deutsch and Journel, 1992). As a result, reservoir and production geologists adopted this new technique as an important step in reservoir

GEOSTATISTICS FOR SEISMIC DATA INTEGRATION IN EARTH MODELS

1. Introduction
2. The Covariance and the Variogram
3. Kriging and Cokriging
4. Conditional Simulation
5. Geostatistical Inversion
6. Quantifying Uncertainties
7. Conclusions

SEG/EAGE DISC 2003

I-1

A HISTORICAL PERSPECTIVE

- Since the 60s, Developments of G. Matheron and the Centre de Géostatistique.
- 70s: Deterministic mapping applications. Development of associated commercial software.
- Mid-80s: 3D stochastic reservoir modelling, public-domain and commercial software.
- Early 90s: First uncertainty quantification approaches.
- Mid-90s: Earth modeling software allows integration between 2D (surface) and 3D (properties) for deterministic *and* stochastic modelling.
- 2000s: New data integration developments.

SEG/EAGE DISC 2003

I-2

modeling workflow (Dubrule, 1998). Commercial sofware also developed quickly, but it remained limited to the generation of heterogeneity models within a somewhat simplistic stratigraphic context.

Another potential application of simulation rapidly emerged — that of uncertainty quantification. The pioneering work of the Norwegian School, led by Prof. H. Omre (Lia et al., 1997), showed that it was possible to combine uncertainties related to geometrical modeling (essentially 2D) with uncertainties associated with 3D heterogeneity modeling. However, if Lia et al.'s work showed that the approach was possible, it also showed that powerful software was needed to make it practical for industrial applications.

The arrival of 3D earth modeling in the mid-1990s, linked with the development of new commercial software (see, for instance, Tinker, 1996, or Dubrule et al., 1997), opened a new era. The dichotomy between 2D surface modeling and 3D property modeling disappeared, thus giving a tremendous boost to cross-discipline integration. Suddenly it became possible to quickly quantify the impact of a change in the time-to-depth conversion velocity model on the predicted production profiles. It also became possible, as Lia et al. had predicted, to quantify the joint impact of subsurface uncertainties, including those associated with time-to-depth conversion, property modeling, and dynamic parameters, on production forecasts (Corre et al., 2000).

It is not a coincidence that, at about the same time that 3D earth modeling software was developing on an industrial scale, new applications emerged that led to a better integration of seismic data in the construction of 3D reservoir models. Cokriging became a viable approach for combining seismic and reservoir data (Xu et al., 1992), and the relationships between geostatistical techniques and the approaches more commonly used by geophysicists tended to narrow. Doyen et al. (1996) showed that collocated cokriging could be simply put in a Bayesian context, whereas Haas and Dubrule (1994) proposed geostatistical inversion as a stochastic high-resolution addition to deterministic stratigraphic inversion techniques.

Over the last 30 years, we have seen that the development of geostatistical techniques, combined with that of earth modeling software, has led to better integration of geostatistics in the disciplines' workflow, whether the discipline is geology, geophysics, or reservoir engineering. This is very good news. However, much progress can still be made, because geostatistics is still regarded by many geoscientists as a random-number-generating black box, disconnected from the constraints provided by the disciplines. This may be partly the geostatisticians' responsibility, because they have developed many new techniques, sometimes using a very heavy mathematical formalism, without clearly indicating what was really important and how it related with what the various disciplines were used to doing.

Now, following this natural-selection period of about 30 years, one of our goals is to clarify the actual achievements of geostatistics and their relationship to the practice of various geoscience disciplines — especially geophysics.

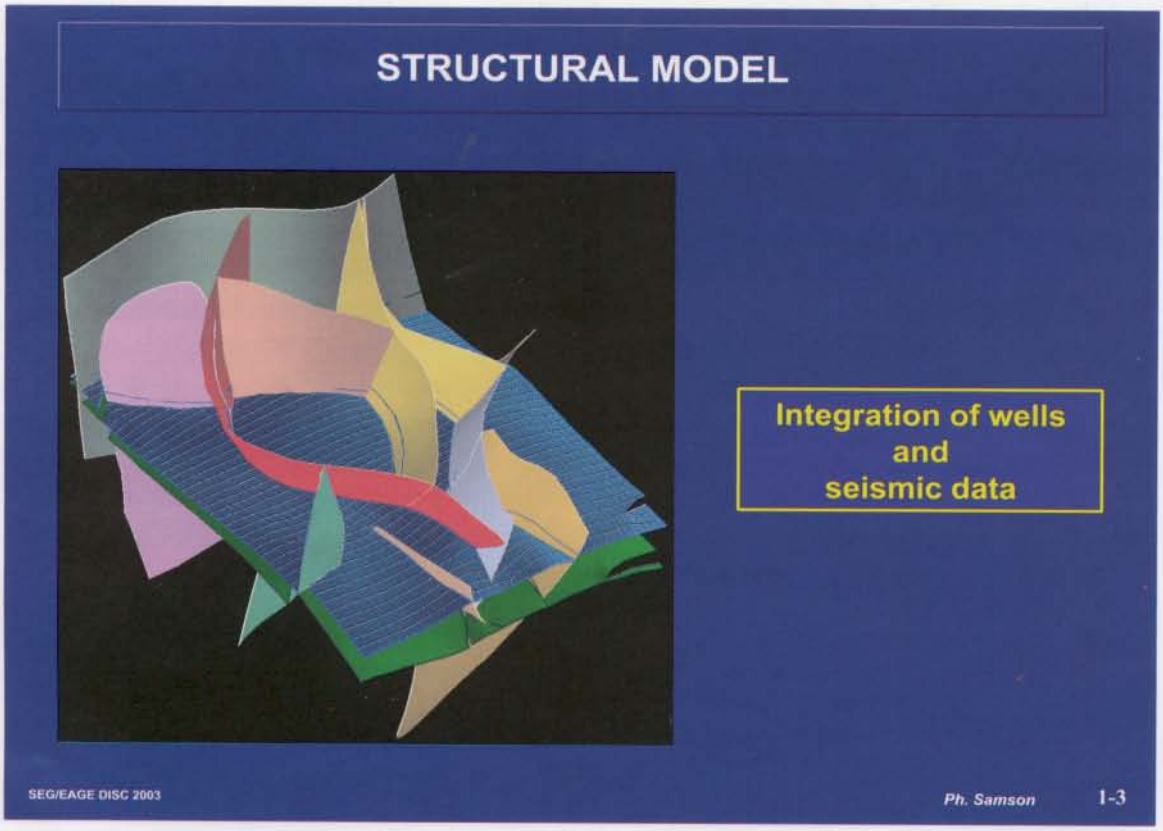
1.2 The Role of Geostatistics at Different Steps of the Earth Modeling Workflow

The first element in the construction of a 3D earth model is the structural framework

(Fig. 1-3). It describes, through faults and surfaces, the skeleton within which the meat of the property model will be incorporated. In most cases, the structural model is constructed from 2D or 3D seismic data, although it may be built from wells alone if no seismic data are available. We will see examples of modeling when only well data are available, but we will spend more time discussing the geostatistical approaches for velocity mapping and for time-to-depth conversion (error cokriging, factorial kriging, external drift approach, and the like). We will also discuss how uncertainties that affect the geometrical model can be quantified with geostatistics.

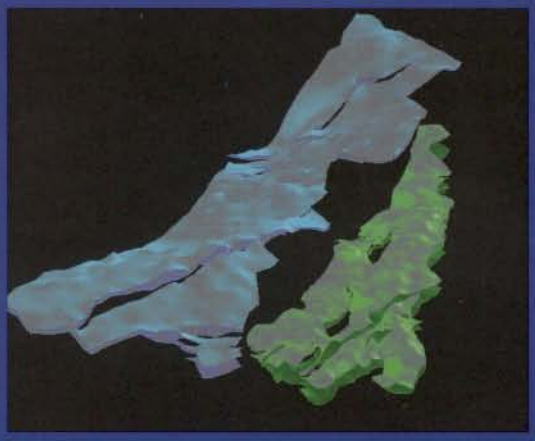
In many cases, the structural model is not sufficient to define the stratigraphic framework. Important stratigraphic surfaces may be recognizable from seismic data and wells or from wells alone. Fig. 1-4 shows examples of the envelope of a turbidite channel complex picked from a good-quality, deep-offshore seismic data set (left) and stratigraphic surfaces picked on wells alone in a shallow-marine reservoir (right). In the latter example, the interpolation of the stratigraphic surfaces identified on wells is guided by the main structural surfaces seen on seismic data, with the assumption, in that case, that the surfaces are subparallel. The model combining structural and stratigraphic surfaces is the 3D geometrical model.

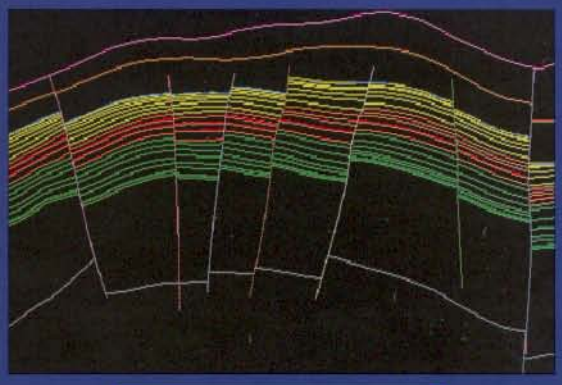
The next step is very crucial and probably is the most important advance brought about by earth modeling software. It is the construction of the stratigraphic grid, which will effectively constitute the link between the geometrical and the property models. The simplest form of stratigraphic grid is shown in Fig. 1-5. Each grid cell in the stratigraphic grid has two coordinates: (x,y,z) corresponding to its absolute location in space,



STRATIGRAPHIC SURFACES AT THE RESERVOIR SCALE

Deep offshore



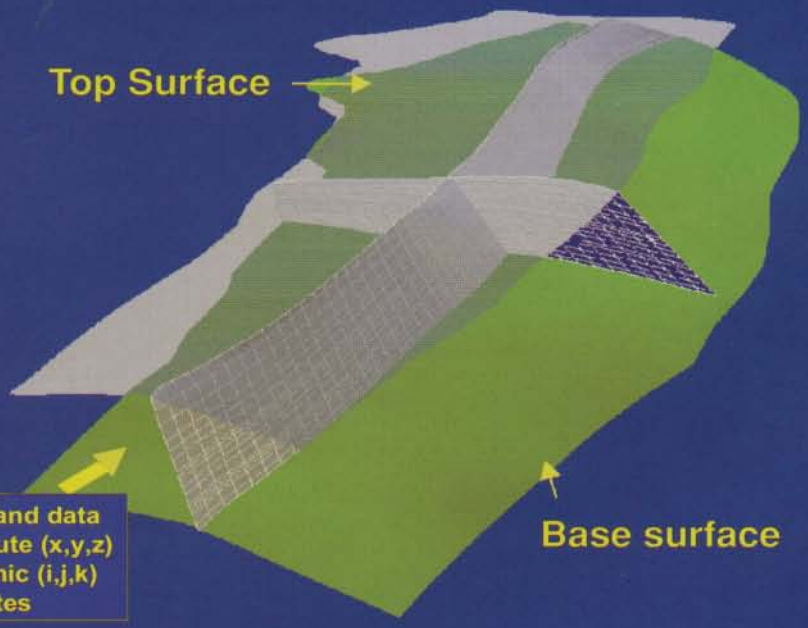


Shallow marine

SEG/EAGE DISC 2003

1-4

EXAMPLE OF A SIMPLE STRATIGRAPHIC GRID



SEG/EAGE DISC 2003

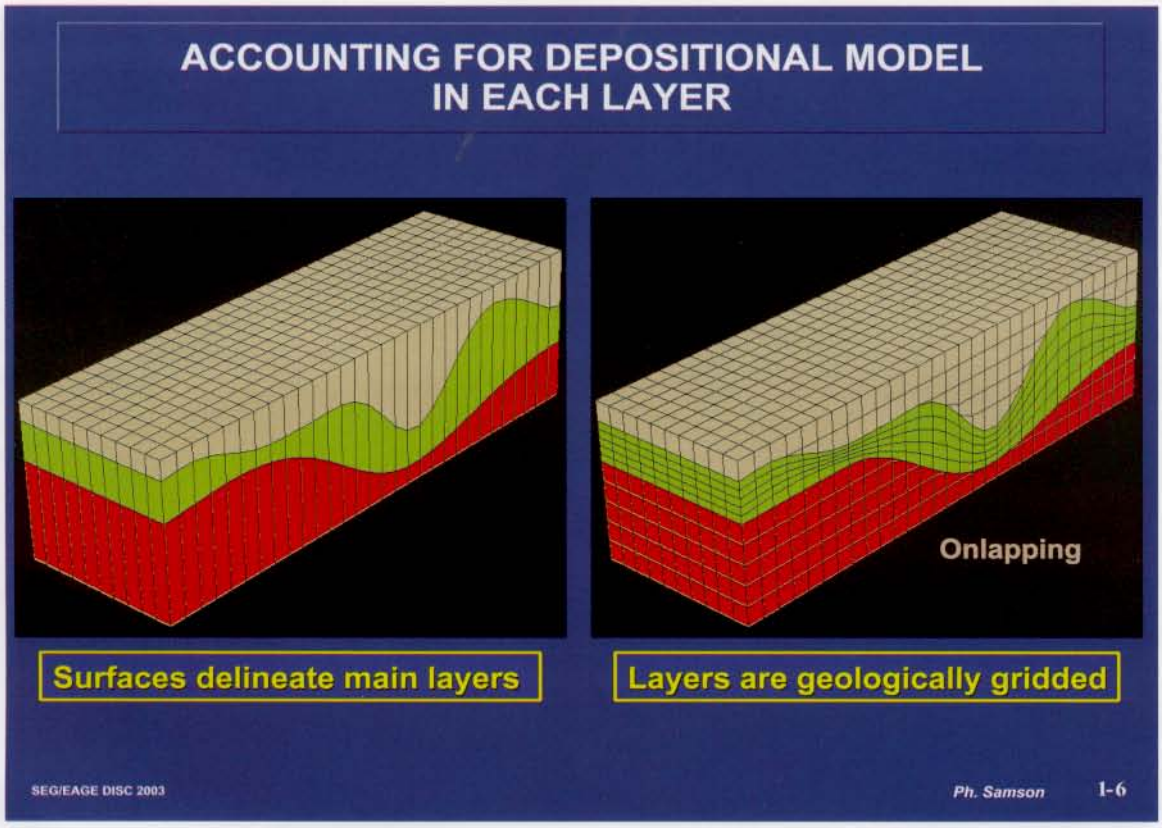
1-5

and (i,j,k) corresponding to its relative location in the stratigraphic grid. The stratigraphic grid can be seen as a map from the (relative) stratigraphic (deposition) space to the (absolute) depth space. Obviously, when interpolating properties in 3D between surfaces of the geometrical model, one must incorporate stratigraphic control. This means that interpolation must take place in the stratigraphic space. However, the result of this interpolation must be displayed in the absolute structural space.

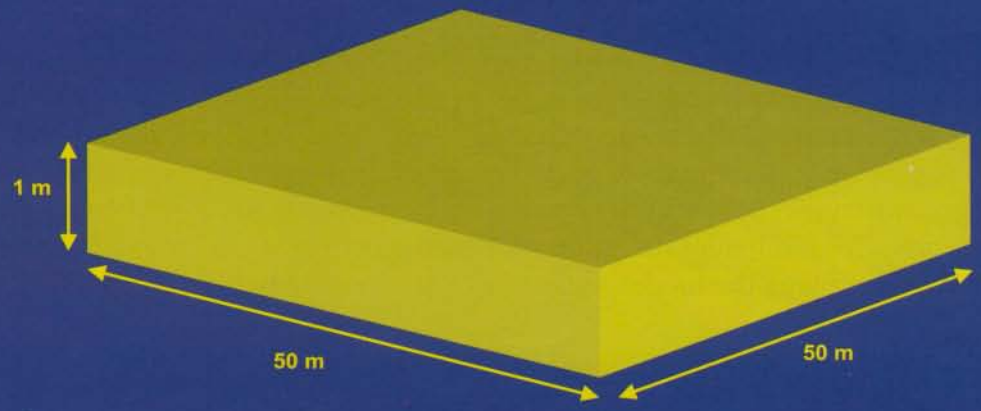
In practice, geologic considerations are used to construct the stratigraphic model (Fig. 1-6). For example, in cases where palaeohighs controlled deposition, onlapping stuctures can be incorporated into the stratigraphic grid. In cases in which a good sequence stratigraphic model has been established, sea-level falls may generate erosional truncations that can also be represented in the earth model. Of course, stratigraphic scenarios can be combined — for instance, onlapping can be combined with erosion.

Experience shows that a typical grid cell of the earth model is around 1 m thick and a few tens of meters wide (Fig. 1-7). This difference between thickness and lateral extent is due to the assumption that geologic variations are much more rapid along the vertical direction than along directions parallel to stratigraphy. This also implies that the number of grid cells in an average-size reservoir often will be on the order of several million.

What will be the impact of the stratigraphic grid assumptions on property modeling? Fig. 1-8 shows that it can be very significant. The three models all share the same well data at their left end and the geostatistical model is also the same. However, just because of the choice of stratigraphy, the onlapping model contains much fewer red



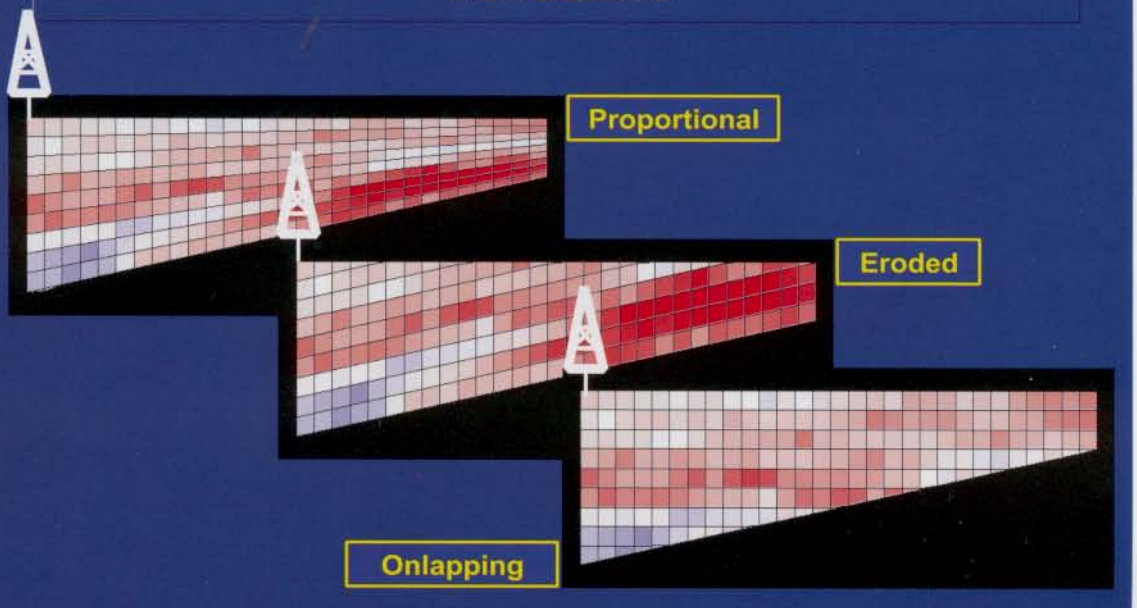
TYPICAL SIZE OF INDIVIDUAL GRID CELL IN 3D EARTH MODEL



Typically, 1 million grid cells in a 5 km x 5 km x 100 m reservoir model

SEG/EAGE DISC 2003 1-7

IMPACT OF STRATIGRAPHIC GRID ON PROPERTY MODELLING



Stratigraphic grid controls lateral correlations

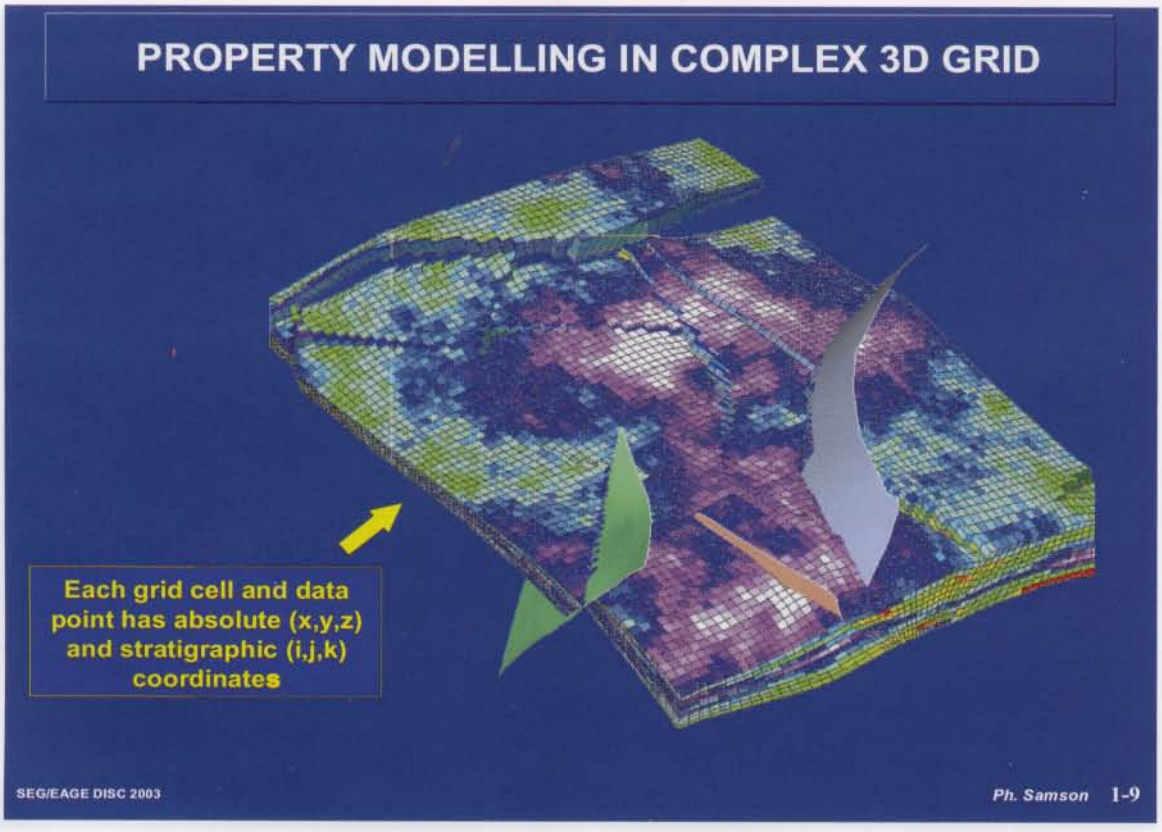
SEG/EAGE DISC 2003 *Ph. Samson* 1-8

values than do the two other models. Fig. 1-9 shows the impact of using a stratigraphic grid in a faulted model: whereas the (x,y,z) grid is obviously affected by faulting the (i,j,k) stratigraphic grid is not, because the assumption has been made that faulting was postdepositional. Therefore two grid cells on each side of a fault are interpolated as if they were neighbors, because they were neighbors at the time of deposition.

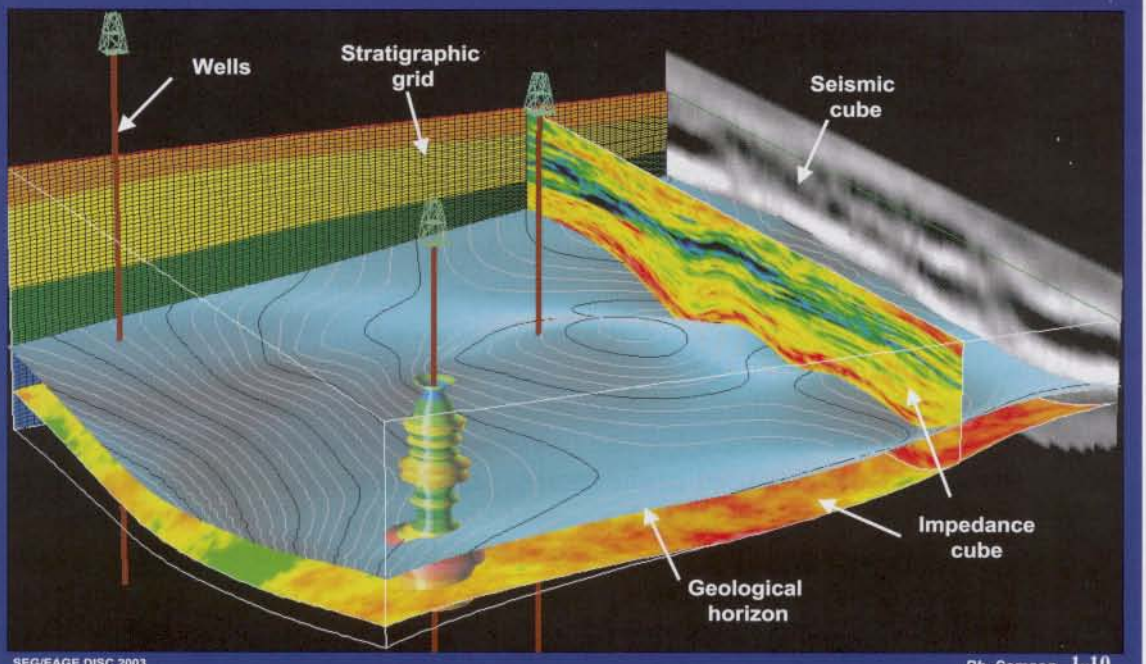
Fig. 1-10 is a nice example of the amount of information that can be integrated and displayed with a 3D earth model. Surfaces and the acoustic-impedance properties along these surfaces are both shown, together with amplitude and acoustic impedances along cross-sections. An acoustic-impedance log is also displayed, together with a cross-section of the stratigraphic grid, with a different color corresponding to each stratigraphic layer. The use of 3D visualization environments, with participation of geoscientists from different disciplines, offers the perfect setup for validating and, if necessary, improving the model — thanks to different input brought by different disciplines.

Standard flow simulators cannot handle more than approximately 100,000 grid cells. The geometry of the simulation grid is also less flexible than that of the earth model grid. This means that an upgridding and an upscaling stage are often required to translate the earth model into a dynamic model. These topics are too broad to be discussed here. Christie (1996) is a good reference on the subject.

Fig. 1-11 identifies the modeling stages for which geostatistics is often used. The different techniques listed on the right-hand side will be discussed during this course. Obviously, geostatistics plays a role in the construction of the geometric model, through



EXAMPLE OF 3D STATIC EARTH MODEL



SEG/EAGE DISC 2003 Ph. Samson 1-10

GEOSTATISTICS AND EARTH MODELLING

Steps in earth model construction

- Structural model
- Layering and well correlations
- Gridding
- Property modelling
- Upscaling and construction of dynamic model

Useful geostatistical functionalities

TIME-TO-DEPTH CONVERSION
Kriging, external drift, error cokriging, collocated cokriging, factorial kriging

UNCERTAINTY QUANTIFICATION
Impact of picking uncertainty and velocity model uncertainty

GENERATION OF HETEROGENEOUS MODEL
Conditional simulation (facies and properties)

MULTIDISCIPLINARY DATA INTEGRATION
Collocated cokriging

UNCERTAINTY QUANTIFICATION
Multi-realization by conditional simulation

1-11

kriging and through the quantification of associated uncertainties. It is also used in the construction of the static and of the dynamic models, through the 3D modeling of geological facies and of petrophysical properties, and through the quantification of their associated uncertainties.

1.3 The Goal of This Course

For both geometry and property modeling, deterministic (kriging-based) or stochastic (conditional-simulation-based) geostatistical approaches can be used. The reasons for using one rather than the other will be discussed later. As the title of the course indicates, we will focus especially on the geostatistical techniques that can be used to incorporate seismic-derived information in the earth model, in combination with well data and geological constraints.

We will clarify the current range of applications of geostatistical techniques, from 2D mapping applications to the generation of 3D heterogeneity realizations fully constrained by seismic data. However, we will not cover all methods, but only those that appear to be the most widely used or the most promising for the future. We will not have time to cover any topic in detail, but, hopefully, after following this course, people will be familiar with the basic concepts and will be able to find their way through the maze of geostatistical terminology and applications.

In the world of geophysics, geostatistics is often perceived as a black-box approach that is often reduced to the generation of random numbers between wells. We will show that this is far from true and that geostatistical applications can be understood as extensions of techniques that are familiar to most geophysicists and interpreters.

We will discuss the relationships between geostatistics and approaches such as Bayesian or regularization-based inversion methods, filtering, Fourier analysis, or splines. A number of short mathematical developments will be given in the text and in the figures, for the benefit of readers interested in more details than it is possible to cover in a one-day course. But these developments are not rigorous. We only wish to give a flavor of where the theoretical relationships may lead, and we apologize — once and for all — for the mathematical shortcuts.

1.4 Basics of Univariate Statistics

To understand geostatistics, one must know a number of basic statistical results. In the following, we focus on the statistical results that appear most relevant to the understanding of geostatistics. The reader familiar with statistics can skip this chapter.

1.4.1 Random variables

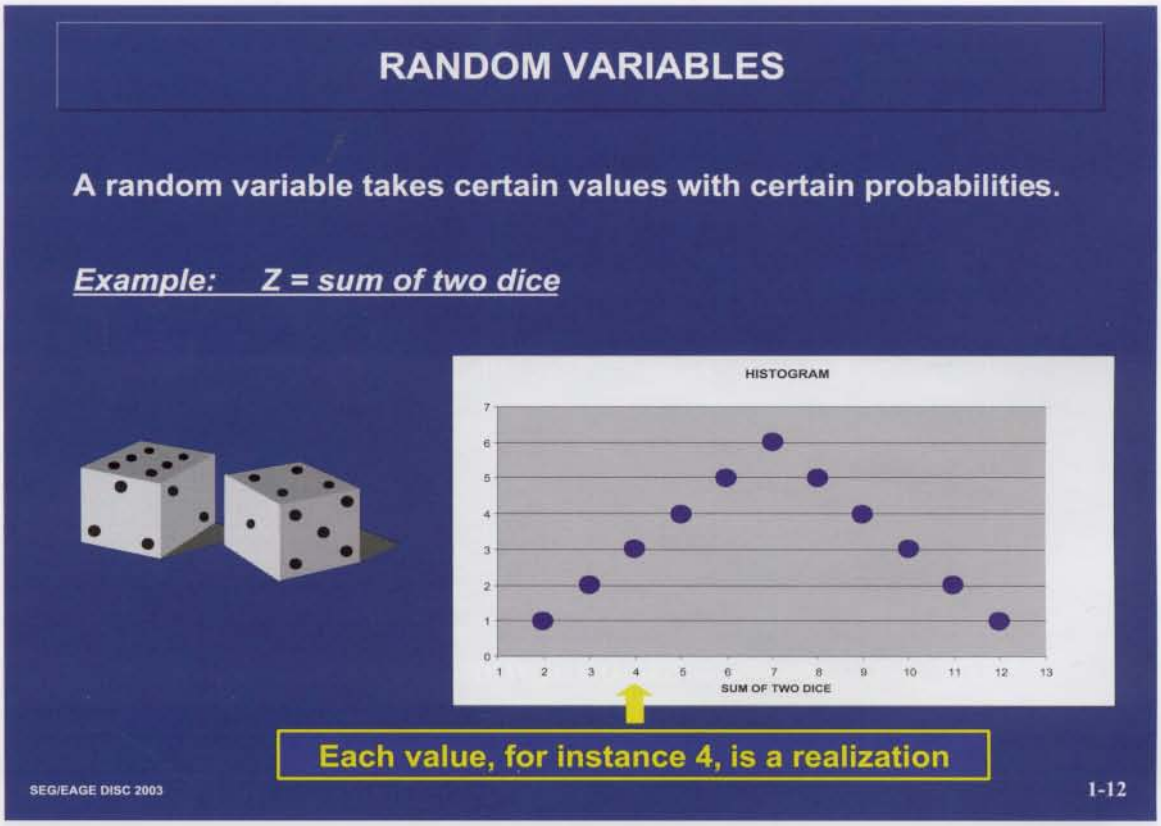
The concept of the random variable is crucial, because the geostatistical model assumes that the value of any property $z(\mathbf{x})$, whether $\mathbf{x}$ is a point along a line, in the plane, or in 3D space, is the realization of a random variable $Z(\mathbf{x})$. There are many complicated mathematical ways to define random variables. Let us avoid them and simply say that a random variable is one that takes certain values with certain probabilities.

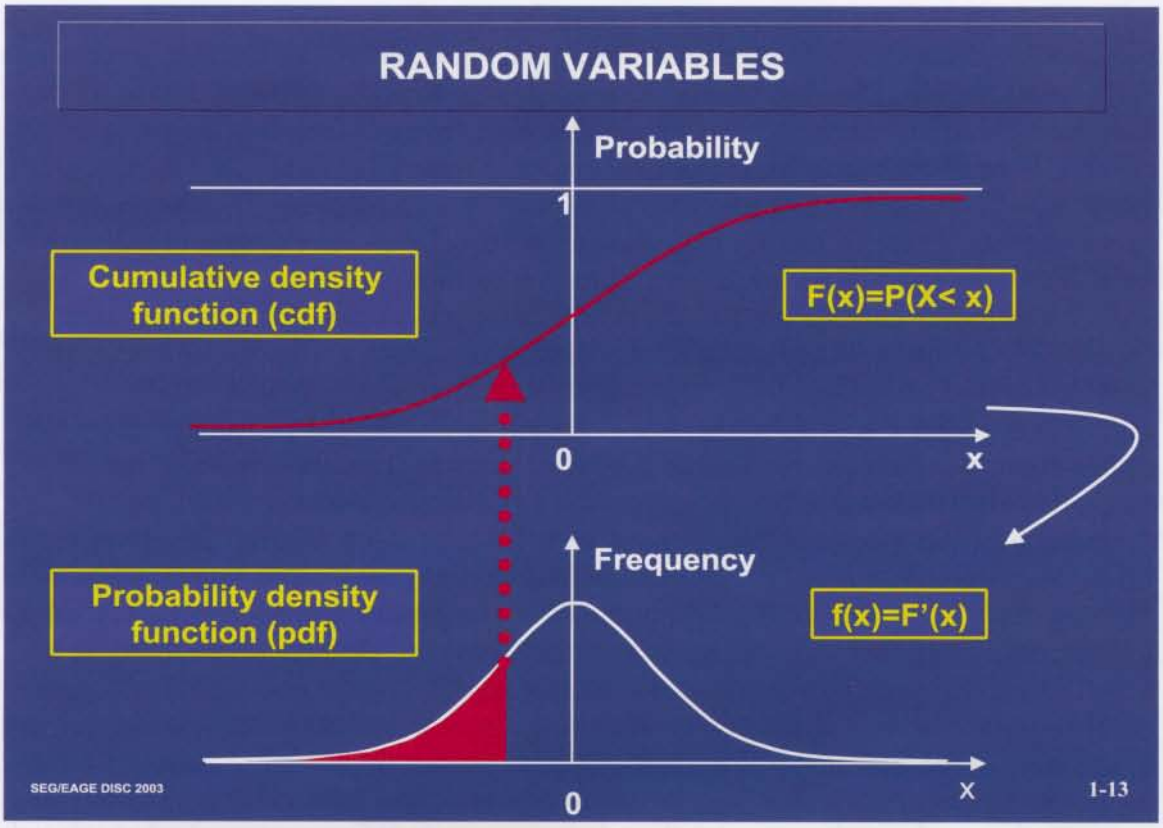
There are discrete random variables, which take only a small number of integer

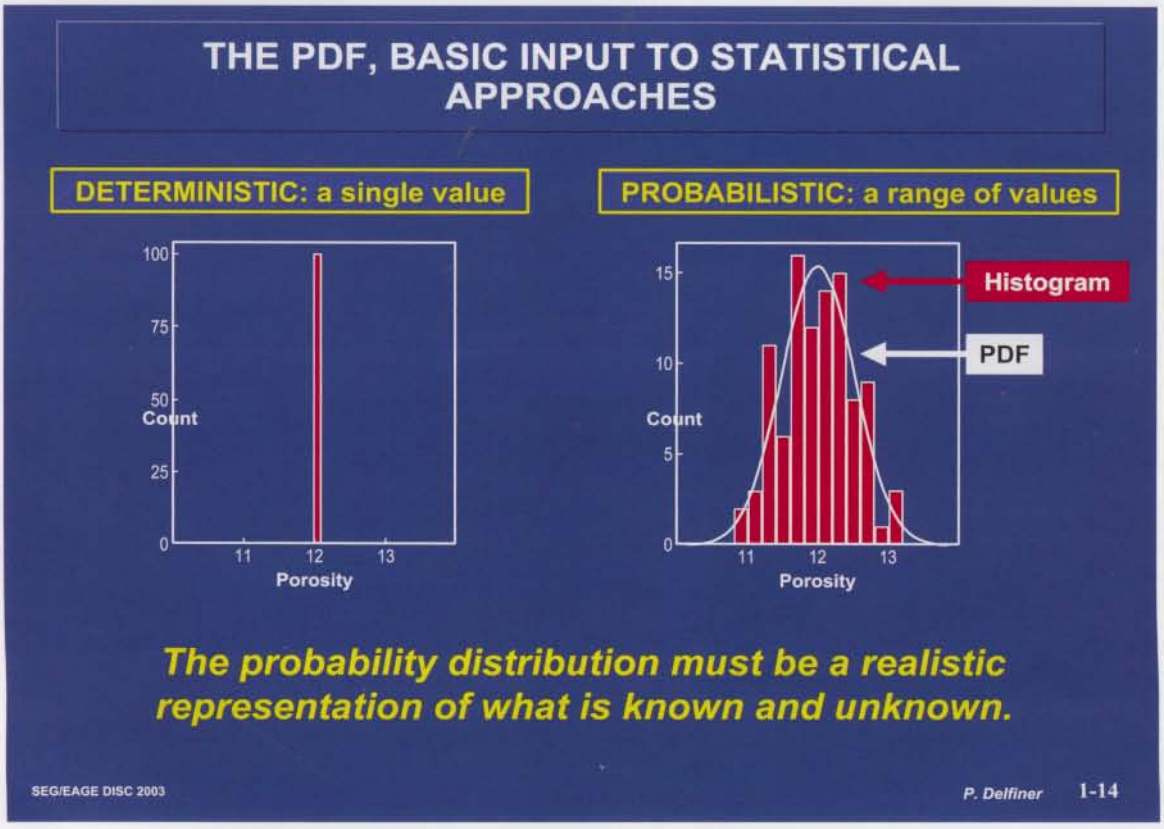
values. Fig. 1-12 shows the simple example of the sum of two dice, which takes values between 2 and 12 with different probabilities. The figure shows the number of ways each value can be obtained out of the 36 possible results of the throw of the two dice. This can be readily translated into probabilities by dividing by 36. This is a discrete variable, because its histogram is made of a finite number of columns. A random variable is usually designated using an uppercase letter (usually X, Y, or Z), whereas one of its realizations is designated by a lowercase letter (usually z). An example of a discrete variable is lithology, which can be, for instance, 0 for shale and 1 for sand.

A continuous random variable is one that takes real values. Porosity, acoustic impedance, and permeability are examples of such parameters. They are characterized by a probability density function (pdf) and a cumulative density function (cdf). The value of the cdf for x is equal to the area at the left of x under the pdf plot (Fig. 1-13). Of course, in situations where a value is perfectly known (porosity measured on a plug), the pdf is reduced to one single bar corresponding to the measured variable. Otherwise, the pdf measures the degree of knowledge available about a parameter (Fig. 1-14). The larger the spread, the poorer this knowledge is.

Usually a pdf is derived from a combination of a priori knowledge that a specialist has about a parameter: "in a braided-streams environment, I expect permeability to be high," and actual data from the area of interest. Examples from Capen (1976) or Rose (2001) have shown how difficult it could be to provide a realistic measure of uncertainty, and we will see later that this is a difficult issue in the application of statistical approaches to natural phenomena. The statistical formalism, as such, is perfectly cor-




RANDOM VARIABLES
Probability
1
Cumulative density function (cdf)
F(x)=P(X< x)
0
x
Frequency
Probability density function (pdf)
f(x)=F'(x)
SEG/EAGE DISC 2003
0
X
1-13


THE PDF, BASIC INPUT TO STATISTICAL APPROACHES
DETERMINISTIC: a single value
PROBABILISTIC: a range of values
100
75
50
Count
25
0
11
12
13
Porosity
15
10
Count
5
0
Histogram
PDF
11
12
13
Porosity
The probability distribution must be a realistic representation of what is known and unknown.
SEG/EAGE DISC 2003
P. Delfiner
1-14

rect, once all the inputs have been quantified. But, if things are not done carefully and the assumptions are not right, the "garbage in, garbage out" law will apply.

1.4.2 Mean, variance, standard deviation, and support effect

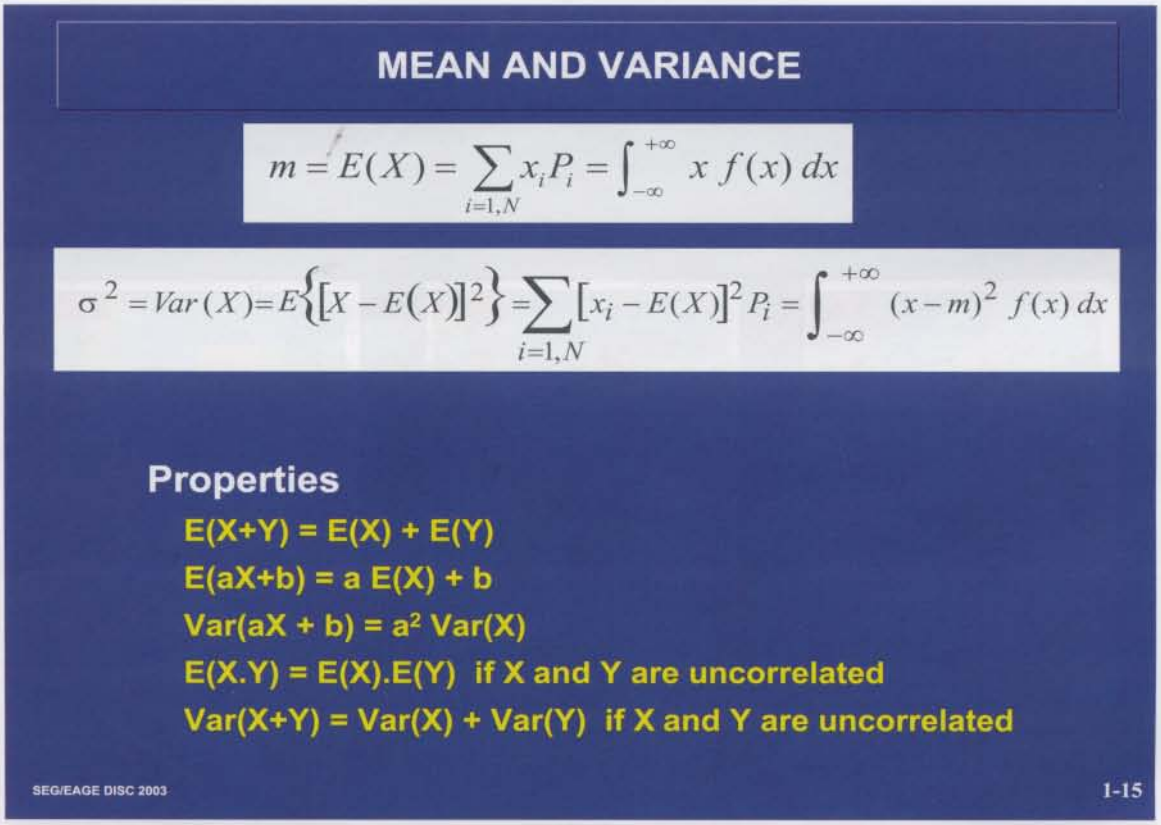
Fig. 1-15 gives basic definitions of mean and variance. For the sake of simplicity, we do not make a distinction in the notations between the population mean and the sample mean, or between the population variance and the sample variance. Fig. 1-16 gives an example of how variance can be calculated. A number of points are worth discussing.

- The mean

The mean of the sum is equal to the sum of the means. This means that if the histogram of a large number of values picked on a map is calculated, the mean of this histogram will be independent of the block size on which it is calculated (Figs. 1-17 and 1-18). It does not matter whether the values have been averaged over a certain block size before the histogram was calculated.

- The variance

On the other hand, the variance of the sum is not equal to the sum of the variances. The formulas show that if a large number of uncorrelated variables are averaged, the variance of the average is inversely proportional to the number of variables that are



AN EXAMPLE OF VARIANCE CALCULATION

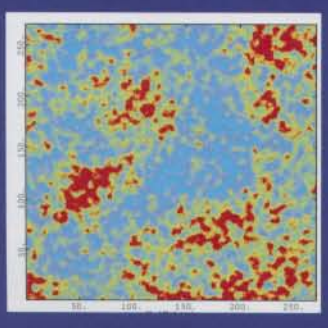
Does not change if a constant is added to all data.

If all data are multiplied by k the variance is multiplied by k^2.

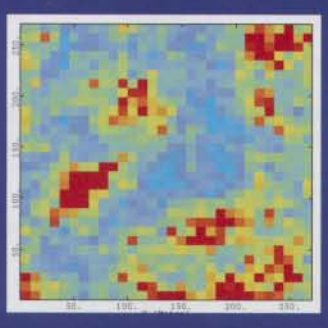
	X_i	X_i - mean	squared	squared X
	27	9	81	729
	0	-18	324	0
	19	1	1	361
	4	-14	196	16
	13	-5	25	169
	50	32	1024	2500
	8	-10	100	64
	13	-5	25	169
	35	17	289	1225
	11	-7	49	121
sum	180	0	2114	5354
mean	18	0	211	535

σ^2=Var(X) = 535 - 18^2 = 211

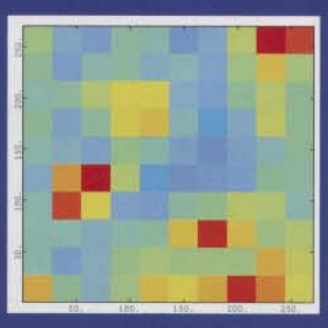
THE IMPACT OF AVERAGING (1) THE DATA



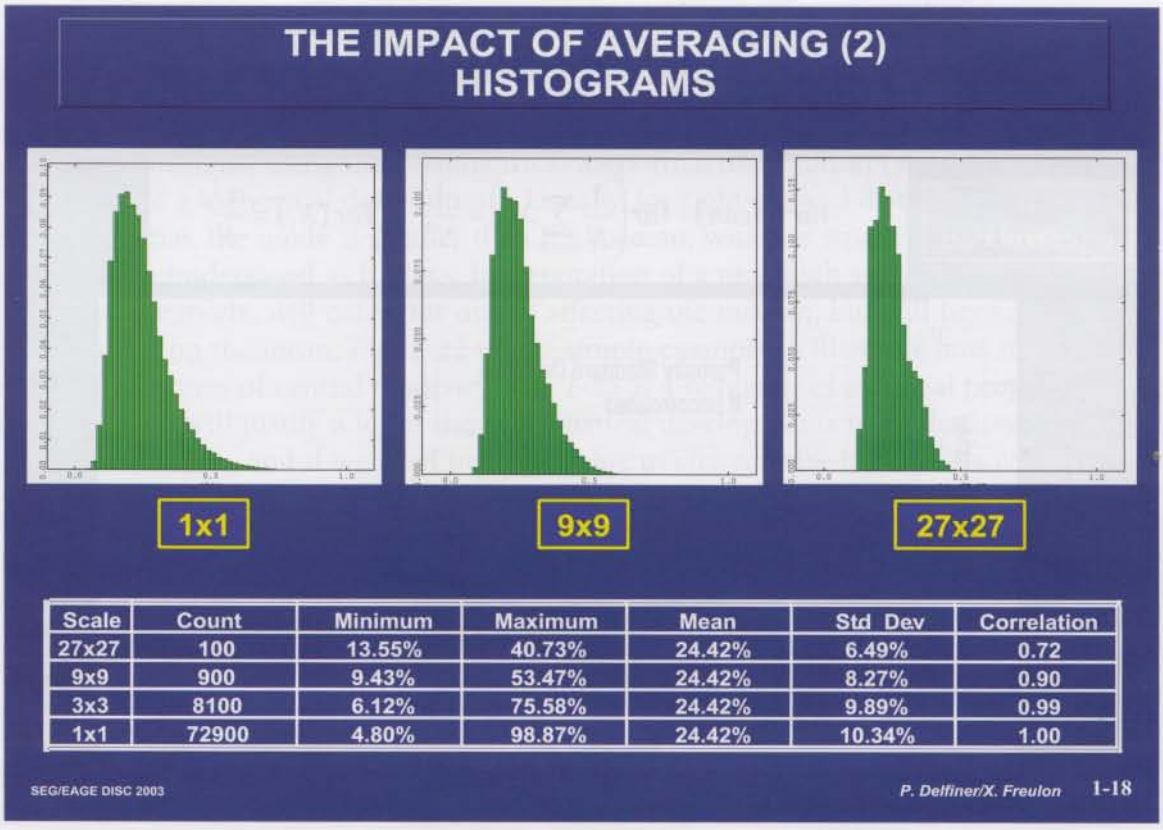
1x1



9x9



27x27



Scale	Count	Minimum	Maximum	Mean	Std Dev	Correlation
27x27	100	13.55%	40.73%	24.42%	6.49%	0.72
9x9	900	9.43%	53.47%	24.42%	8.27%	0.90
3x3	8100	6.12%	75.58%	24.42%	9.89%	0.99
1x1	72900	4.80%	98.87%	24.42%	10.34%	1.00

averaged (Fig. 1-19). In petroleum applications, we will see that the phenomenon is somewhat attenuated. If porosity values within a volume are averaged, the variance of the averaged values will decrease, but not as fast as given by the equation of Fig. 1-19, because the values are usually correlated in space. This is what happens with the examples of Figs. 1-17 and 1-18: Spatial correlation between neighboring porosity values implies that the compensation effect from one to another is not as extreme as in the situation where data are completely uncorrelated (Fig. 1-19). This is what geostatisticians call the "support effect."

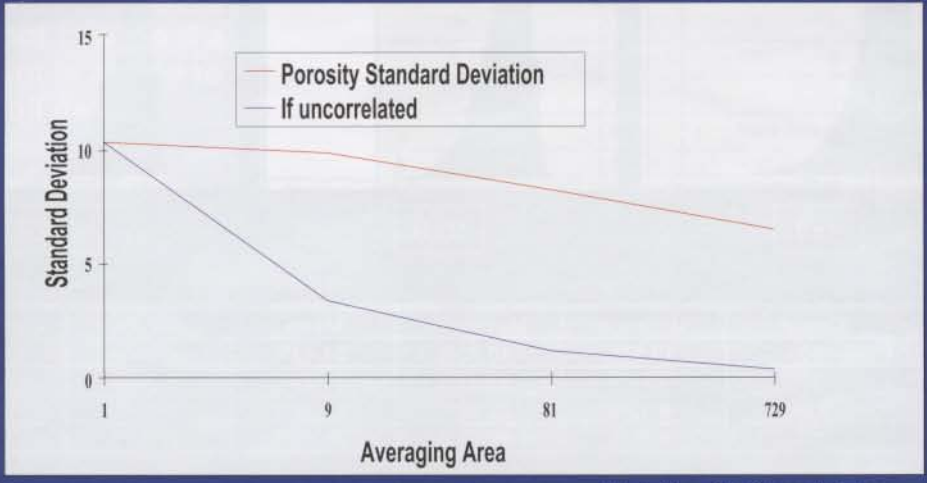
Because of the support effect, it is meaningless to talk about "the variance of porosity" over a given reservoir (while it is correct to talk about the mean of porosity, at least if there is no systematic trend). The variance of porosity values is support-related, in the sense that the variance of plug values of porosity will usually be larger than the variance of values derived from logs, because the latter are averaged over a larger volume than the former. Fig. 1-20, from Kelkar (2000), shows a model of how porosity and its variance vary as a function of the averaging scale.

This also applies to an even larger scale. It would be completely meaningless to use the histogram of plug porosities to derive the variance associated with field-averaged porosity values! We will see later that, thanks to their use of the spatial covariance or variogram, geostatisticians can theoretically predict how variance changes as the averaging volume increases.

THE IMPACT OF AVERAGING (3) VARIANCE DECREASES AS AVERAGING AREA INCREASES

If uncorrelated

$$Var(Mean) = Var\left(\frac{1}{N}\sum_{i=1,N} X_i\right) = \frac{1}{N^2}\sum_{i=1,N} Var(X_i) = \frac{\sigma^2}{N}$$

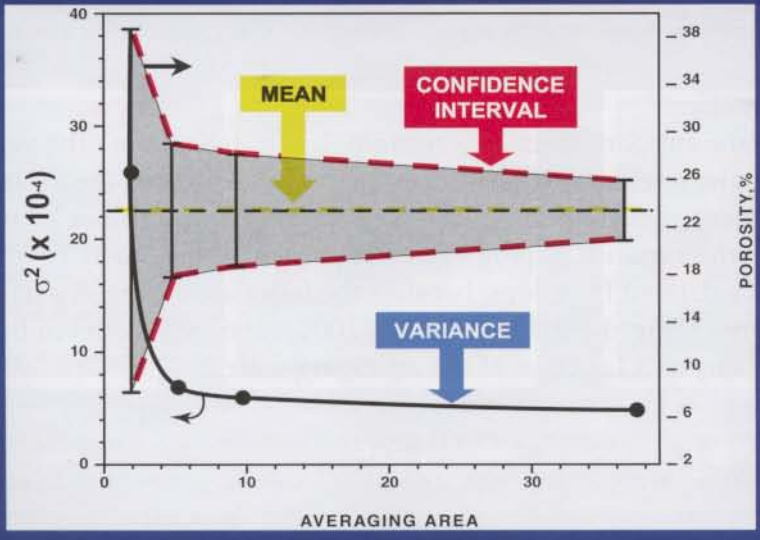


SEG/EAGE DISC 2003

Adapted from P. Delfiner and Freulon

1-19

SPREAD AND VARIANCE OF POROSITY AS A FUNCTION OF AVERAGING VOLUME (KELKAR, 2000)



SEG/EAGE DISC 2003

1-20

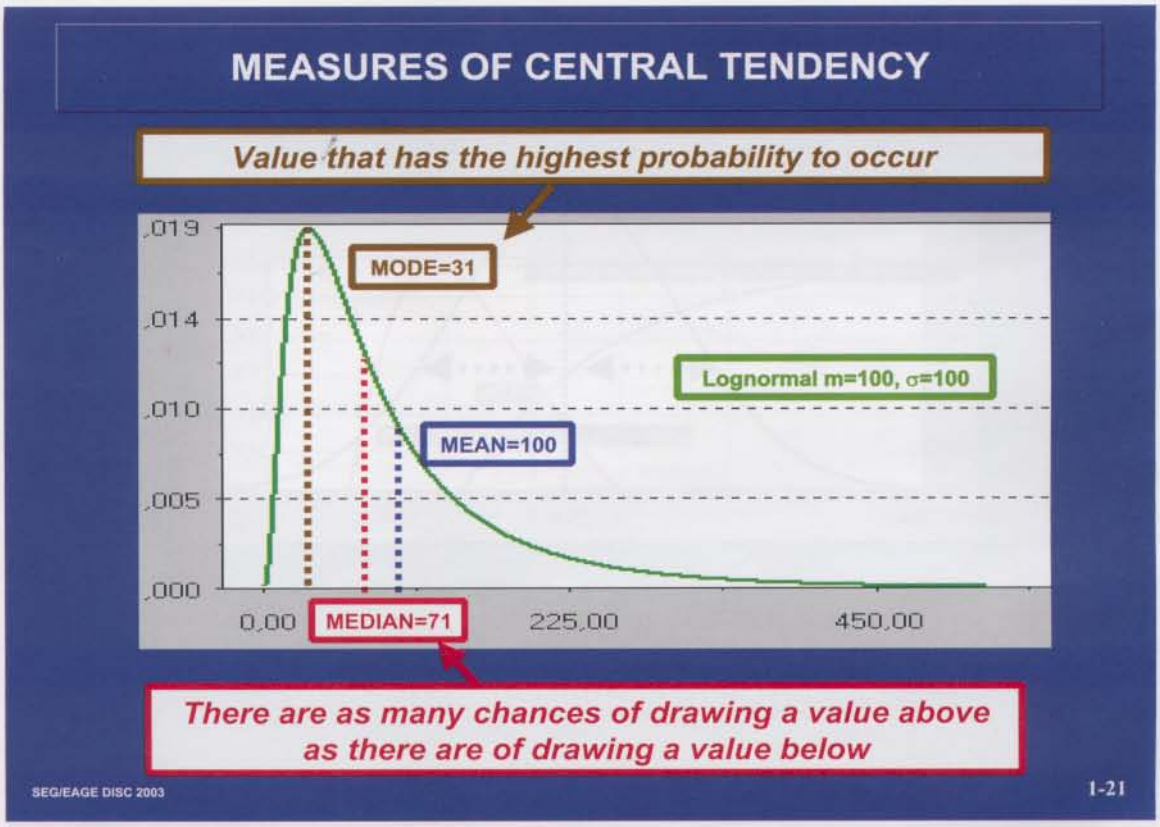
1.4.3 *Measures of central tendency and quantiles*

- Measures of central tendency

There is often confusion between various measures of central tendency. Fig. 1-21 clarifies the definitions, using the asymmetrical lognormal distribution (see Fig. 1-27 for the definition of a lognormal distribution). Usually, for right-skewed distributions such as the lognormal, the mode is smaller than the median, which is smaller than the mean. This can be understood as follows: Incorporation of a new, high value will have no effect on the mode, will count for one in affecting the median, but will have a significant impact on the mean. Fig. 1-22 uses a simple example to illustrate how to calculate various measures of central tendency. Fig. 1-23 is a reminder of a crucial property of the mean, which will justify a lot of the geostatistical developments to be discussed later on. If a pdf is known, and if we need just one value to characterize it in such a way that the average error made is minimal, the best value to use is the mean.

- Quantiles

In situations where a probabilistic approach has been used, SPE-WPC guidelines for reporting reserves (SPE/WPC, 1997) are based on quantiles. The quantile definition is best understood by using the cumulative density function (Fig. 1-24). Rather than directly associating proven, probable, or possible reserves figures to quantiles, the practice in the industry is often to define "P" values. This amounts to taking a positive view of things, basing the definitions on the probability of having "more than" (P90 means

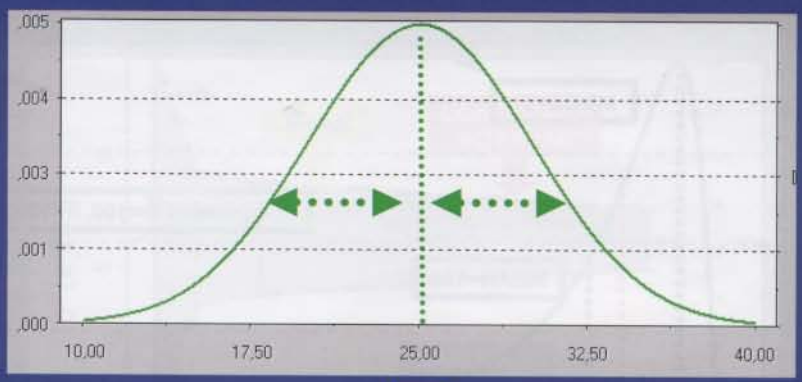


EXERCISE: CENTRAL TENDENCY COMPARISON

DATA	MEAN	MODE	MEDIAN
10, 20, 30, 40, 50	30	N/A	30
10, 20, 30, 40, 50, 50	33.33	50	35
1, 3, 3, 3, 7, 23, 100	20	3	3
1, 1, 2, 3, 3	2	tie: 1, 3	2
2, 2, 40, 4000	1011	2	21

SEG/EAGE DISC 2003 P. Delfiner 1-22

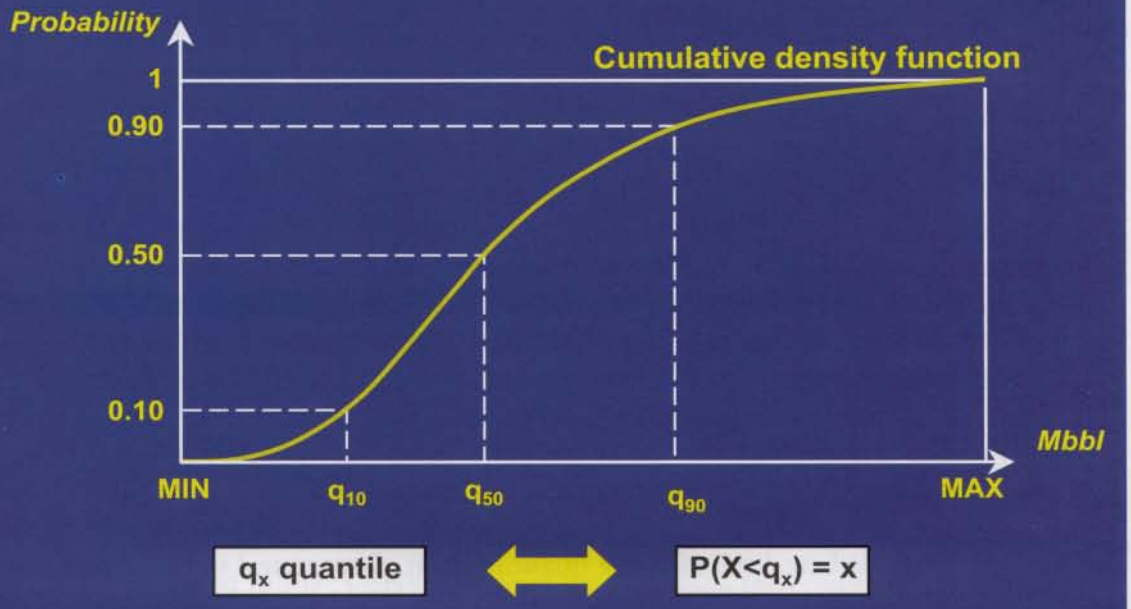
A CRUCIAL PROPERTY OF THE MEAN



Mean[(Z-a)^2] is minimum for a=m

SEG/EAGE DISC 2003 1-23

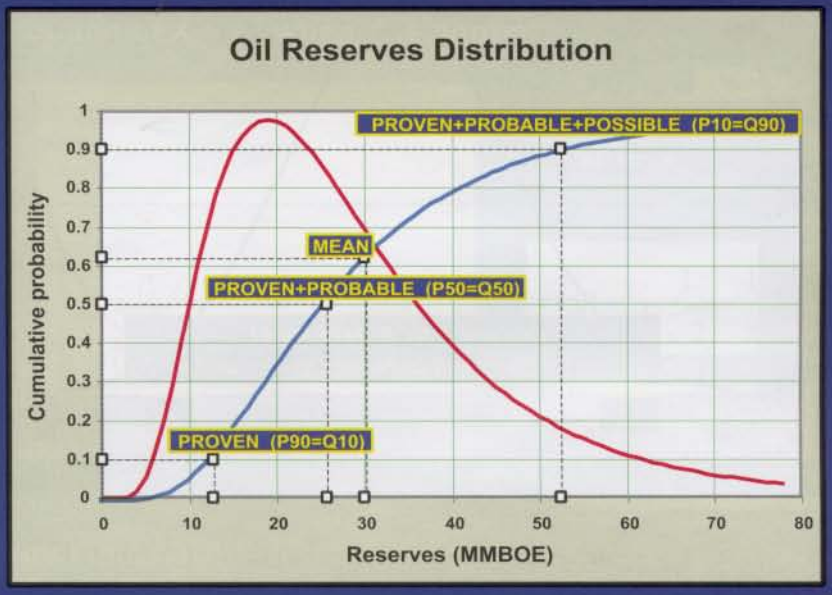
QUANTILES



SEG/EAGE DISC 2003

1-24

DEFINITIONS USED WHEN REPORTING RESERVES



SEG/EAGE DISC 2003

P. Delfiner

1-25

that there is 90% chance to find more) rather than "less than." SPE/WPC recommends taking the 10% quantile as the value of proven reserves in situations in which a probabilistic approach has been used (Fig. 1-25).

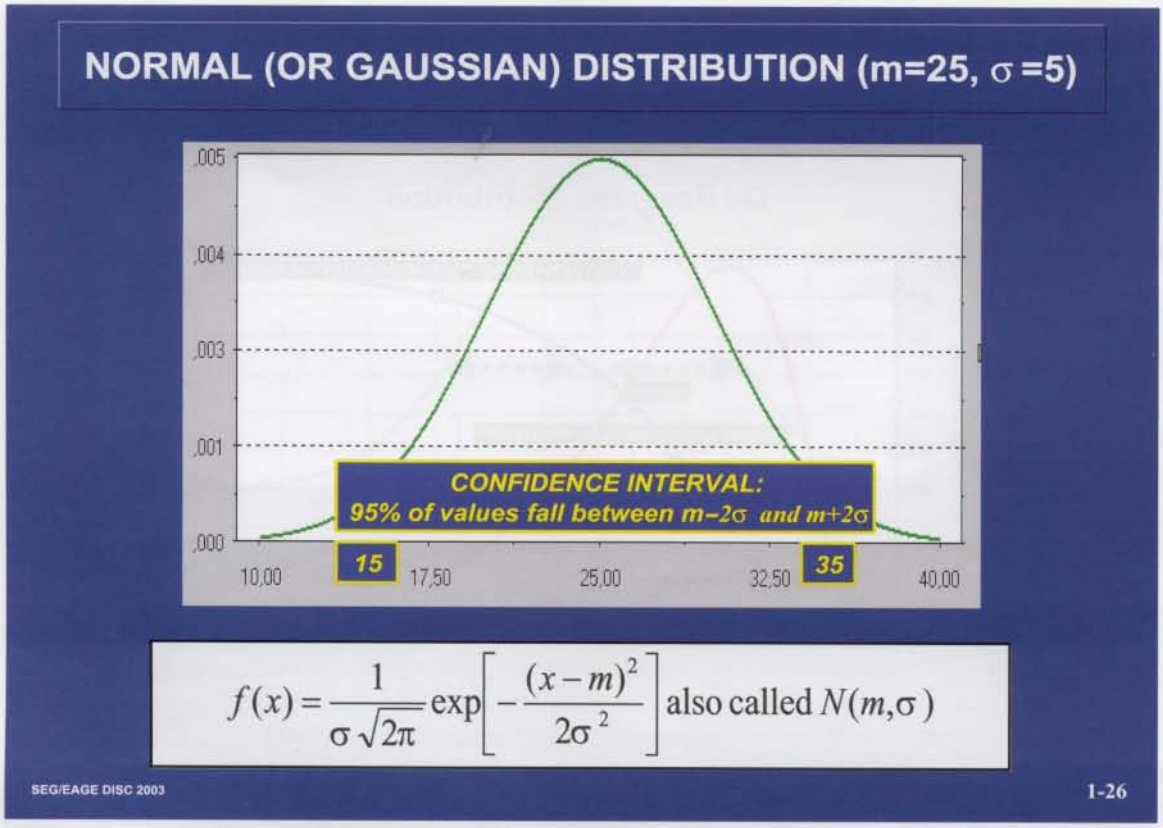
A result that will prove useful later is that, if a random variable is transformed into another random variable through a continuously decreasing or increasing transformation, then quantiles are also transformed into quantiles.

1.4.4 Two important distributions

- Normal distribution and confidence interval

The normal (or Gaussian) distribution and its mathematical expression are given in Fig. 1-26. The normal distribution is often used to represent the pdf of porosity or of random errors. We will see below, when discussing the central-limit theorem, why the normal distribution is so important in statistics. It is a symmetrical distribution (mean = mode = median), such that 95% of the values fall between $m - 2\sigma$ and $m + 2\sigma$.

If we make a random draw from a Gaussian pdf, we have 95% confidence that this draw falls within this interval, also called the confidence interval. On the basis of Fig. 1-23, we can also say that the best estimate of a value from a normal distribution is the mean, and that the confidence interval around this estimate is plus or minus twice the standard deviation. When discussing the quantification of structural uncertainties (Section 6.2), we will see that this definition of confidence interval will help us define the normal distribution associated by an interpreter to his/her interpretation at a given location.



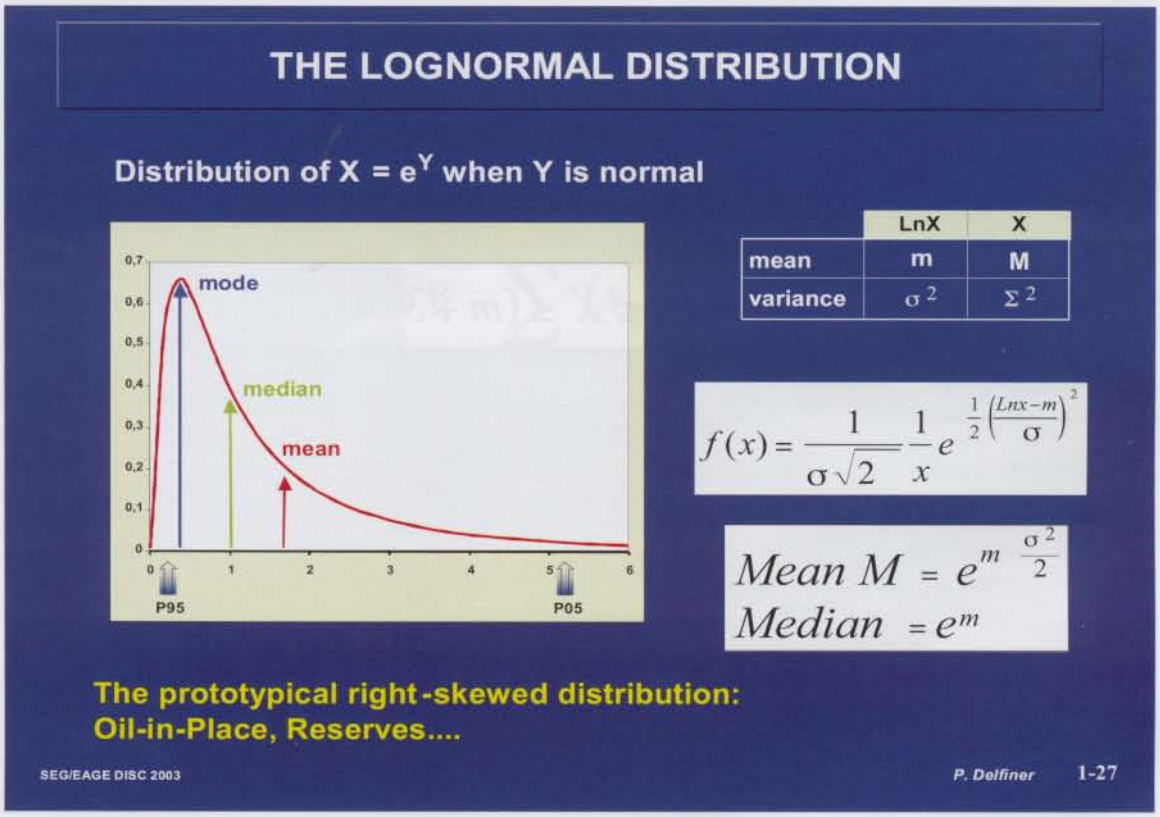
- Lognormal distribution

The lognormal distribution is closely related to the normal distribution. It is often used to model permeability. It is important to understand the relationships between the mean and standard deviation of a lognormal distribution and the mean and variance of its associated normal distribution. We saw earlier that all quantiles must be preserved through the exponential transformation. This means that the median of the lognormal distribution must be transformed into the median of the normal distribution, which is also its mean. On the other hand, the mean of the lognormal distribution is related to the mean of the normal distribution through a relationship that incorporates the variance of the normal distribution (Fig. 1-27). This implies that any mean-preserving statistical operation in the normal distribution will usually NOT be transformed into a mean-preserving statistical operation in the lognormal domain. A simple consequence of this, often ignored in practice, is that an unbiased regression model calculated between the logarithm of permeability and porosity will not simply translate into an unbiased relationship for predicting permeability itself.

1.4.5 Two important theorems

- Generalizing the confidence interval

The inequality given in Fig. 1-28 generalizes the notion of the confidence interval introduced above for the normal distribution. It shows that the knowledge of the mean and standard deviation of any continuous and unimodal distribution may result in a 95%



confidence interval for this variable. As stressed by Chilès and Delfiner (1999), the penalty for not knowing the distribution is an interval of width 6σ instead of 4σ. Later, this will help us better interpret kriging, which consists of calculating mean and standard deviation maps for a spatial parameter.

- Central limit theorem

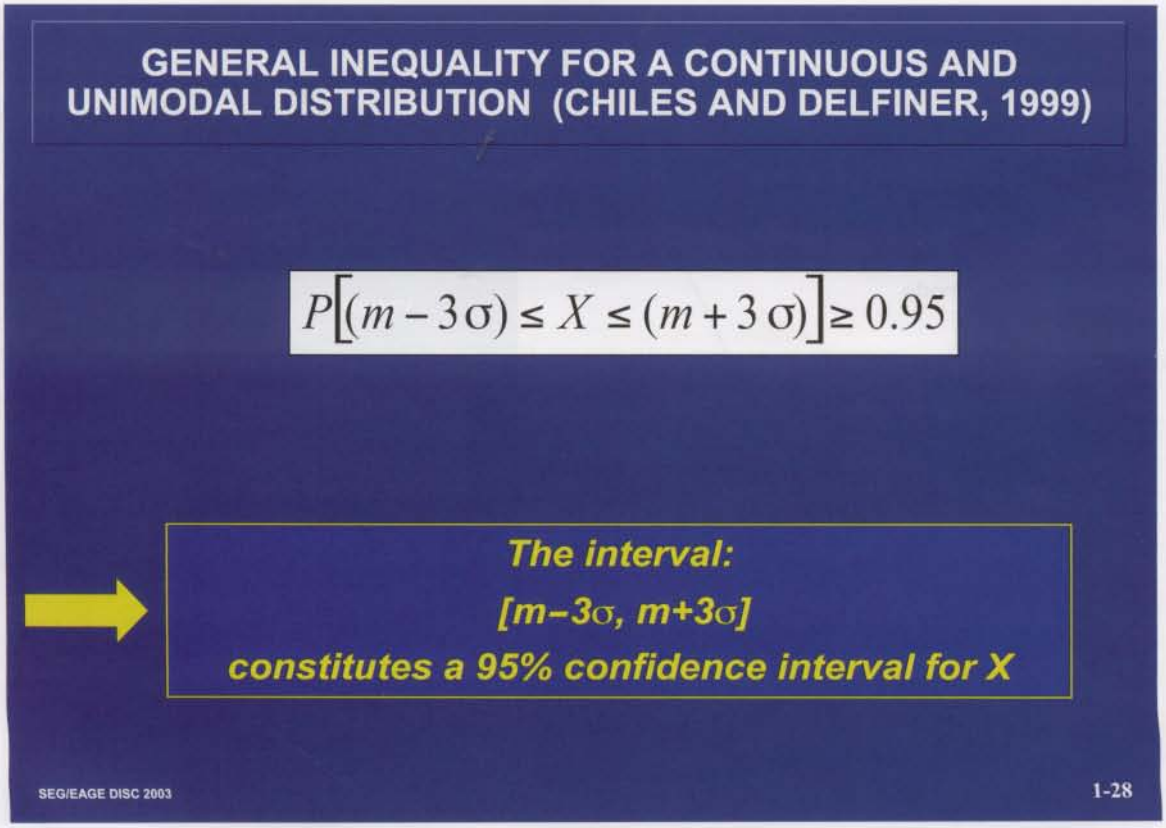
The central limit theorem (Fig. 1-29) must also be mentioned at this stage. It justifies the importance of the normal and lognormal distributions, respectively, as limits of sums and products of random variables. Because many of the geostatistical conditional simulations discussed later are obtained through the sum of a number of independent random variables, the result of these conditional simulations will tend to be distributed normally.

1.5 Basics of Bivariate Statistics

During this course, we will see the importance of using bivariate relationships, usually between a seismic attribute and a parameter measured at wells. Such relationships will be crucial when we are predicting the parameter of interest away from the wells. Thus, some time needs to be spent discussing bivariate relationships.

1.5.1 Covariance and correlation coefficient

The basic tool to measure the relationship between two random variables is the covari-



ance (Fig. 1-30). The development of the variance of the sum of two random variables as a function of the covariance is a very important relationship used to derive the kriging equations. To remove the effect of the variance of the two variables, for instance in situations where the two variances are of different orders of magnitude, the correlation coefficient is preferred, because it normalizes the covariance by the two variances.

The correlation coefficient measures the degree of linear relationships between two parameters, X and Y. In practice, it is calculated using the formula given in Fig. 1-31. Fig. 1-32 gives an example of a correlation coefficient between core-porosity and porosity derived from logs. Increases in the absolute value of the correlation coefficient correspond to convergence of the cluster toward a line (Fig. 1-33). Nevertheless, spurious effects can be caused when the data cluster is strongly affected by outliers, that is, by pairs of points that are clearly inhomogeneous with the rest of the cluster (Fig. 1-34).

1.5.2 Fitting a regression line

The regression line of Y against X is the line corresponding to the linear transformation of X that best predicts Y. In order to obtain it, we minimize the sum of squared differences between actual and predicted values. The value of the slope is a function of the correlation coefficient and the two standard deviations, whereas the intercept is simply calculated by forcing the line to go through the point associated with the mean of the two variables (Fig. 1-35).

CENTRAL LIMIT THEOREM FOR SUM AND PRODUCT

N independent random variables X_i of means m_i & variances σ_i^2

Define:

$$X = \sum_{i=1,N} (X_i - m_i)$$

X has mean zero and variance:

$$\Sigma^2 = \sum_{i=1,N} \sigma_i^2$$

X/Σ converges towards a normal distribution N(0,1)

SEG/EAGE DISC 2003

1-29

COVARIANCE

$$Cov(X,Y) = E\{[X - E(X)][Y - E(Y)]\}$$

Fundamental properties

Var(X+Y) = Var(X) + Var(Y) + 2 Cov(X,Y)

Cov(X,Y) = 0 if X and Y are uncorrelated

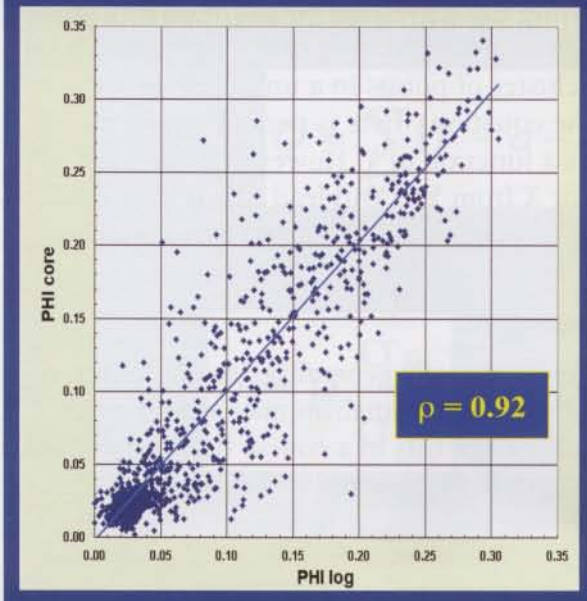
The basis for all the kriging maths!

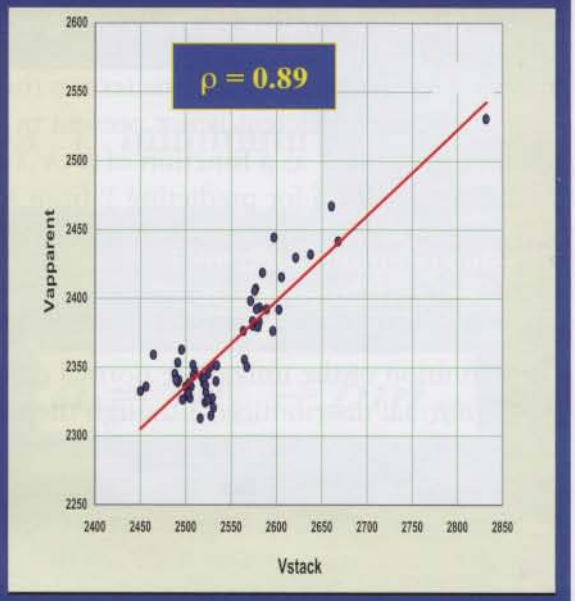
CORRELATION COEFFICIENT DEFINITION

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

$-1 \le \rho \le 1$	follows from a general mathematical inequality (Schwarz)
ρ	does not depend on the means nor on the scales of X and Y
$\rho = 0$	no linear relationship between X and Y
$\rho = 1$	perfect positive linear relationship between X and Y
$\rho = -1$	perfect negative linear relationship between X and Y

CORRELATION EXAMPLES

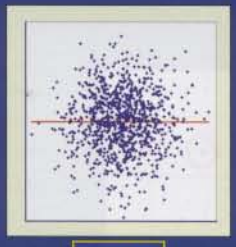




SEG/EAGE DISC 2003

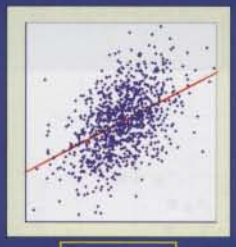
P. Delfiner 1-32

DIFFERENT CORRELATION COEFFICIENT VALUES

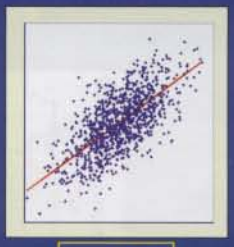


ρ = 0

ρ = 0.3



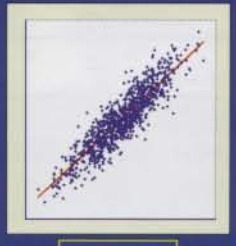
ρ = 0.5



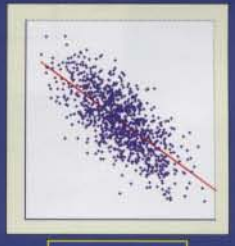
ρ = 0.7

As ρ increases

- the slope of the regression line increases
- the dispersion around the regression line decreases



ρ = 0.9



ρ = - 0.7

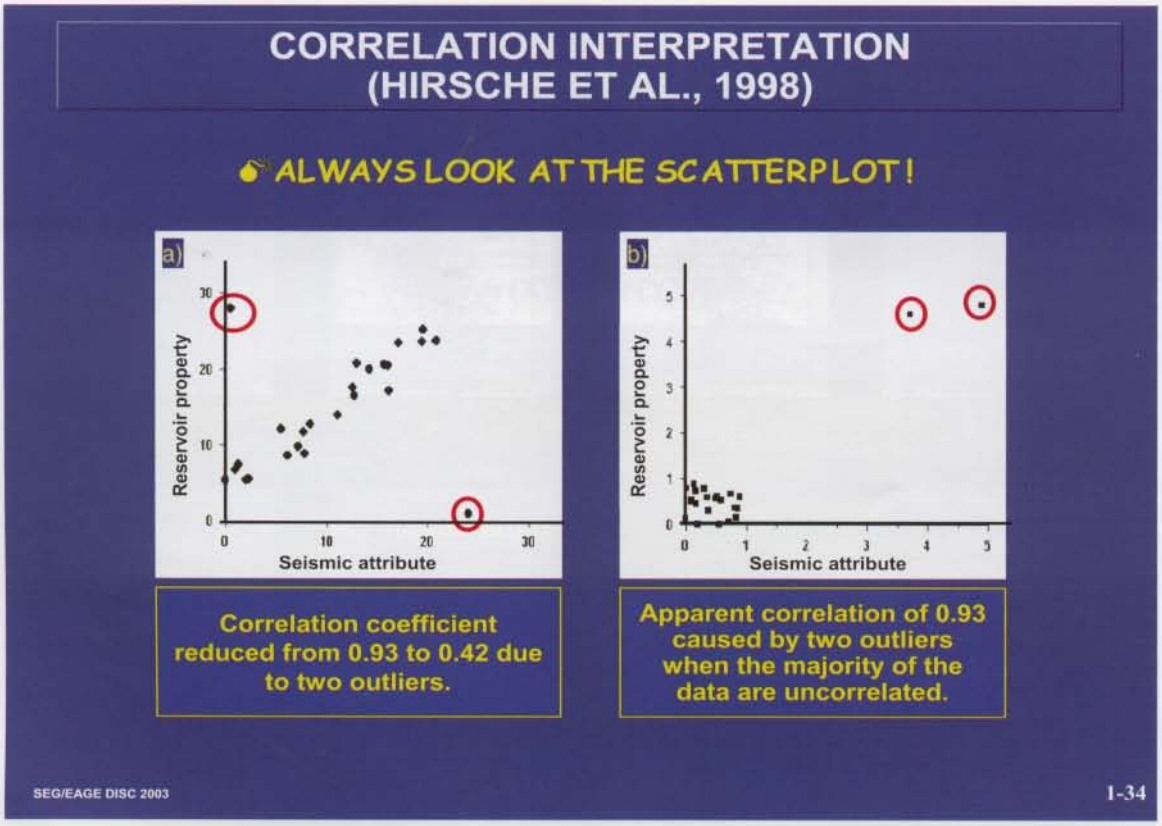
SEG/EAGE DISC 2003

P. Delfiner 1-33

Sometimes, people have difficulty understanding why there are two different regression lines, one of Y against X and one of X against Y (Fig. 1-36). There are two lines because the two lines are calculated using different criteria, respectively that of the minimization of vertical and horizontal differences. As a result, the regression of X against Y cannot be obtained simply by inverting the regression equation of Y against X. This is explained on Fig. 1-37, where a third line is figured, that of the Reduced Major Axis (RMA). This last line characterizes the cluster of points in a unique way. Because the correlation coefficient is not present in the equation, there is perfect symmetry whether we write Y as a function of X or X as a function of Y. However, if the RMA is used as an equation for predicting Y from X or X from Y, it will lead to a poorer prediction in terms of the sum of squared differences between predicted and actual values.

1.6 The Multivariate Normal Distribution

The definition of the univariate normal distribution easily generalizes to that of multivariate normal distribution, although the mathematical formalism may appear more intimidating (Fig. 1-38). Why do we want to mention this in a course that should limit the theoretical aspects? Because, if we look carefully at the term in the exponential, we see that it is simply a quadratic form of the vector z where the inverse of the variance-covariance matrix is present. Later in the course, this formula will help us understand why energy-based and stochastic modeling approaches are closely related.



CONSTRUCTING THE REGRESSION LINE OF Y ON X

Find coefficients a and b minimizing the sum of squared errors

$$Q = \sum_{i=1}^{N} (Y_i - a - bX_i)^2 \text{ minimum}$$

Line goes through center of gravity of the cloud

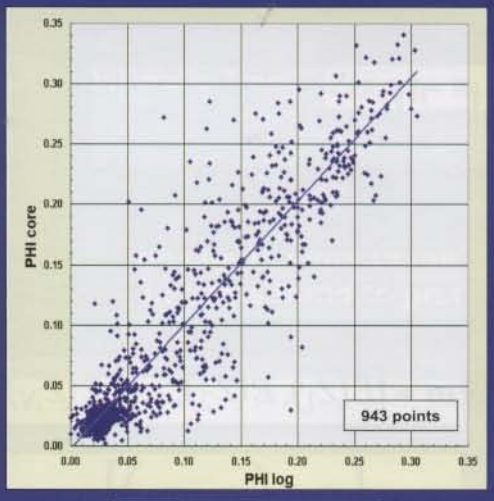
$$a = m_Y - bm_X$$

$$b = \rho \frac{\sigma_Y}{\sigma_X}$$

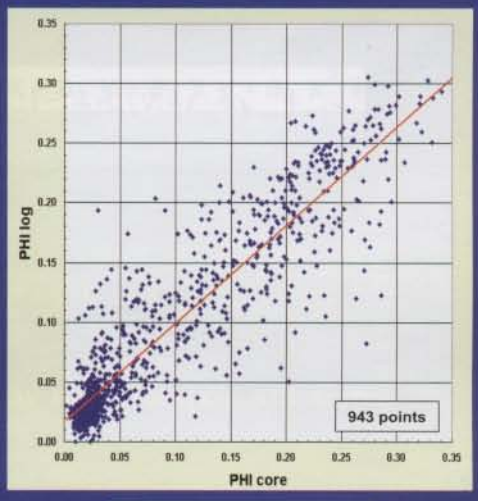
Correlation coefficient

Standard deviations of X and Y

PREDICT Y FROM X OR X FROM Y?



Y = 1.0235 X - 0.0017
R^2 = 0.8433

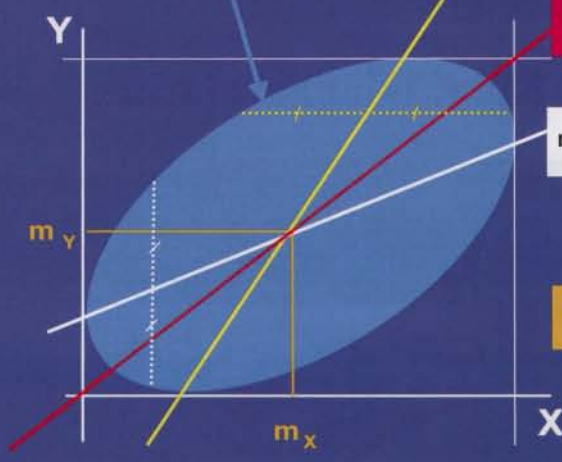


Y = 0.824 X + 0.0161
R^2 = 0.8433

Why this difference ?

BIVARIATE NORMAL DISTRIBUTION

Rough sketch of the scatterplot



$Y = m_Y + \frac{\sigma_Y}{\sigma_X}(X - m_X)$ (Reduced Major Axis)

$m_Y + \rho\frac{\sigma_Y}{\sigma_X}(X - m_X)$ (Regression of Y against X)

Lines go through center of gravity (m_x, m_y)
Product of slopes = ρ^2

$m_X + \rho\frac{\sigma_X}{\sigma_Y}(Y - m_Y)$ (Regression of X against Y)

SEG/EAGE DISC 2003 — P. Delfiner — 1-37

MULTIVARIATE NORMAL DISTRIBUTION

$$f(\mathbf{z}) = (2\pi)^{-N/2}|\Sigma|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{z}-\mathbf{m})'\Sigma^{-1}(\mathbf{z}-\mathbf{m})\right]$$

z is the vector $\mathbf{z} = (z_1, z_2, \ldots\ldots, z_N)$

m is the expectation vector $\mathbf{m} = [E(Z_1), E(Z_2), \ldots\ldots, E(Z_N)]$

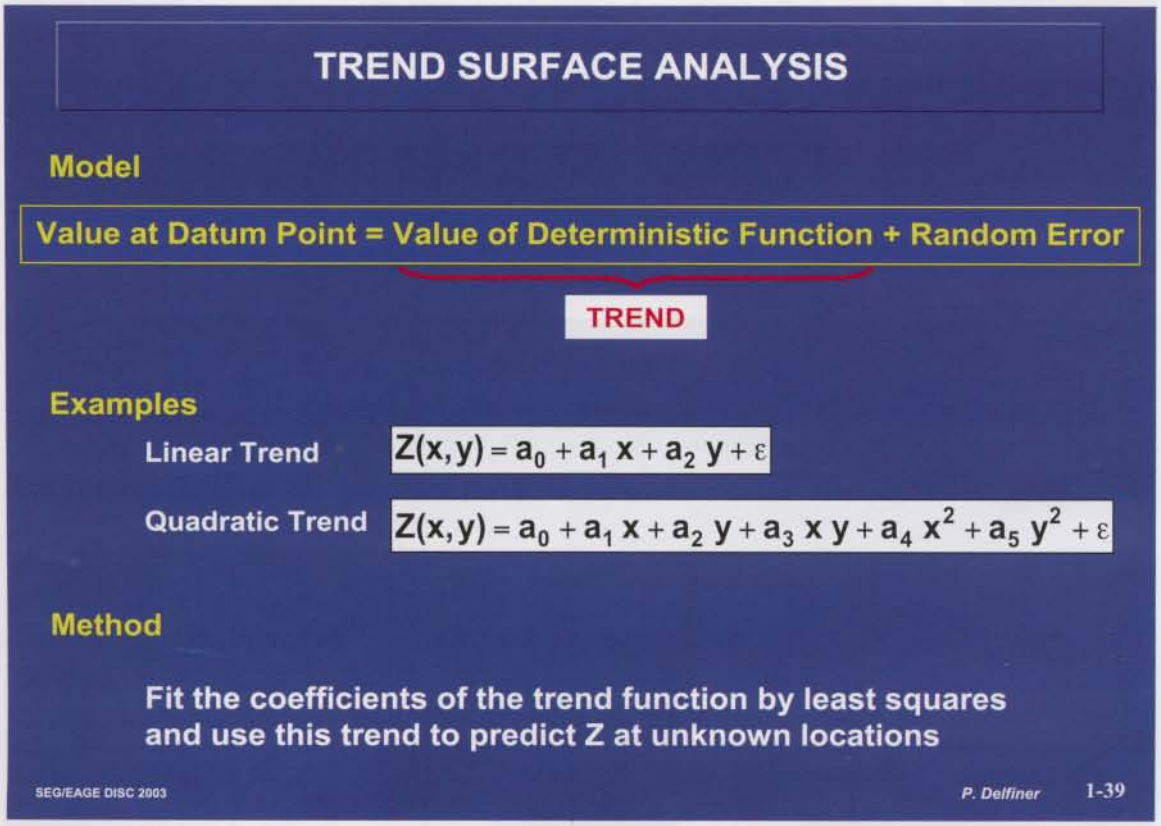
Σ is the variance-covariance matrix $\Sigma = [Cov(Z_i, Z_j)]_{i,j=1,N}$

SEG/EAGE DISC 2003 — 1-38

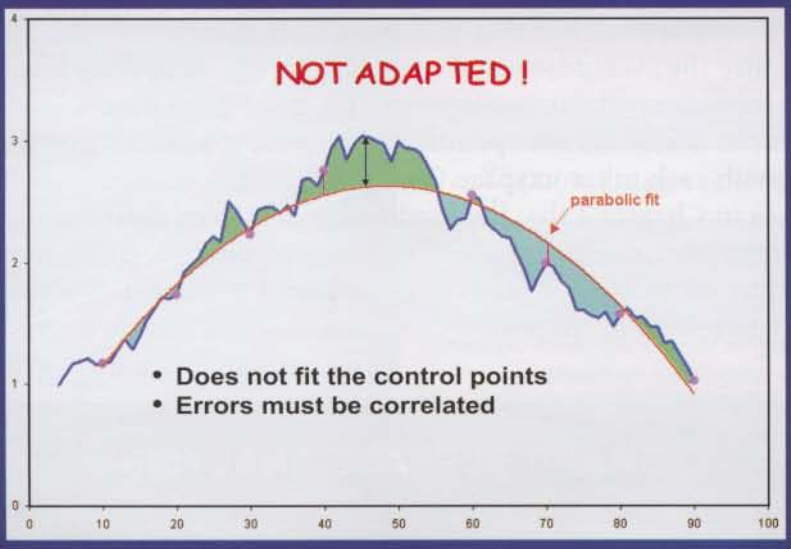
1.7 Trend Surface Analysis

Trend surface analysis follows the same formalism as does linear regression (1-39). The goal is to predict one variable, whether it is defined in one or two dimensions, using a linear combination of coordinates x and possibly x^2 in one dimension, or x, y, x^2, y^2 in two dimensions. The approach is that of least squares. Coefficients are calculated so that they minimize the average squared error. This is a good approach for fitting a linear or polynomial trend to spatial data. However, it is not a good interpolation technique, because it does not honor the data points, and because it assumes that the residuals are not correlated with each other in space (Fig. 1-40).

We will see in Chapter 2 that the model of geostatistics also assumes that the variable of interest can be decomposed as the sum of a polynomial trend and a residual. However, contrary to the trend surface analysis model, geostatistics assumes that the residual is also correlated in space.



TREND SURFACE ANALYSIS: TOO SIMPLISTIC FOR MAPPING



GEOSTATISTICS

SEG/EAGE DISC 2003 P. Delfiner 1-40

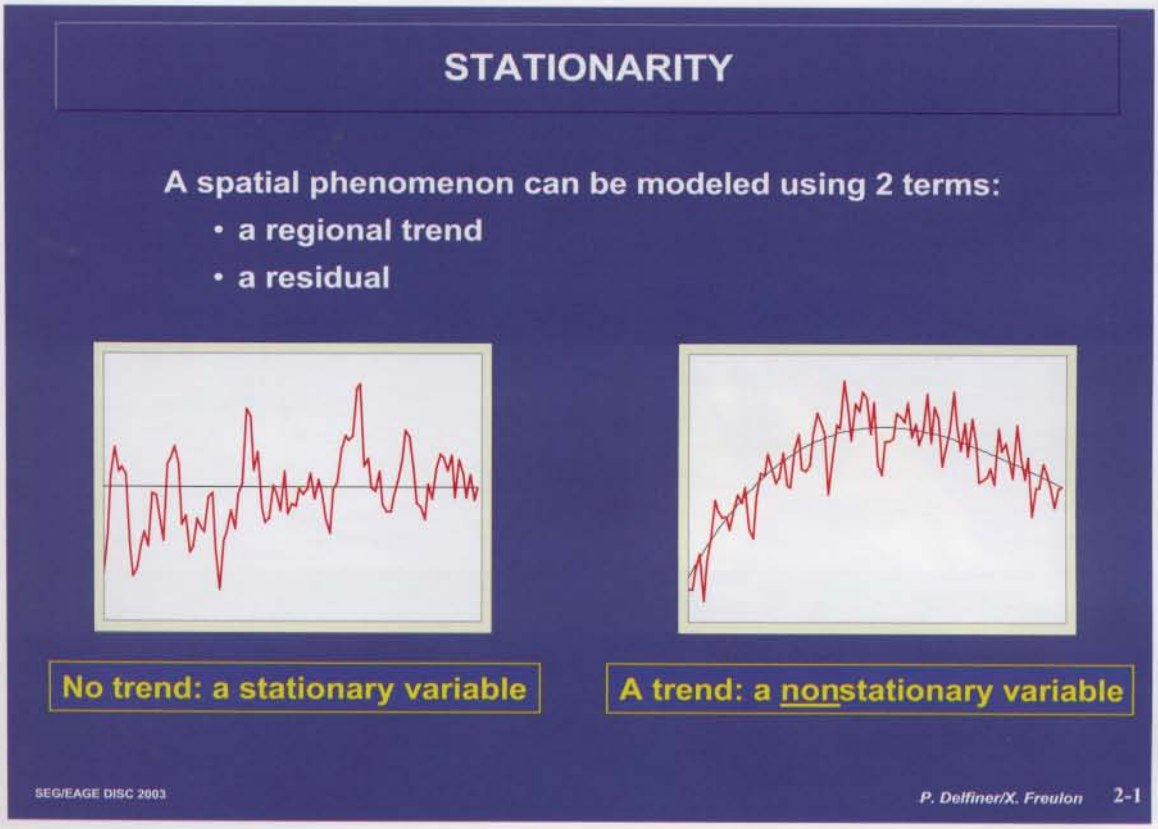
2 The Covariance and the Variogram

2.1 Stationarity Versus Nonstationarity

Nature often behaves in a very complicated way, and geology is no exception. In petroleum applications, where we are dealing with reservoirs at depths of several kilometers that are recognized by only a few wells and some seismic, we need to simplify the description of these reservoirs by means of models. A model is a simplification of nature and should never be identified with the natural phenomenon it seeks to describe. However, a model has the advantage of reducing our understanding of the reservoir to the estimation of a few parameters. The best approach is to explain the concepts using a 1D example.

Fig. 2-1 shows, on the left, a variable that varies around a constant mean. At any location, the behavior of the variable, although complicated, can be qualified as "homogeneously heterogeneous." On the average, it behaves the same everywhere, in the sense that we would make the same kind of error at any location if we were to predict the value of the variable from the value of the horizontal line. This will be discussed later as the stationarity hypothesis.

The picture on the right of Fig. 2-1 shows a different behavior. There is a systematic trend in the data that can be fitted using a parabolic model. The variable can be decomposed between a parabolic trend and a random pattern that varies around this trend. Thus, it can be modeled as the sum of a smooth polynomial trend plus a "sta-

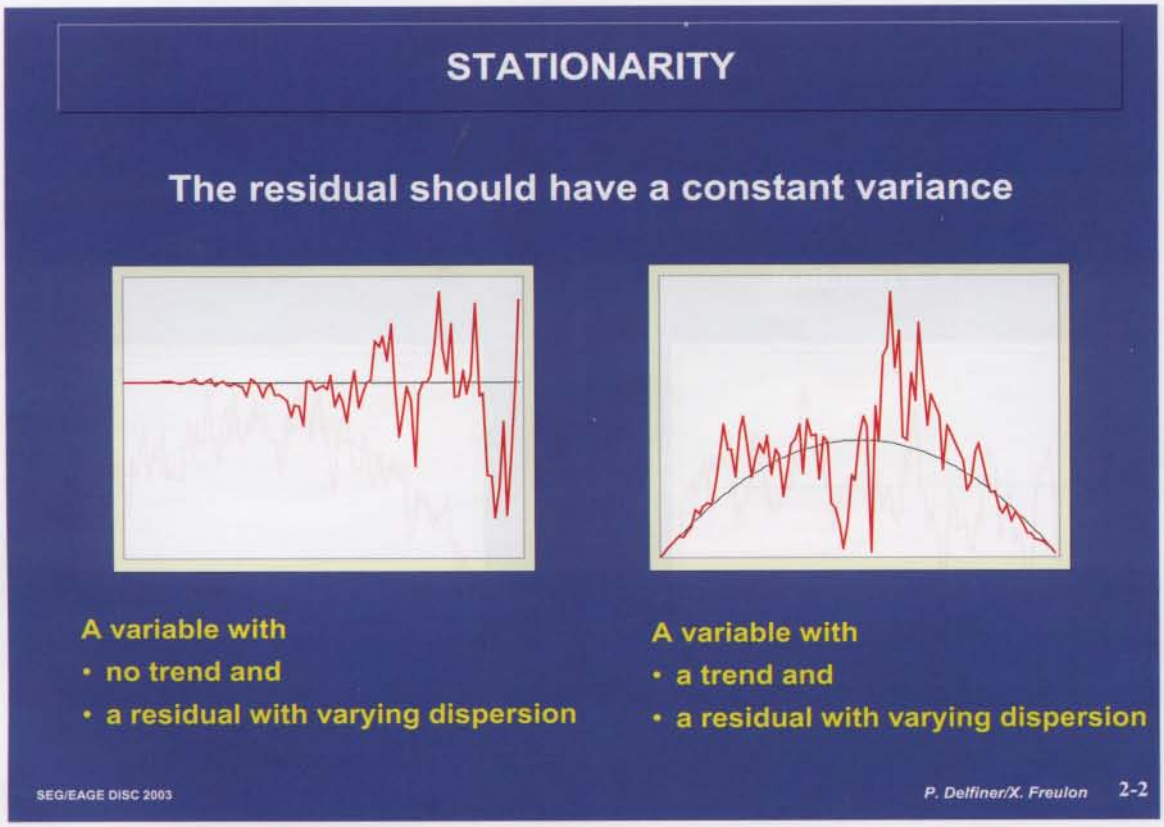


tionary" residual (to be better defined later). The model remains simple, in spite of the fact that it is a bit more complicated than that of stationarity.

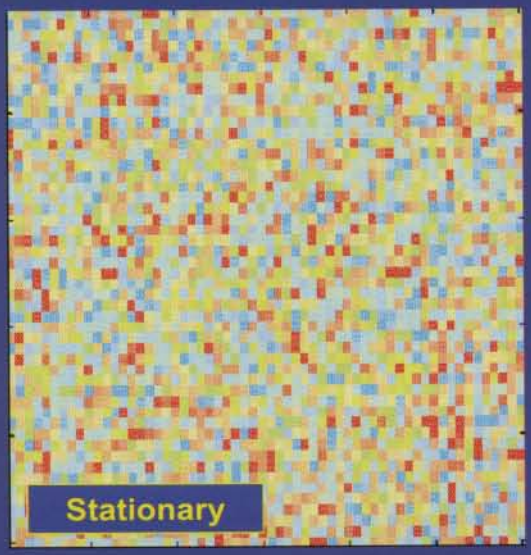
Fig. 2-2 shows variables whose behavior is harder to model. The curve on the left shows more scatter on the right side than on the left side. Similarly, the variable on the right of Fig. 2-2 cannot be modeled using a parabolic trend plus a stationary phenomenon. These two variables cannot be approached using the "trend + stationary residual" model.

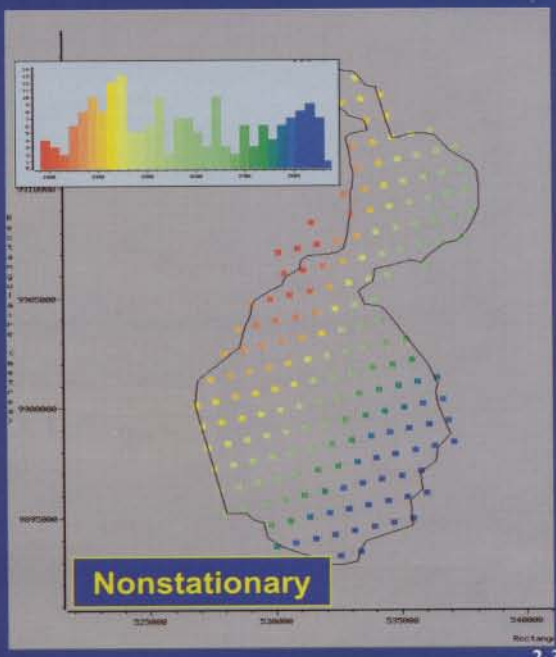
Fig. 2-3 generalizes the discussion from 1D to 2D. If we were to guess which color is present at any given pixel in the left square, we would probably give the same probability to each color: The statistical properties of the picture are independent of location, thus the phenomenon is stationary. On the other hand, there is clearly a systematic trend in the right picture (velocity data), which can be approached using the "polynomial trend + stationary residual" model. Fig. 2-4 shows two other examples of stationary and nonstationary surfaces. The surface on the right would probably lend itself to a "linear trend + stationary residual" model, which constitutes the geostatistical model.

Obviously, reducing a geological variable to such a model is a significant oversimplification. The beauty of the approach is that it will allow us to characterize the trend and residual using a few model parameters that we will try to estimate using the few data available. This is what we will study next. Before moving to the next paragraph, note that the choice of a model is always scale-dependent! In Fig. 2-5, if only the blue zone is of interest, a stationary model is perfectly suitable. Similarly, in the yellow area, a linear-trend model plus a residual will be satisfactory. In most applications, the



STATIONARITY

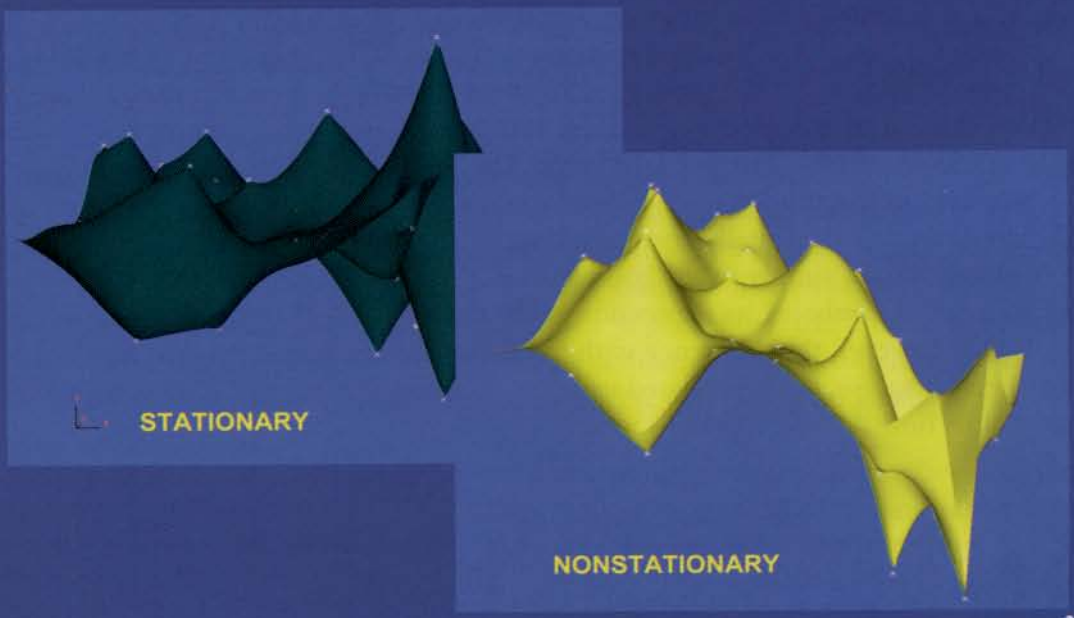




SEG/EAGE DISC 2003

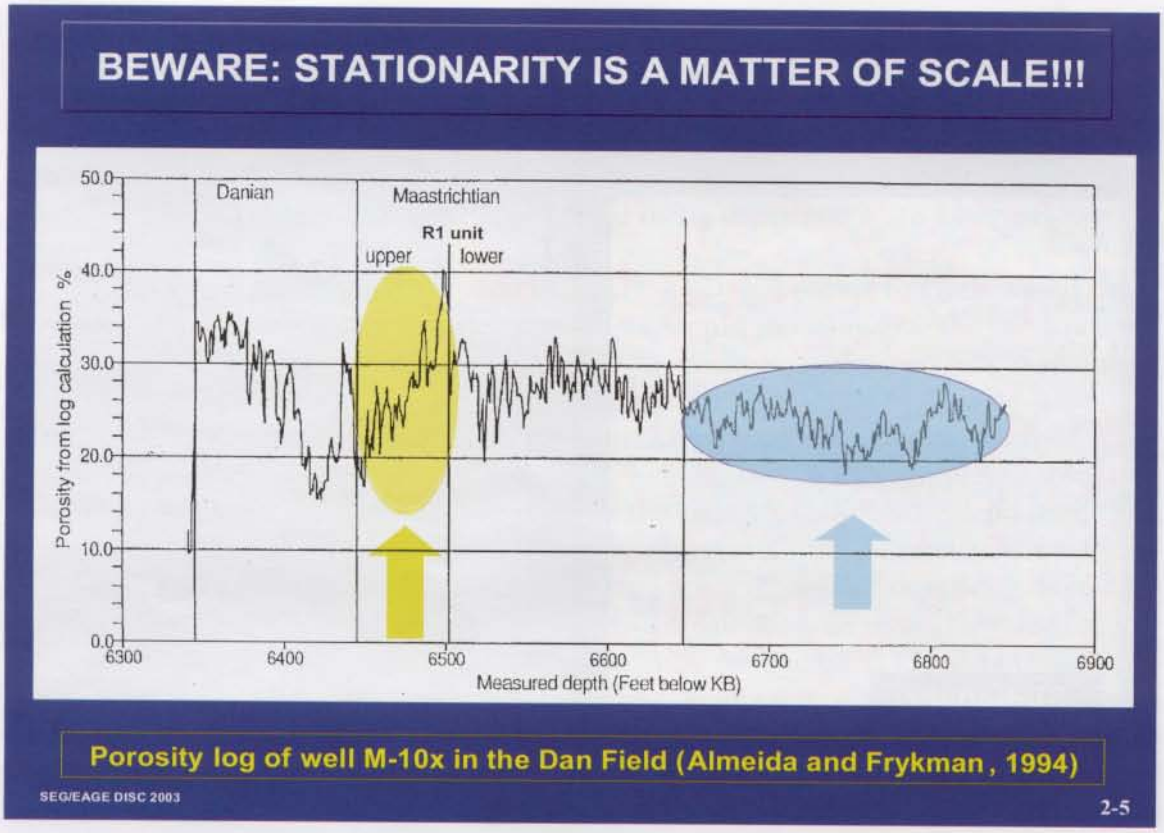
2-3

EXAMPLE OF STATIONARY AND NONSTATIONARY SURFACES



SEG/EAGE DISC 2003

2-4

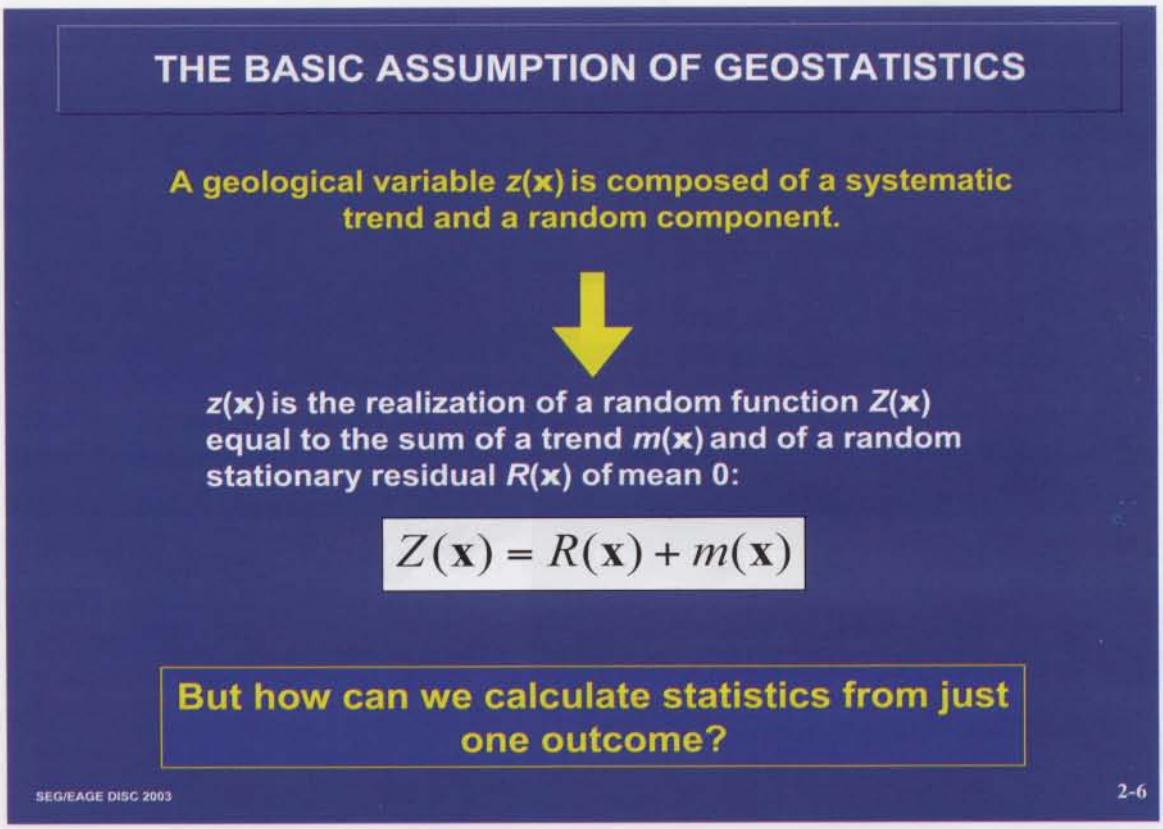


zone of interest — the well log here — is subdivided into subzones that are associated with different depositional environments (in Fig. 2-5 these zones are separated by vertical lines). Then, a different geostatistical model may be fitted to the different subzones. We will come back to this example later.

Fig. 2-6 summarizes the main mathematical assumption behind geostatistics. If $\mathbf{x}$ represents a point in 1D [$\mathbf{x}$ reduced to one coordinate (x)], 2D [$\mathbf{x}$ represented by two coordinates (x,y)] , or 3D [$\mathbf{x}$ represented by three coordinates (x,y,z)], the geological variable $Z(\mathbf{x})$ will be modeled as a random function — simply a mathematical object that is a random variable at every location $\mathbf{x}$. The geostatistical model will consist of modeling $Z(\mathbf{x})$ as the sum of a polynomial trend $m(\mathbf{x})$ and a residual random function $R(\mathbf{x})$. The trend $m(\mathbf{x})$ is usually constant, linear, or parabolic in the coordinates of $\mathbf{x}$. If it is constant, we are simply in the context of a stationary model of mean equal to m. If $m(\mathbf{x})$ is linear or parabolic, its expression is the same as that of trend surface analysis (Fig. 1-39). However, the difference with trend-surface analysis is that the residual of trend-surface analysis is uncorrelated in space, whereas the residual $R(\mathbf{x})$ is correlated in space. Let us now discuss what this means.

2.2 The Stationary Model

How can we calculate statistics from just one outcome of a random function? Indeed, there is little use in using a random-function model if we only are dealing with one real-



ization $z(\mathbf{x})$ of this function. It is the same as having to predict the number of balls of each color in a bag on the basis of drawing only one ball! However, the stationarity assumption will allow us to infer statistical properties of the random function $Z(\mathbf{x})$ on the basis of this single realization $z(\mathbf{x})$. The first assumption is that of a constant mean. This will imply that this mean can be inferred by averaging the measured values of $z(\mathbf{x})$ at different locations. The second assumption is that the variance of $Z(\mathbf{x})$ is also independent of location. This means that we are dealing with a parameter that is behaving somewhat as that of the left side of Fig. 2-3 and oscillating around a constant mean with an amplitude that is statistically the same everywhere.

This last property generalizes into the covariance property in the stationary situation — covariance between measurements at two locations depends only on the vector between these two locations. This means, as shown in Fig. 2-7, that we will be able to infer this covariance by combining pairs of points taken at different locations. Thus, thanks to the stationarity assumption, we are able to get around the limitation of having only one realization by calculating statistics that combine values at different locations. This would not be possible if mean and variance were dependent on location. The covariance $C(\mathbf{h})$ measures the spatial correlation.

In geostatistical practice, it is preferable to use the variogram tool rather than the covariance. The variogram is simply (half of) the variance of the increments (Fig. 2-8). The variogram is often preferred to the covariance, because it can be calculated directly from the data without needing to calculate the mean. The variogram is also more general than the covariance, because it only requires the stationarity of the increments of

THE STATIONARITY ASSUMPTION

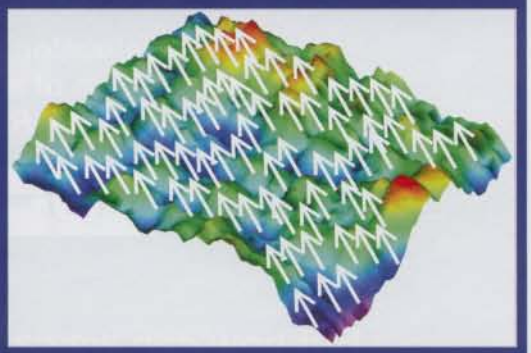
Mean and variance are independent of location:

$$E[Z(\mathbf{x})] = E[Z(\mathbf{x}+\mathbf{h})] = m$$

$$Var[Z(\mathbf{x})] = Var[Z(\mathbf{x}+\mathbf{h})] = \sigma^2$$

Covariance between values at two locations only depends on distance:

$$Cov[Z(\mathbf{x}), Z(\mathbf{x}+\mathbf{h})] = C(\mathbf{h})$$



SEG/EAGE DISC 2003

2-7

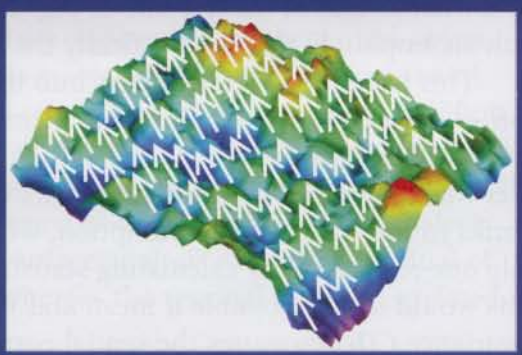
A MORE GENERAL ASSUMPTION THAN ORDER 2 STATIONARITY: THE STATIONARITY OF INCREMENTS

Increments are stationary

$$E[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = 0$$

$$\frac{1}{2}E\left[(Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x}))^2\right] = \gamma(\mathbf{h})$$

Variogram



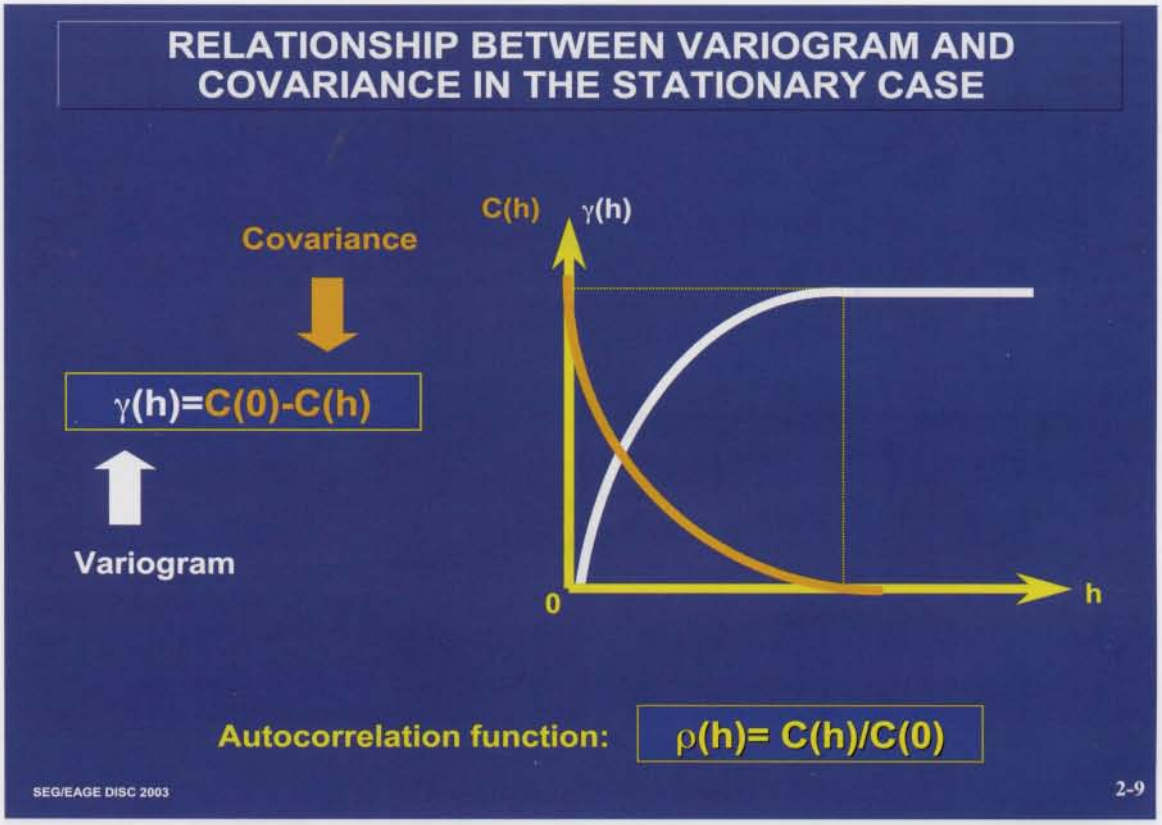
SEG/EAGE DISC 2003

2-8

$Z(\mathbf{x})$. We will see later that variables that have an unbounded variogram are not stationary, whereas their increments are stationary. This is the case for the power-law variogram model, which will be discussed later with relation to fractals. In the case of a stationary variable, the variogram flattens out at a certain distance, and we have a simple relationship between variogram and covariance (Fig. 2-9). In most cases, because covariance and variogram are even functions of distance, they are only plotted for positive distances. Fig. 2-9, as a reminder, also includes the relationship between covariance and autocorrelation function.

2.3 Calculation of a Variogram

Fig. 2-10 is an example showing how to calculate a variogram in a very simplistic 1D case. The approach consists of classing pairs of points by distances and then calculating the mean squared difference between pairs corresponding to each distance. When data are distributed in two dimensions, the isotropic (independent of direction) variogram is calculated in a similar way (Fig. 2-11). For each possible pair of data, the difference between values measured at the two locations can be plotted as a function of distance. This constitutes the "variogram cloud," which is a useful quality-check tool that shows which pairs and hence which data points seem to be outliers of the distribution. Differences corresponding to pairs associated with different bins of distances are then averaged. Fig. 2-12 is an example of filtered migration velocity data from West Africa. The isotropic variogram corresponding to these data is shown in Fig. 2-13 (consider



SIMPLE EXAMPLE OF 1D VARIOGRAM

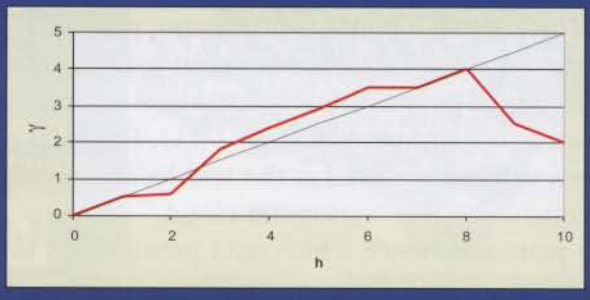
Data

x	0	1	2	3	4	5	6	7	8	9	10
z	0	-1	-2	-1	-	1	2	1	2	1	2

Experimental variogram

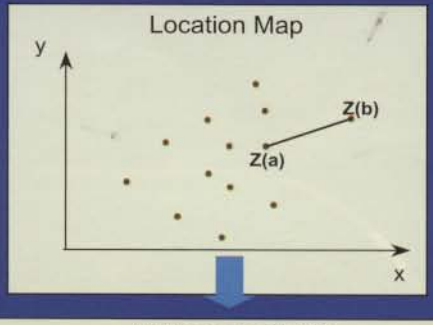
$$\gamma(h) = \frac{1}{2N(h)} \sum_{i} (z_i - z_{i+h})^2$$

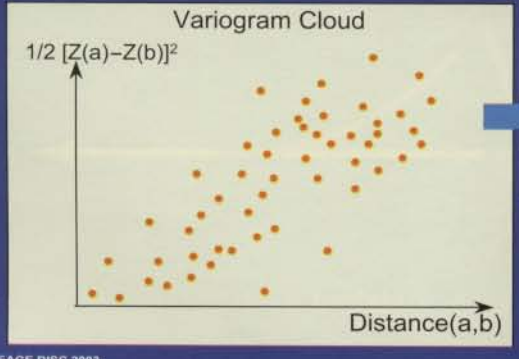
h	N(h)	γ(h)
1	8	0,50
2	7	0,57
3	6	1,83
4	5	2,40
5	5	2,90
6	4	3,50
7	4	3,50
8	3	4,00
9	2	2,50
10	1	2,00

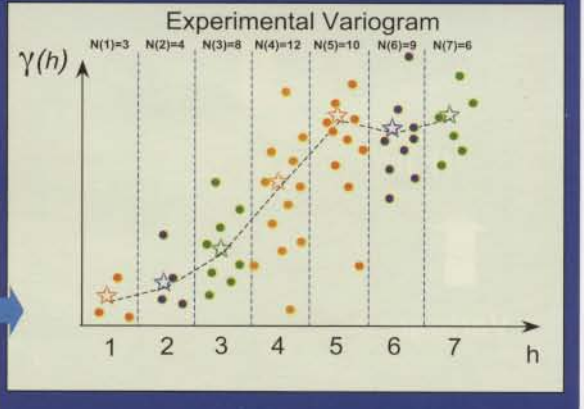


SEG/EAGE DISC 2003 — P. Delfiner/X. Freulon 2-10

HOW TO CALCULATE A 2D ISOTROPIC VARIOGRAM

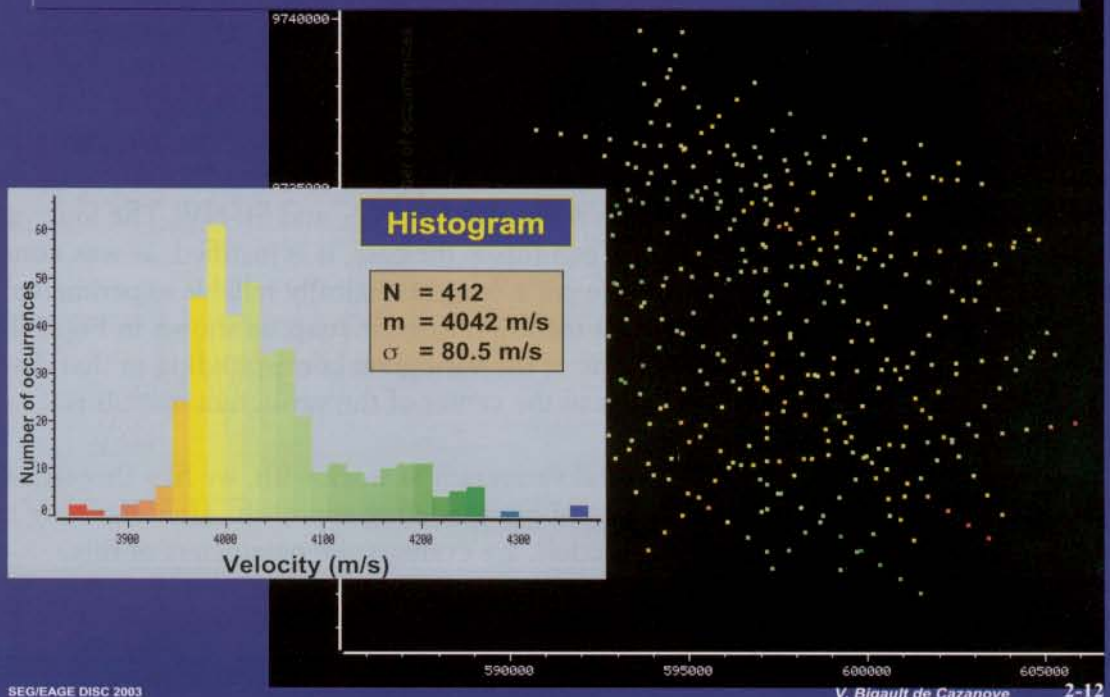






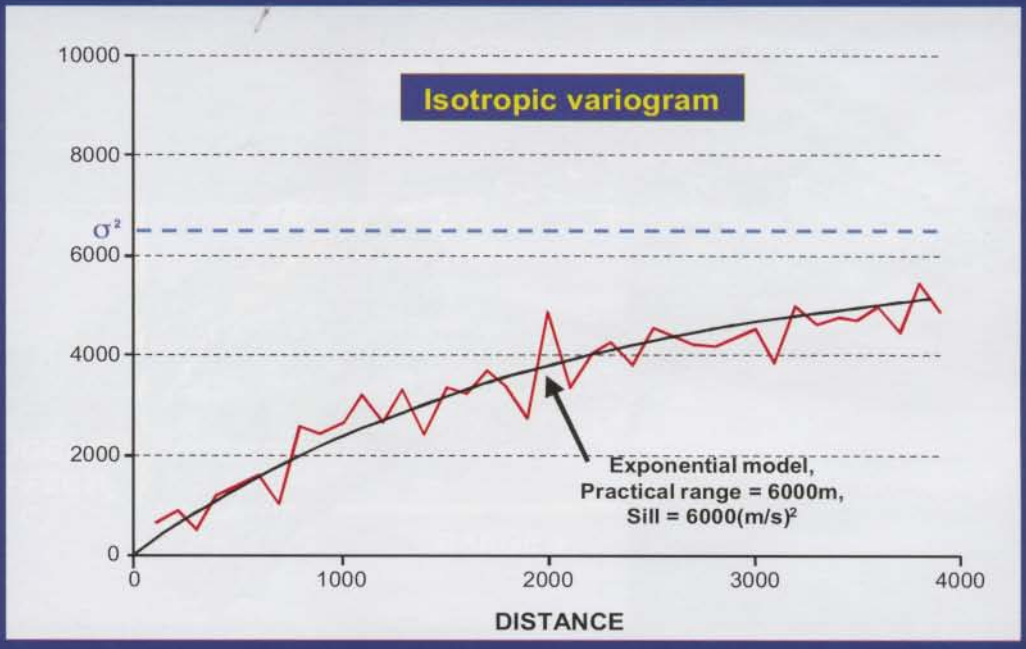
SEG/EAGE DISC 2003 — J.M. Guéméné 2-11

VELOCITY DATA EXAMPLE (1)



SEG/EAGE DISC 2003 V. Bigault de Cazanove 2-12

VELOCITY DATA EXAMPLE (2)

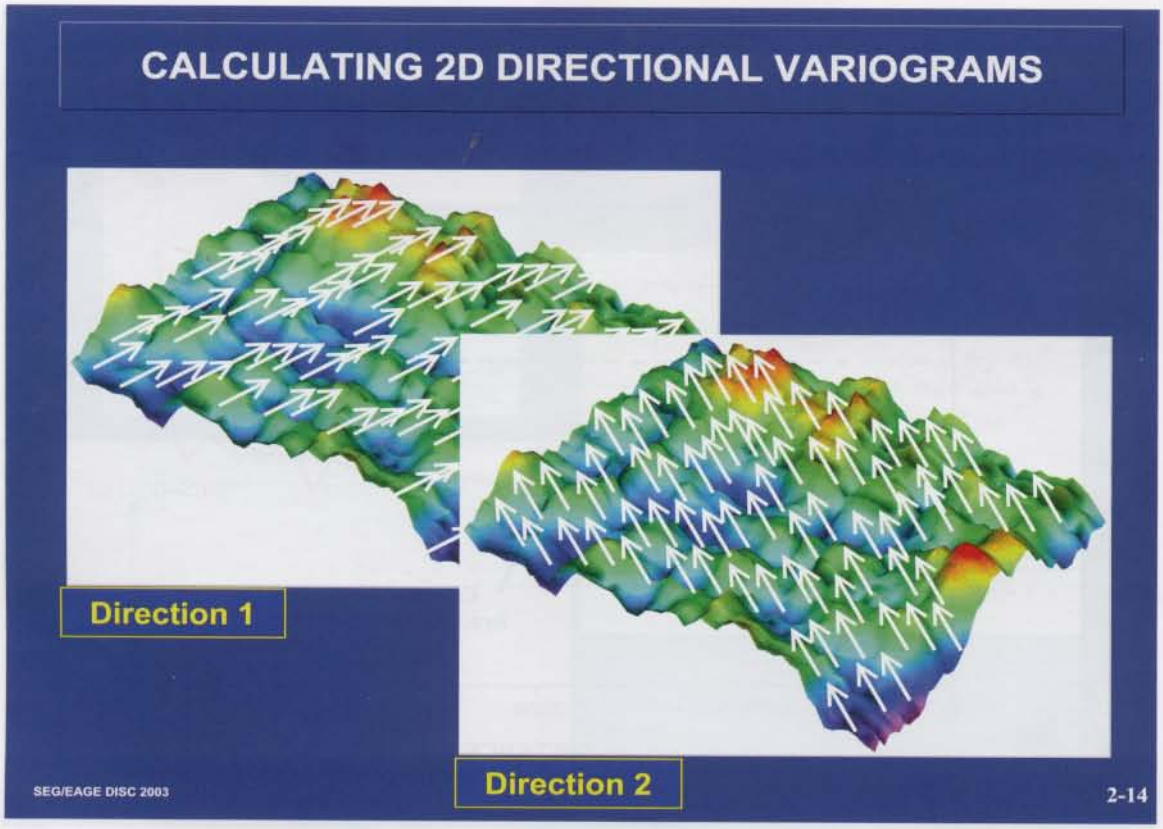


SEG/EAGE DISC 2003 V. Bigault de Cazanove 2-13

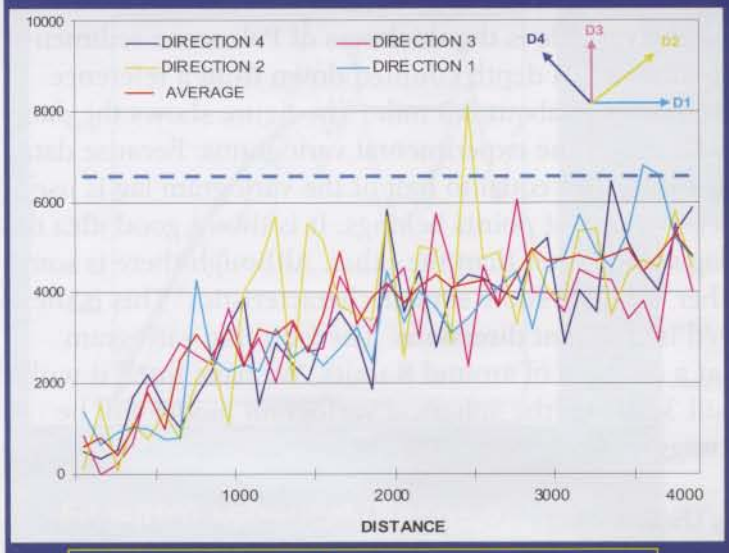
only the broken line at this stage). It shows, as expected, a slow increase as a function of distance. Because these data have been filtered, the variogram behavior at the origin is smooth, indicating that random noise has been removed.

Note that, in the above calculations, directions associated with each pair of points are not taken into account. In many cases variability changes with direction, which justifies the calculation of a variogram in different directions, as illustrated in Fig. 2-14. The difference with the isotropic variogram is that pairs are this time also binned according to their direction. In the velocity-data example, as in many 2D cases, pairs are classed according to four main directions: E-W, SW-NE, N-S, and SE-NW. The four variograms (Fig. 2-15) look quite similar. When this is the case, it is justified, as was done in Fig. 2-13, to merge all the directions to get a more statistically reliable experimental variogram. It is also possible to construct the 2D variogram map, as shown in Fig. 2-15. Each point on this map provides the value of the variogram corresponding to that distance and direction calculated in relation to the center of the variogram map. It is remarkably isotropic in our example.

Once we have a reliable experimental variogram to work with, we fit a theoretical model to it. This constitutes the next step of our modeling approach. After choosing the geostatistical model (here a stationary model), we evaluate the parameters of this model. Fig. 2-16 summarizes the important features of a variogram, at least in the stationary case: the nugget effect (discontinuity at the origin), the range (distance at which it becomes flat), and the sill (value of the plateau). Fig. 2-13 (smooth line) shows the result of the model-fitting exercise on our experimental variogram. The fitted model is

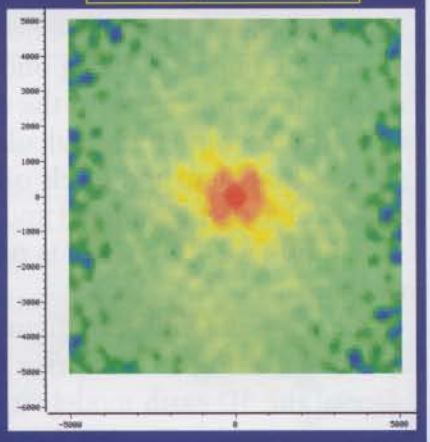


VELOCITY DATA EXAMPLE (3)

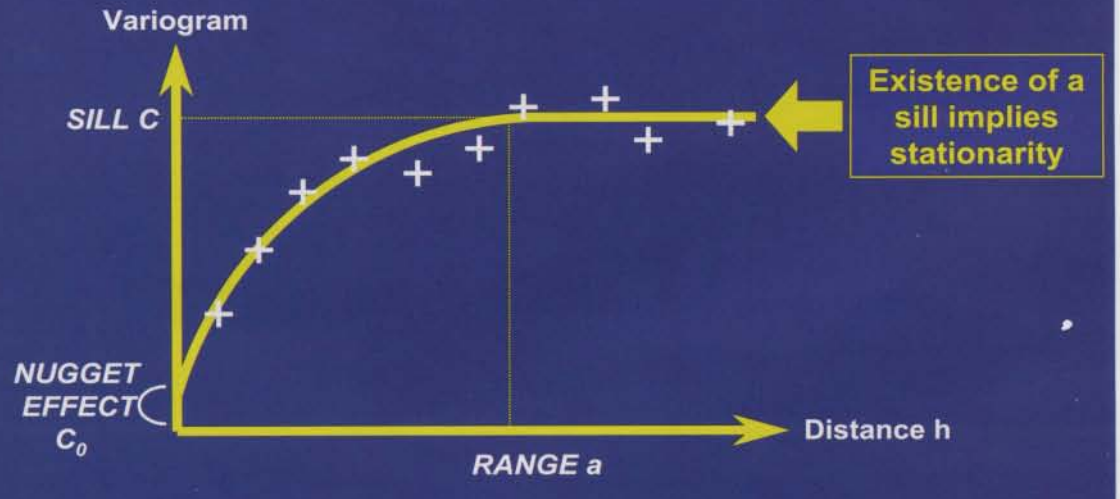


Variograms in four main directions

Variogram map



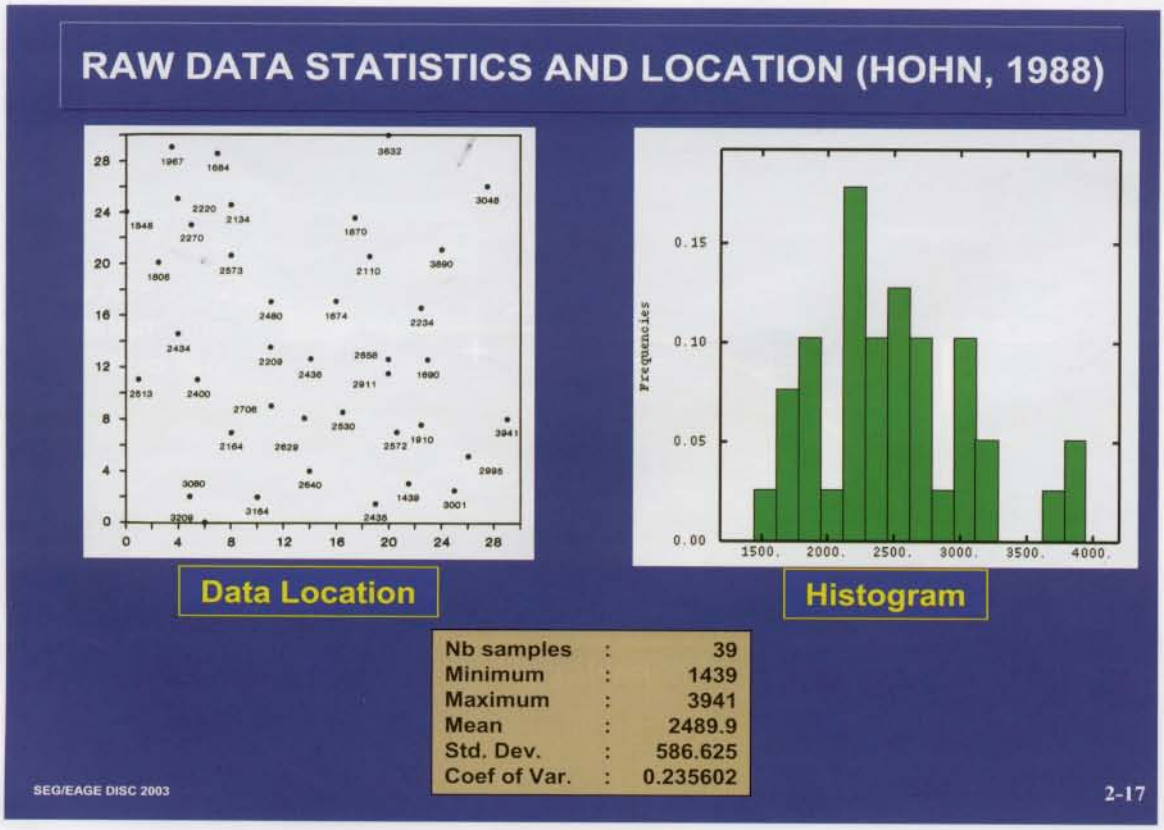
EXPERIMENTAL VARIOGRAM AND ITS FITTED MATHEMATICAL MODEL



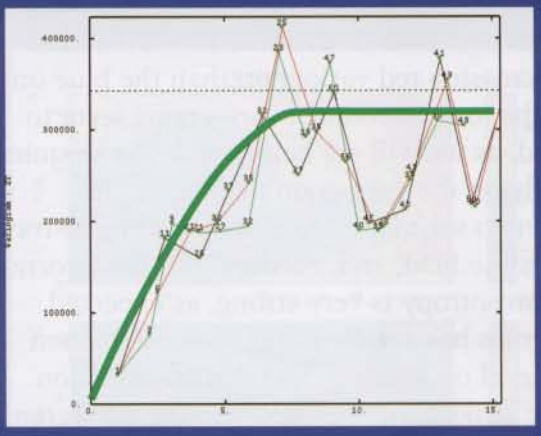
an exponential variogram with no nugget effect, a practical range of 6000 m, and a sill of 6000 $(m/s)^2$. We will explain later what an exponential variogram model is (Fig. 2-23).

Fig. 2-17 is a simple example taken from Table 2.1 of Hohn (1988), which we will use a few times during this course. The variable is the thickness of Paleocene sedimentary rocks in Libya. Here we will assume it is a depth counted down from a reference surface. A spacing of one unit corresponds to about 0.3 mile. The figure shows the data location and the histogram. Fig. 2-18 shows the experimental variograms. Because data are scarce and irregularly sampled, a tolerance equal to half of the variogram lag is used when deciding to which distance class a pair of points belongs. It is thus a good idea to measure the impact of using one lag value rather than the other. Although there is some variation from one curve to the other, all plots show similar characteristics. This is the same with the variograms calculated in different directions. The isotropic variogram increases from zero and levels off at a distance of around 8 units. We have fitted it with a spherical model of range 8 and sill 320,000 (the spherical variogram model will be defined in Fig. 2-23). There is no nugget effect. We will see below the implication of this choice.

Some of the following figures show various examples of experimental variogram models in 2D. Fig. 2-19, from a field in West Africa, shows a variogram calculated along a slice of the 3D earth model stratigraphic grid. It has been fitted using a slightly anisotropic model: The NE-SW direction has the smallest range, about .09, whereas the NW-SE direction has the largest range, about .12 (the side of the square that constitutes

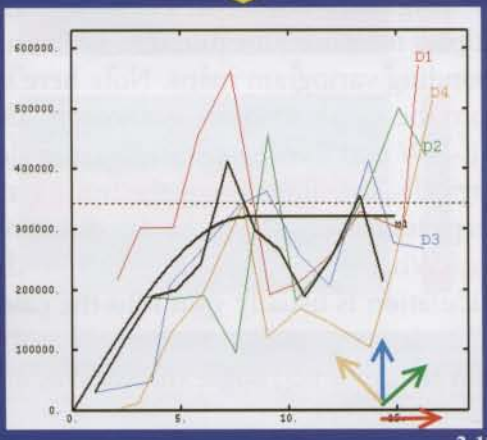


ISOTROPIC VARIOGRAM AND MODEL (HOHN DATA)



Isotropic model displayed with variograms for different lags

Isotropic model displayed with different directional variograms

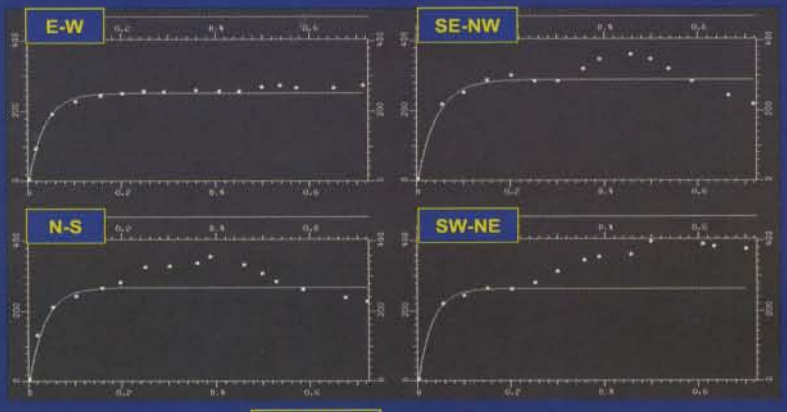


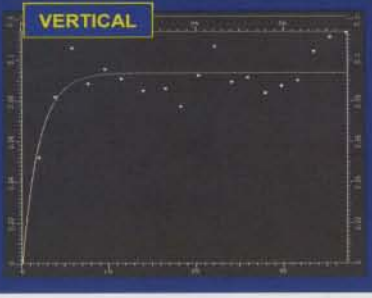
Fit with a spherical (Range 8):
$\gamma(h) = 320\,000\ \mathrm{sph}(8)$

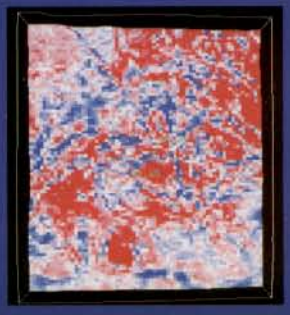
SEG/EAGE DISC 2003

2-18

2D AND VERTICAL VARIOGRAM EXAMPLE







HORIZONTAL AMPLITUDE SLICE

SEG/EAGE DISC 2003

L. Barens, TFEEUK Geoscience Research Centre

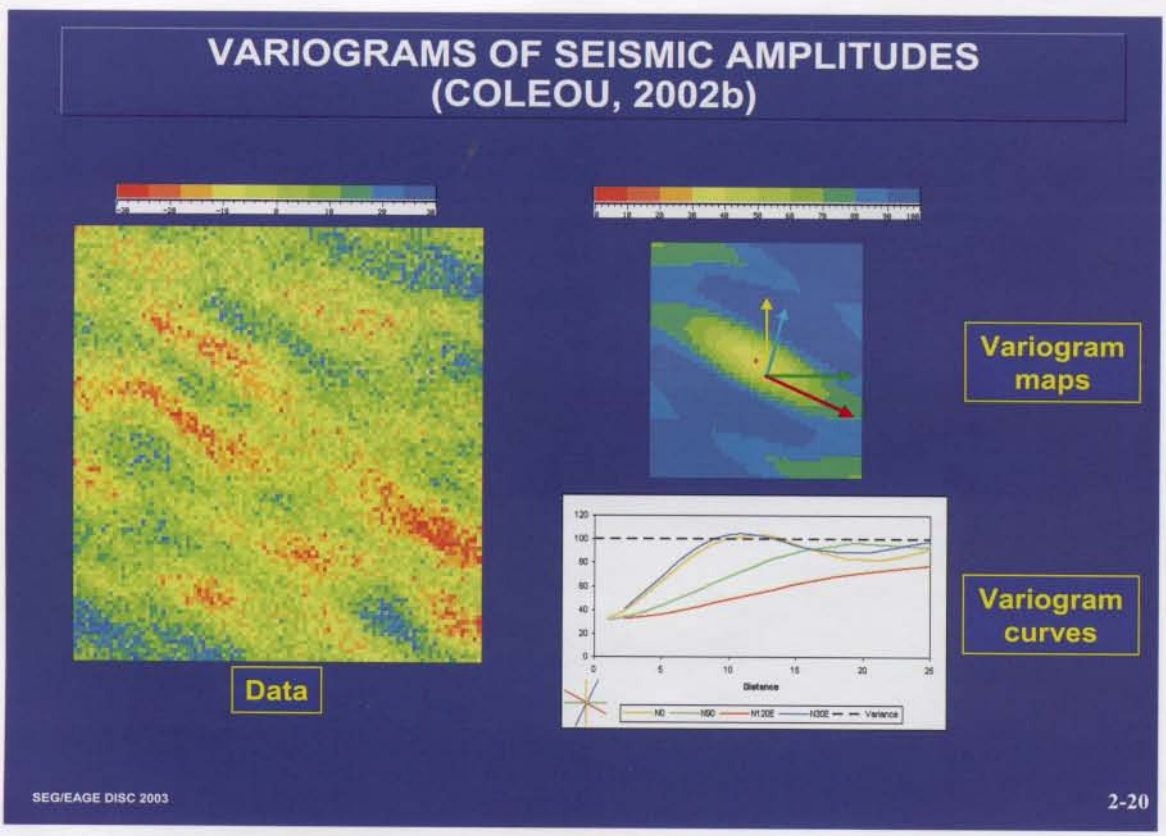
2-19

the area of interest is normalized to a distance of one). We have also displayed the vertical variogram calculated from the well data and the model that was fitted to it.

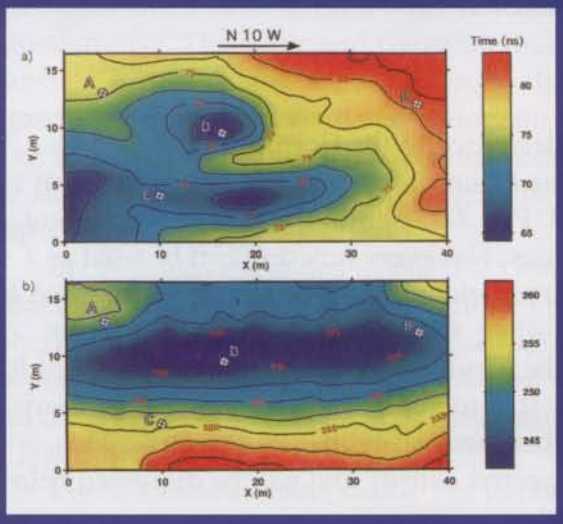
The anisotropy is stronger on another seismic example (Fig. 2-20). The variogram is that of seismic amplitudes along a horizontal slice of a 3D seismic block. There is clearly more continuity in the red (NW-SE) direction than in the blue (NE-SW) direction, which translates into a more gradually increasing red variogram than the blue one in the individual variogram curves. Note that the four directional variograms seem to have a similar nugget effect, which is associated, as we will see below, with the seismic noise. Again, the anisotropy appears very clearly on the variogram map.

Fig. 2-21 is another example, this time of two seismic maps corresponding to traveltimes measured for different horizons of the same field, and, on the right, their corresponding variogram maps. Note here that the anisotropy is very strong, as expected from the maps. Also, for both maps, the variogram has a finite range in one direction (SE-NW and E-W, respectively), but does not level off in the perpendicular direction. Logically, the direction of the finite range is the strike direction, whereas the variogram keeps increasing along the dip direction.

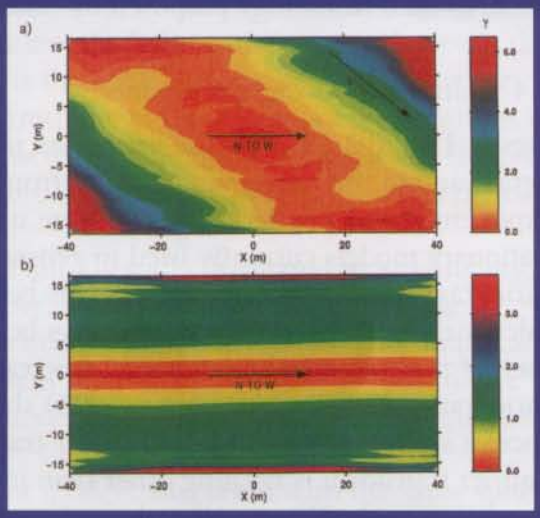
What if we are dealing with 3D data? In this situation, the experimental variogram calculation is usually split into the calculations of one 2D and one 1D variogram (Fig. 2-22). It is especially important that the 2D variogram is calculated in the (i,j,k) rather than in the (x,y,z) space, to avoid combining pairs of points that are not stratigraphically related. The chosen 1D direction is usually the vertical direction. Obviously, the vertical variogram shows variations over a much smaller distance than do horizontal vari-



VARIOGRAM MAP TO VISUALISE 2D VARIOGRAM (SZERBIAK ET AL., 2001)

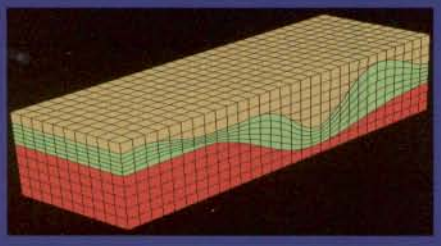


Two maps of two-way traveltime

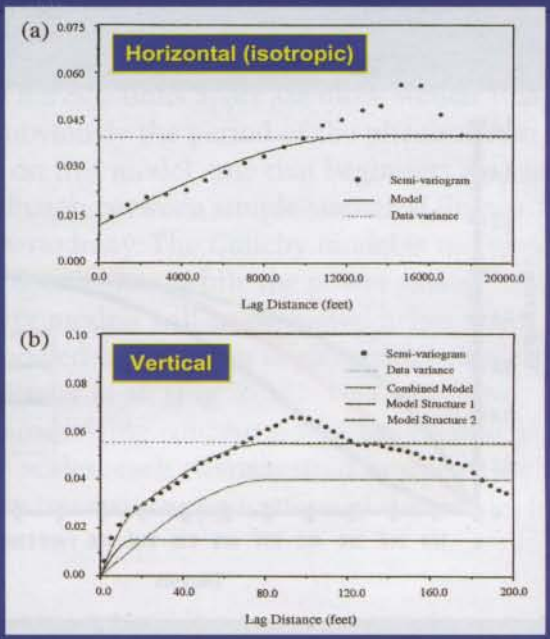


Their associated variogram maps

3D EXPERIMENTAL VARIOGRAMS (PAWAR ET AL, 2001)



In most cases, horizontal (2D) variograms and vertical (1D) variograms are calculated parallel to stratigraphy and along the vertical direction.



Shale fraction variogram

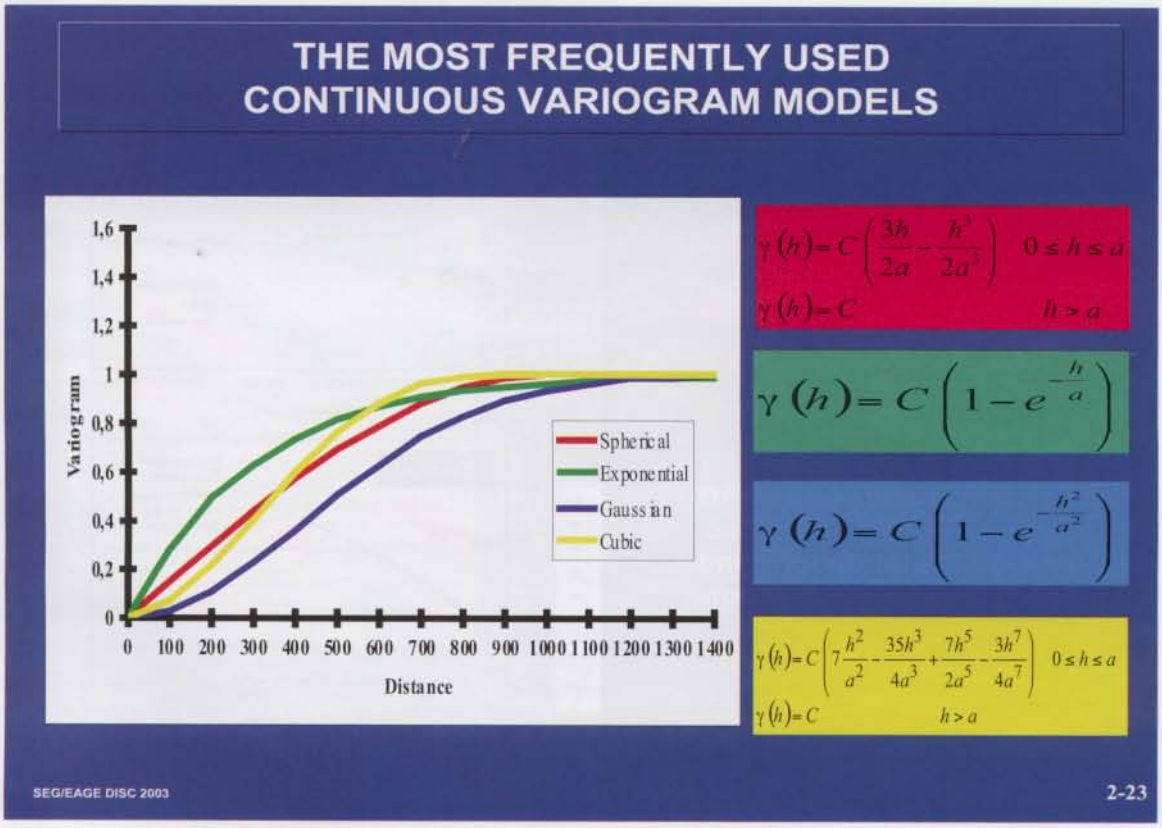
ograms, because geology changes more rapidly perpendicular to stratigraphy than it does along stratigraphy.

Note that Pawar et al. (2001) and other authors use the term "semi-variogram," because of the ½ factor in front of the variance of the increments. We prefer to stick to the original terminology proposed by Matheron.

2.4 Stationary Variogram Models

Figs. 2-13 and 2-18 show the variogram models fitted to the isotropic velocity variogram and to the variogram obtained from the Hohn data set. It was mentioned that an exponential and a spherical model were used. Fig. 2-23 shows the four main types of stationary models currently used in geostatistics. Not every function can be used as a variogram or covariance model. This is because variograms and covariances are used to calculate variances, which must always be positive. Covariances must belong to the class of positive definite functions, whereas the condition is somewhat more relaxed for variograms. Chilès and Delfiner (1999) show that to check that a function is a covariance, it suffices to calculate its Fourier transform and verify that it is positive. This Fourier transform is nothing other than its spectral density and will be discussed below in relation to the Wiener-Khinchin relationship.

The exponential and Gaussian variograms are bounded, but they only asymptotically reach the sill C. The practical range of the Gaussian and exponential variograms, that is, the distance at which they reach 95% of their sill, are $a\sqrt{3}$ and 3a, respectively.



As the range of the variogram becomes smaller and smaller, the covariance tends toward the "pure nugget effect" or "white noise" covariance. The spherical and the exponential models both have a linear behavior at the origin. However, the exponential model, for the same practical range, "climbs" faster than the spherical, which means that two measurements taken close to each other tend to differ more in the exponential than in the Gaussian case. The cubic and Gaussian models are parabolic at the origin. However, the Gaussian is much smoother at the origin than is the cubic, which means that, for the same practical range, a Gaussian variogram corresponds to much smoother variations of $Z(\mathbf{x})$. Fig. 2-24 illustrates what has just been discussed: the higher (respectively lower) the slope of the variogram at the origin, the more (respectively less) continuous the corresponding variable.

Because all variograms in Fig. 2-23 have a sill, they all correspond to a stationary model. But beware, because, as mentioned before, stationarity is a matter of scale: A spherical variogram of range a will correspond to a nonstationary variable for practical purposes, if this variable is studied over an area of interest of size $a/2$! The covariance functions associated with the four variogram models can be derived using the equation of Fig. 2-9.

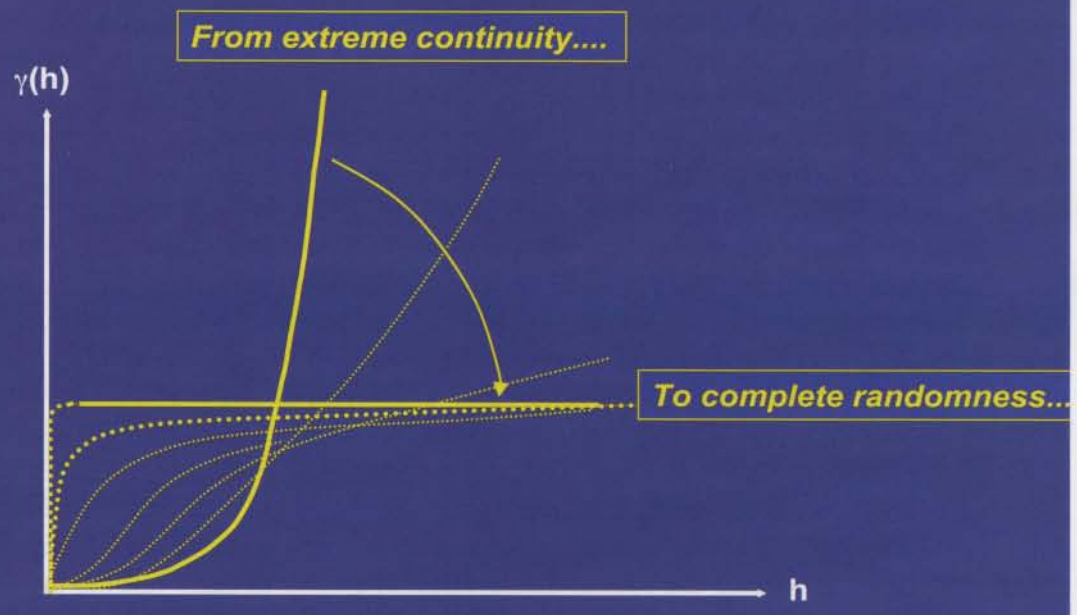
Fig. 2-25 shows more exotic but important models, displayed with the same spherical model as in Fig. 2-23. The "De Wijs" variogram has historical interest because it proved useful for modeling the spatial distribution of grades in early mining applications of geostatistics. It was introduced by Matheron (1962) to model the fact that, in some cases, the variance of grades within mining blocks would plot linearly on a log-log scale as a function of the block size. We will also see later that this model is closely related to fractals because, if the log of a parameter follows a De Wijs model, then the variogram of the parameter itself follows a power-law model (Agterberg, 1994). Matheron (1987) gives another interesting discussion of the relationship between the De Wijs model and fractals. The "hole-effect" model is characteristic of variables showing a strong periodicity, which is indicated by the decrease in the variogram around distance 800. This means that two measurements 800 units apart are more similar than two measurements 400 units apart. 800 is obviously the period of the phenomenon. However, we prefer not to dwell too much on this model, one that beginners may tend to use too often, because it can lead to confusion between simple statistical fluctuations of the experimental variogram and actual periodicity. The Cauchy model is not often used, except in the context of gravity or magnetic data, while the power variogram is related with fractals. Both Cauchy and power models will be discussed below.

An experimental variogram can be modeled using a sum of elementary models — see, for instance, the vertical variogram of Pawar et al. (Fig. 2-22). This will prove important later in this course, because the underlying concept is that the variable is the sum of a number of components of various scales, each characterized by one of the elementary models. Factorial kriging will show interesting applications of this model for filtering purposes.

2.5 Examples of Anisotropic Experimental Models

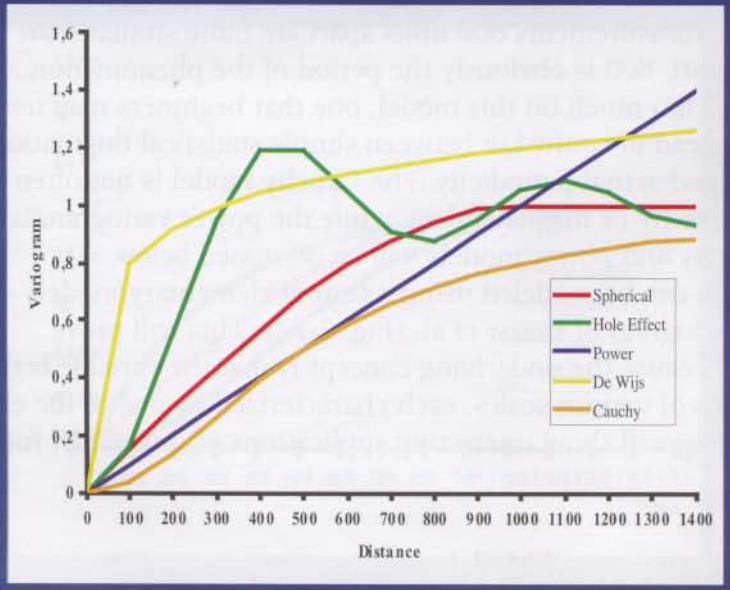
We have already seen examples of anisotropic variograms. Often, horizontal (by hori-

VARIOGRAM BEHAVIOUR AT THE ORIGIN



SEG/EAGE DISC 2003 V. Bigaut de Cazanove 2-24

OTHER CONTINUOUS VARIOGRAM MODELS



$$\gamma(h) = C\left(\frac{3h}{2a} - \frac{h^3}{2a^3}\right) \quad 0 \le h \le a$$
$$\gamma(h) = C \quad h > a$$

$$\gamma(h) = C\left(1 - \frac{a}{h}\sin\frac{h}{a}\right)$$

$$\gamma(h) = C\,h^{\alpha} \quad 0 < \alpha < 2$$

$$\gamma(h) = C\,Logh$$

$$\gamma(h) = C\left(1 - \frac{1}{1 + \frac{h^2}{a^2}}\right)$$

SEG/EAGE DISC 2003 2-25

zontal we mean parallel to stratigraphy, that is, in the (i,j) space) variograms are anisotropic, which can be easily explained. For structural parameters, such as time or depth to a horizon, the variogram will differ along the strike and dip direction. For parameters such as porosity or velocity, which are heavily controlled by sedimentology, features such as the paleocoast or the channel deposition direction will obviously have an impact on the variogram. In almost all applications, horizontal variations are modeled using a "geometrical anisotropy" model — the range of the variogram changes according to direction, but the sill is the same. This can be modeled using a simple rotation of the coordinates. The three examples of Figs. 2-19 to 2-21 show geometrically anisotropic variograms, with that of Fig. 2-19 showing the anisotropic model together with the experimental curve.

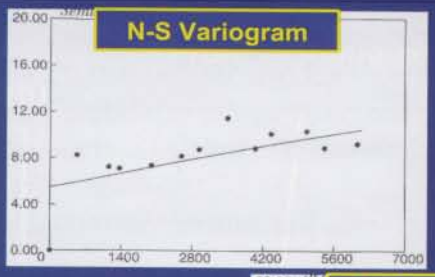
Things may be more difficult in 3D, when we deal with anisotropy between the vertical direction and the direction parallel to stratigraphy (call it horizontal for simplicity). Usually the vertical variogram is calculated over tens of meters, whereas the horizontal variogram is calculated over kilometers. Also, there are more high-frequency variations along the vertical than along the horizontal direction. Thus, the sill of the vertical variogram may be different from that of the horizontal variogram (see the example of Pawar et al. in Fig. 2-22). In this situation, the horizontal (whether isotropic or not) and the vertical variogram can be separately fitted by a different model, but then a merge between the two models must be performed to obtain a 3D model that will be computable for any direction in 3D, including directions that are not strictly vertical or horizontal. A zonal-anisotropy model is usually applied, in which the global 3D variogram is modeled as the sum of three different terms, the first term depending on the three coordinates, the second one equal to the horizontal model, and the third one equal to the vertical model. The use of the first term allows a smooth transition between the vertical and the horizontal variogram. Pawar et al. (Fig. 2-22) separately fitted the vertical and the horizontal (isotropic) model, but it is not clear how they derived the 3D variogram from these two models. Another possibility is that of modeling the covariance as the product of two models, one depending only on horizontal coordinates, the other depending only on vertical coordinates. This model was applied to the variogram of Fig. 2-19. In some situations, however, a geometrical anisotropy will still prove satisfactory for fitting a 3D experimental variogram. Exercise 1 at the end of this book discusses such an example taken from Chu et al. (1994) and presented in Fig. 2-26.

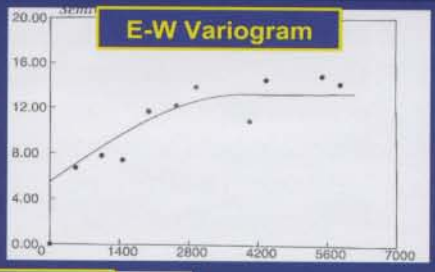
2.6 Unbounded Variogram Models and Their Relationship with Fractals

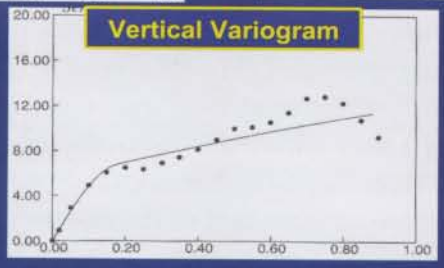
Fig. 2-25 shows two unbounded variogram models, the De Wijs and the power-law models. Both are interesting because they have close relationships with fractals. It can be demonstrated that, for a power-law model to be a valid variogram model, the exponent of h must be smaller than two. Power-law variograms correspond to variables that are not stationary but that have stationary increments. Structural variables, such as time or depth, often follow power-law models. Interestingly, some of the very first applications of geostatistics in the petroleum industry found power-law models.

The time maps associated with the variogram of Fig. 2-27 clearly appear nonsta-

MODELING 3D EXPERIMENTAL POROSITY VARIOGRAMS (CHU ET AL, 1994)

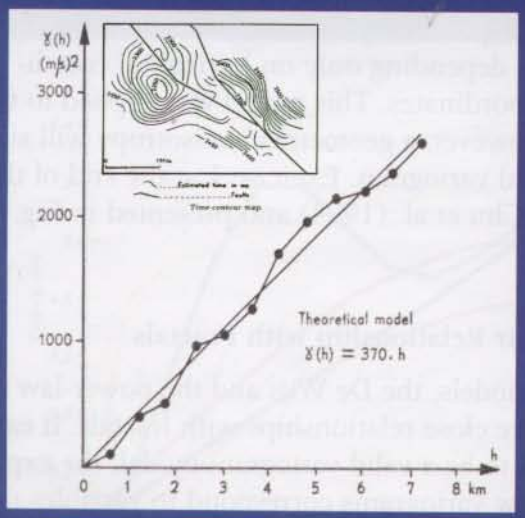




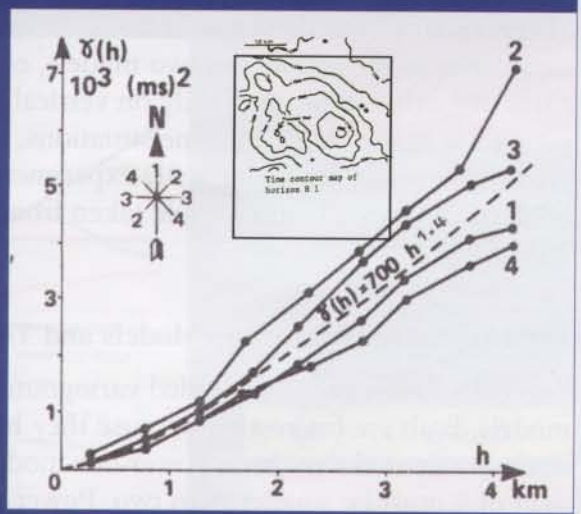


$$\gamma(h_i,h_j,h_k) = 5.5Sph\left(\sqrt{\frac{h_i^2+h_j^2+h_k^2}{0.2^2}}\right) + 8Sph\left(\sqrt{\frac{h_i^2}{4000^2}+\frac{h_j^2}{14000^2}+\frac{h_k^2}{1.9^2}}\right)$$

EARLY EXAMPLES OF POWER-LAW VARIOGRAMS (1)



Reflection times at top of hydrocarbon reservoir (Haas and Viallix, 1974)

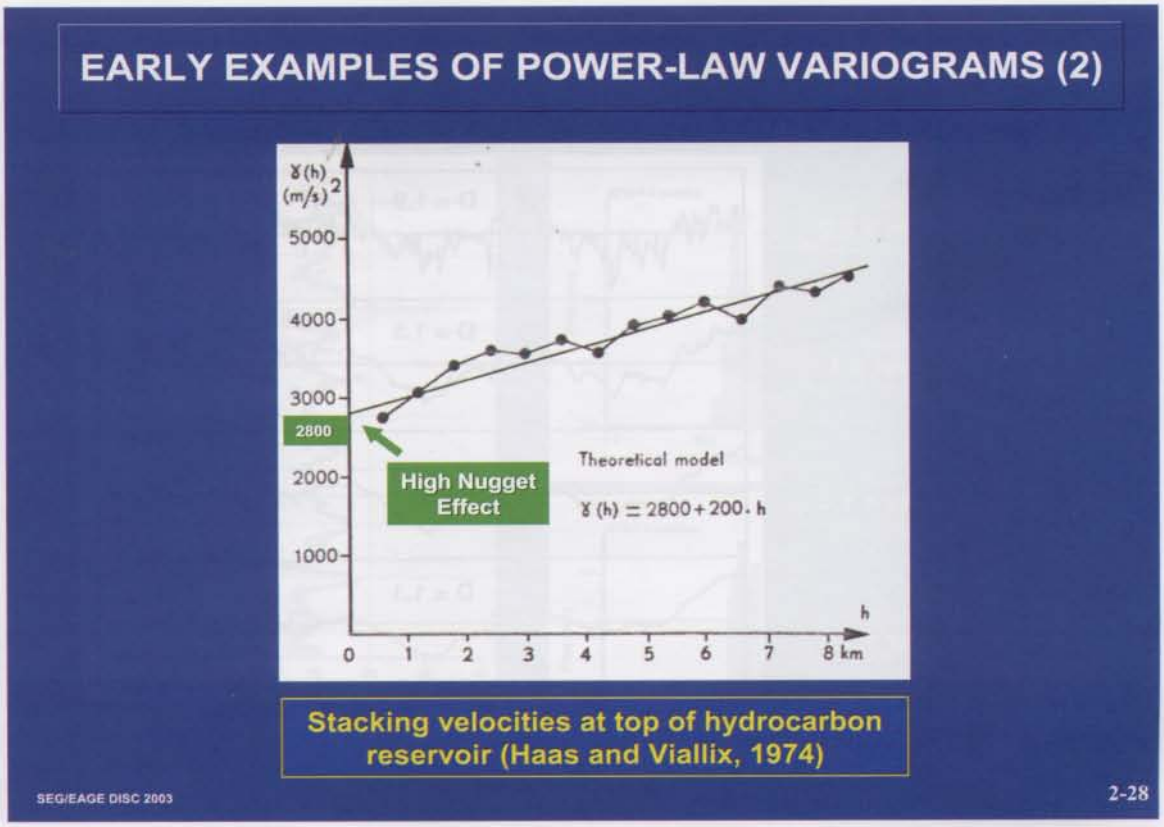


Reflection times at top of hydrocarbon reservoir (Haas and Jousselin, 1976)

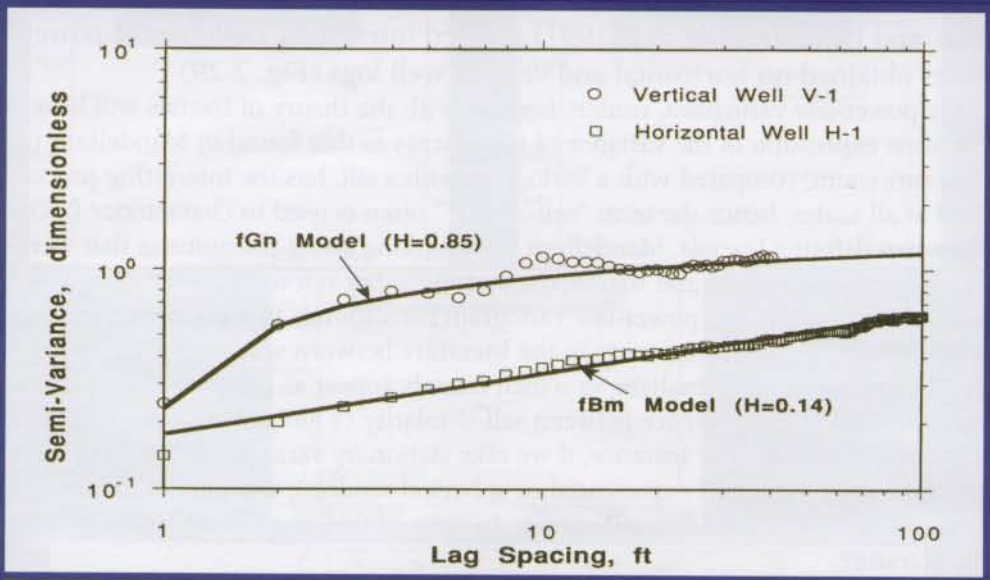
tionary, but at the same time no simple trend can be defined. Fig. 2-28 also shows a rare example of velocities following a power-law variogram. The high nugget effect is characteristic of stacking-velocities variograms, which will be discussed later. In the reservoir engineering literature, Hewett (1986) was a strong promoter of fractal models in the 1980s, and Perez and Chopra (1991) showed interesting examples of power-law variograms obtained on horizontal and vertical well logs (Fig. 2-29).

In the power-law variogram, readers familiar with the theory of fractals will have recognized the same expression of the variance of increments as that found in Mandelbrot (1982). A power-law variogram, compared with a variogram with a sill, has the interesting property that it is identical at all scales, hence the term "self-similar" often is used to characterize fractals. However, when defining fractals, Mandelbrot was thinking about phenomena that were new kinds of mathematical objects and that were continuous but not differentiable (such as the coast of Britain or Norway). Since a power-law variogram corresponds to a phenomenon that is both self-similar and fractal, confusion arose in the literature between self-similar and fractal phenomena. The geostatistical formalism, in which fractals appear as a subset of a wider class of models, helps clarify this difference between self-similarity (a global behavior) and nondifferentiability (a local behavior). For instance, if we take stationary variograms that have a linear behavior at the origin (such as exponential or spherical models), we see that they are associated with fractal models, which are not self-similar. In spite of this, the two concepts are often mixed up in the literature.

The relationship between the power of the variogram and fractal dimension is shown in Fig. 2-30. As the power of the variogram model tends toward zero, the fractal dimension tends toward the dimension of the space (1, 2, or 3) plus 1. This is another



EARLY EXAMPLES OF POWER-LAW VARIOGRAMS (3) (PEREZ AND CHOPRA, 1991)



Variogram of porosity logs shown in logarithmic coordinates

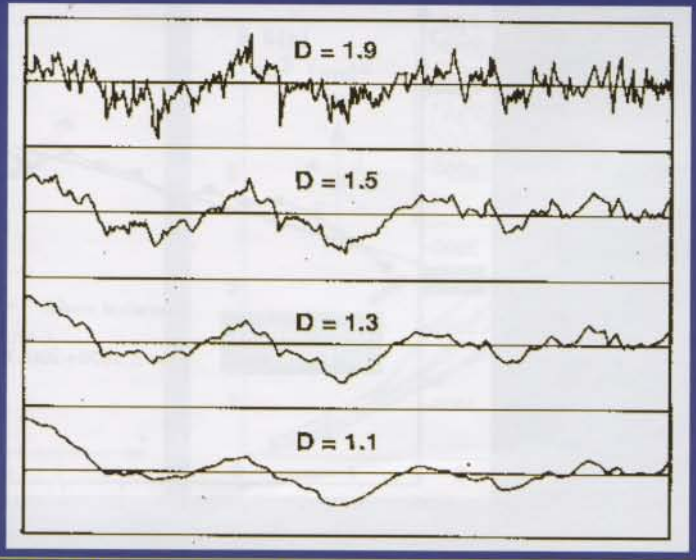
RELATIONS BETWEEN POWER-LAW VARIOGRAMS AND FRACTALS (PEREZ AND CHOPRA, 1991)

Power = 0.2

Power = 1

Power = 1.4

Power = 1.8



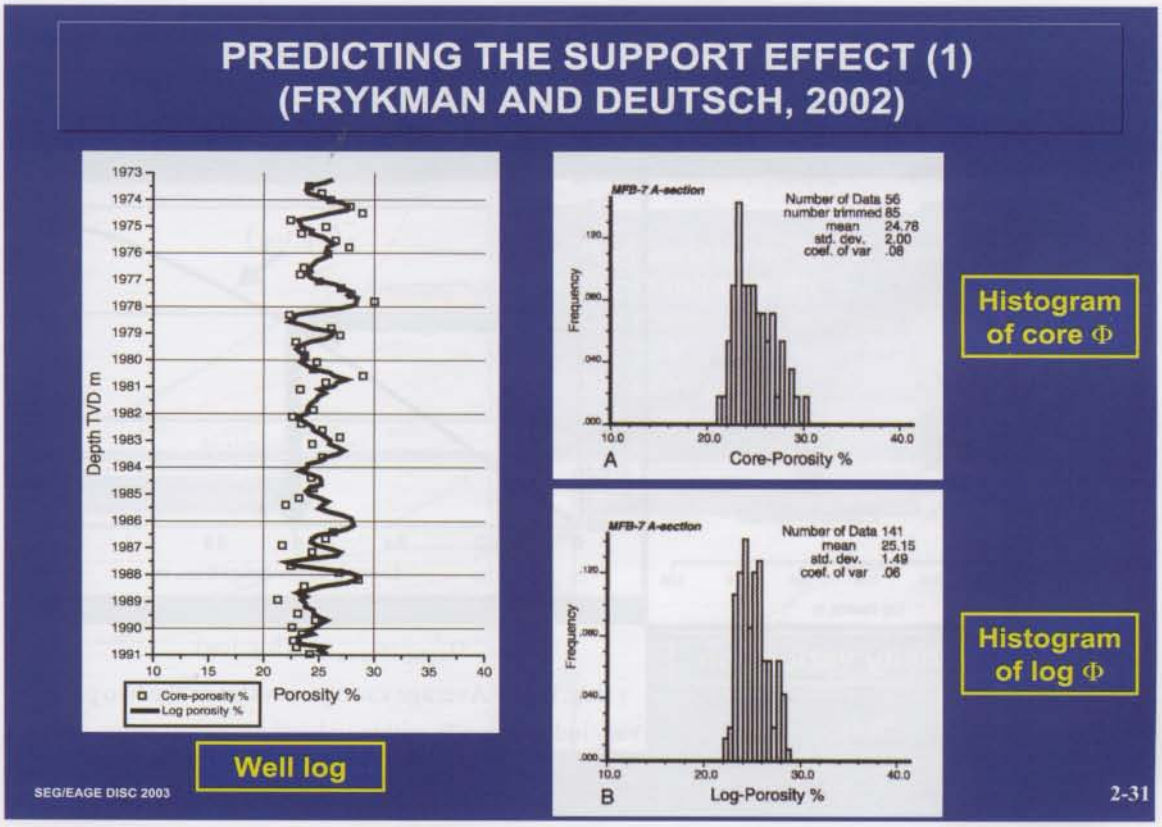
Power of the variogram = 2 (dimension of space +1 - fractal dimension)

way of saying (as shown in Fig. 2-24) that the closer the power of the variogram is to zero, the more random is the variable. A very noisy variable "fills" more space than a smooth one.

2.7 First Uses of the Variogram: Predicting the Support Effect

We saw earlier (Section 1.4.2) that, contrary to the mean, the variance is heavily dependent on the averaging volume. For instance, plug porosities are expected to show more scatter (or a higher variance) than are log-derived porosities, which results from a measurement over a larger volume. Thanks to formulas developed for the mining industry (Journel and Huijbregts, 1978), knowledge of the variogram on plug porosities and of the averaging volume for log measurements can be used to predict the variance reduction on the log-derived porosity data, as compared with the plug data. Strangely enough, this important result of geostatistics is almost never used in petroleum applications.

In a recent paper, Frykman and Deutsch (2002) address this issue using porosity measurements from a well in a chalk reservoir of the Dan field in the North Sea (Figs. 2-31 and 2-32). Along this well, they calculate the standard deviation of plug porosities and log-derived porosities. As expected, the standard deviation of log-derived porosities is smaller than that of plug porosities because of the support effect. Based on the variogram of plug porosities and thanks to the theoretical relationships of Journel and Huijbregts, they can predict how the variance of log-derived porosities will decrease as a

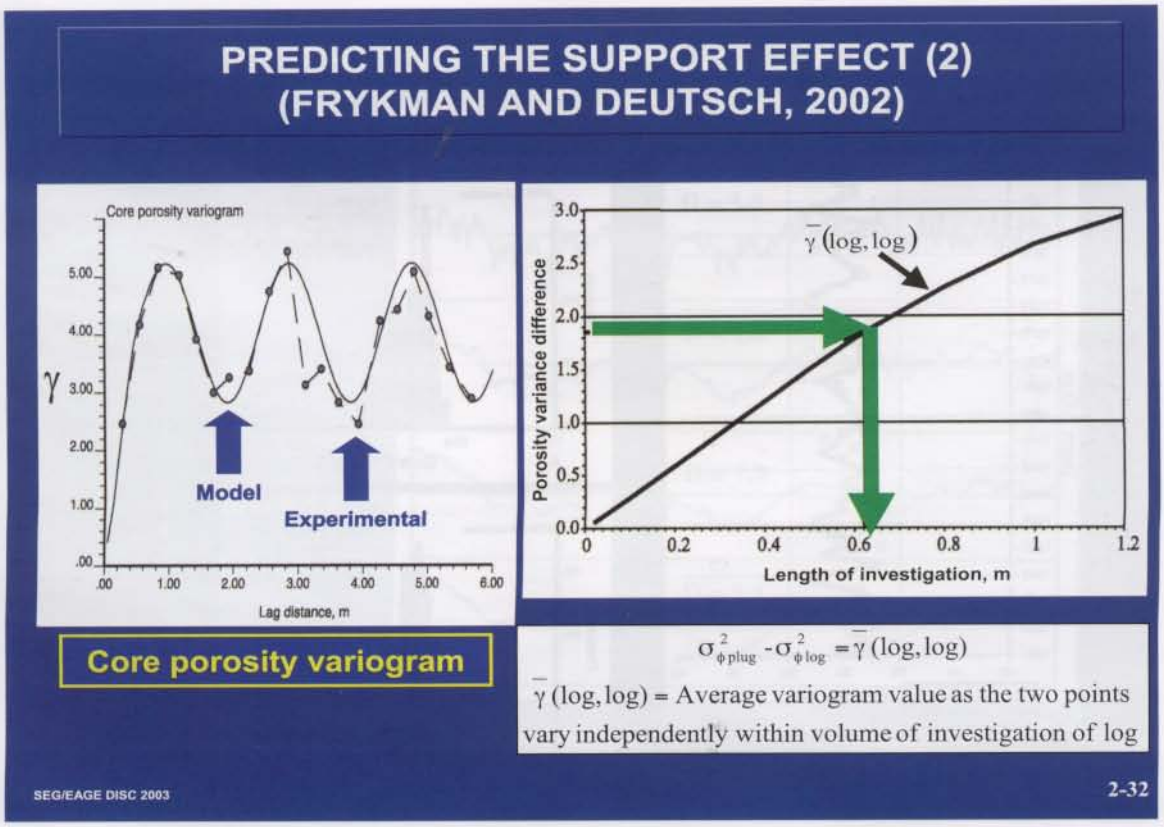


function of the volume of investigation of the log (making the simplification that it is a 1D, along-hole length of investigation). But, because they already know this variance, they invert the relationship. To achieve this, they plot the theoretical difference in variances between plug- and log-derived porosities, as predicted by the variogram model, under all possible assumptions about the length of investigation of the log. Then they pick the value that provides the best match between the experimental variance difference and the theoretical variance difference. They find, after a correction accounting for the well deviation, a length of 0.74 m, a value close to the 0.60 m predicted from the physics of the logging tool, which is an LDT (Litho-Density Tool).

At the end of their paper, Frykman and Deutsch list a number of reasons why this sort of approach is not routinely used in the petroleum industry. A key reason is that many properties — acoustic impedance or permeability, for instance — do not average linearly. The approach would also need to be simplified to be applied to a large number of wells and formations in a reservoir but, today, there is no industry-accepted methodology.

2.8 Cross-covariance and the Variogram

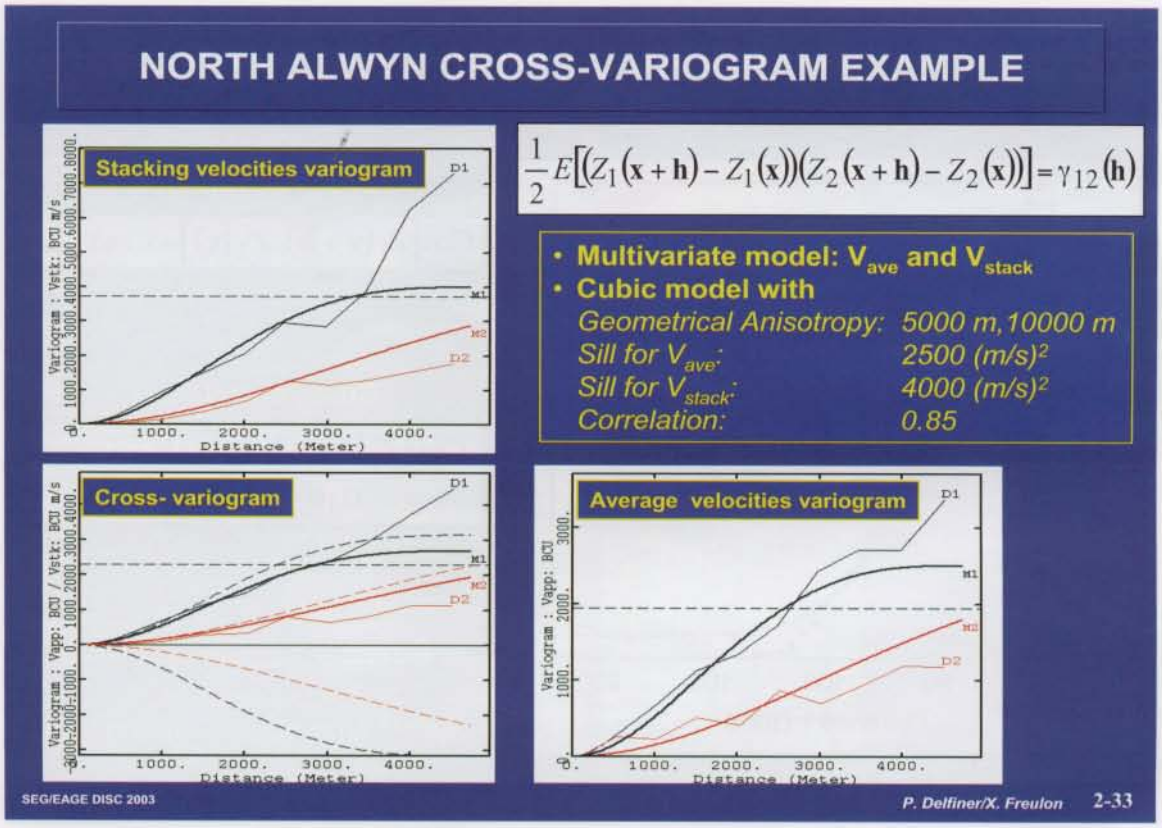
As we will see later, some of the most interesting applications of geostatistics will consist of combining different sets of data, such as porosity and acoustic impedance, or depth and two-way seismic time. How can we measure the relationships between such data sets? The correlation coefficient was presented earlier, but it does not incorporate



any spatial component, it simply evaluates the linear correlation between measurements of two parameters made at identical locations of space. What about the correlation coefficient between a value of porosity at a location **x** and a value of acoustic impedance, say, 1 km away? The cross-covariance function, or the cross-variogram, provides such information. The cross-variogram is defined in Fig. 2-33, which shows the cross-variogram of average and stacking velocity calculated over 59 wells. In this example, average velocity is equal to marker depth at the well divided by seismic one-way time, whereas stacking velocities have been previously smoothed. The cross-variogram is positive because the two variables are positively correlated. When the correlation is negative, the cross-variogram is negative. Note that the models used to fit the three experimental variograms are all proportional to a cubic model.

The cross-covariance and cross-correlation functions are illustrated in Fig. 2-34, which is taken from Doyen's (1988) important paper. In this case there were only 10 wells, which were not enough to calculate the porosity covariance. Note that, again, we are dealing with two positively correlated parameters, thereby resulting in a positive cross-covariance function. Doyen's approach was to fit a Gaussian model to both experimental variograms and then to assume that the porosity variogram was also proportional to this Gaussian model (see section 3.6.2).

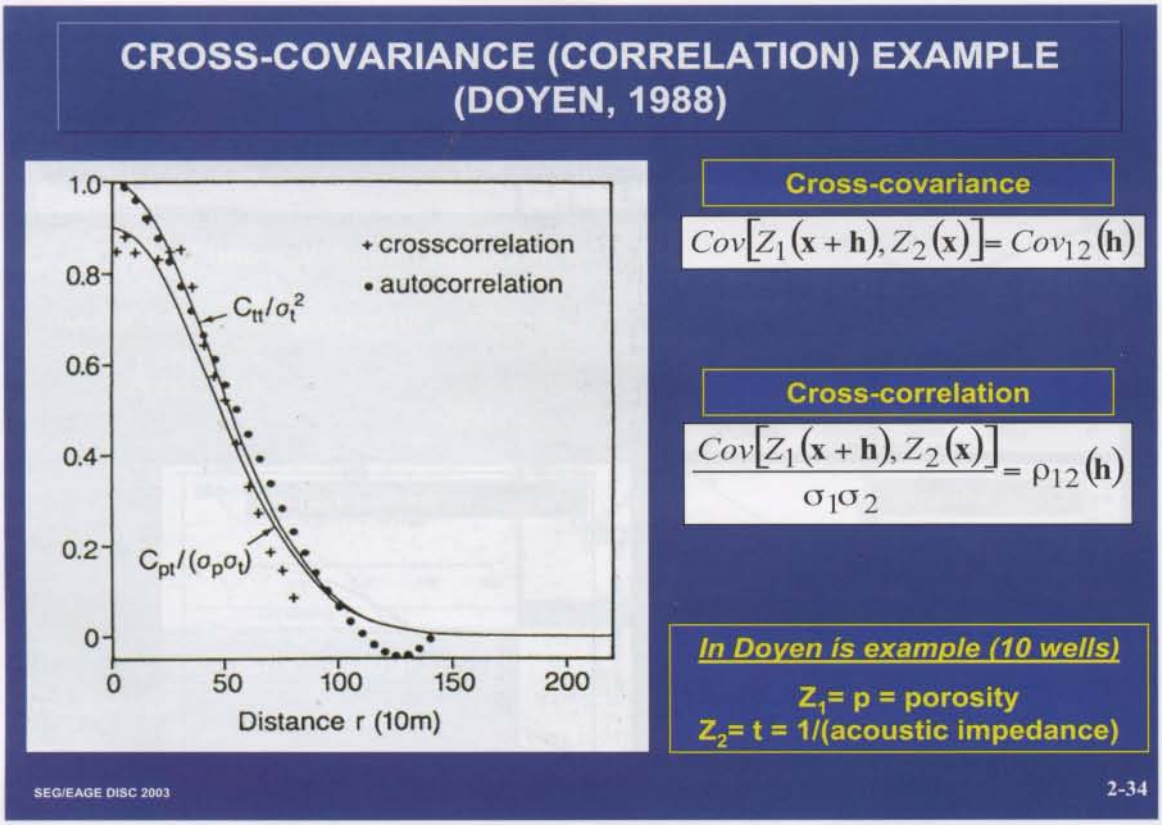
In many applications, the first variable $Z_1(\mathbf{x})$ (which we will call the primary variable) is measured at wells only, whereas $Z_2(\mathbf{x})$ (the secondary variable) is a parameter sampled on the seismic grid, which provides information about the primary variable. A paper by Xu et al. (1992) formalizes the modeling of cross variograms using propor-



tional-variogram models. If the cross-covariance is proportional to the covariance of the primary variable, we are dealing with a "Markov" model (which will lead to the collocated cokriging simplification of cokriging). In most practical applications, this Markov model — which uses only the primary variable variogram model and a proportional cross-variogram — is applied. The variogram of the secondary variable is usually of no use in collocated cokriging, as we will see below. Does this mean that it is of no use at all? We will see later that, in cases where there are very few well data, the variogram derived from seismic may provide useful insight into the variogram of the primary variable (see section 3.6.2).

2.9 Practical Considerations about the Variogram

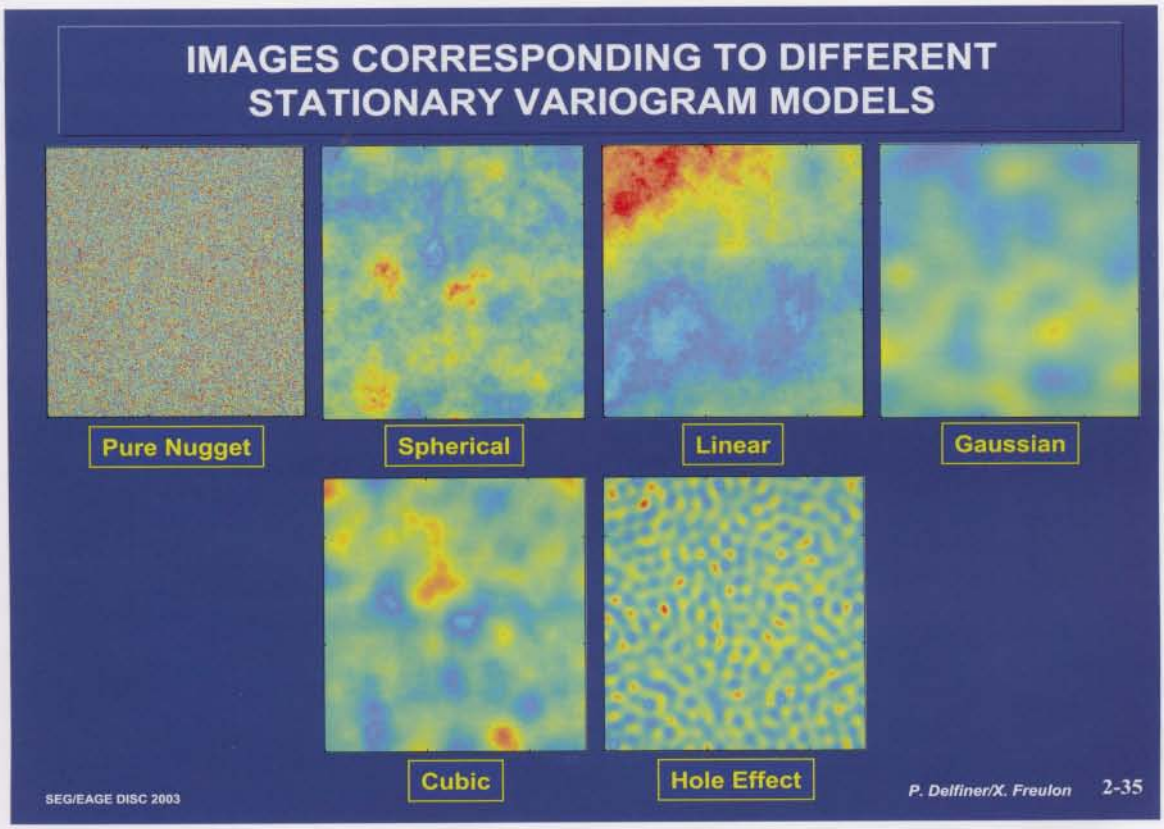
Let us try to understand the practical meaning of the different features of a variogram. Fig. 2-35 shows synthetic images corresponding to various variogram models. The pure nugget effect is clearly associated with the image of a white noise. The spherical, cubic, and Gaussian models all correspond to stationary images, which differ from one another by their degree of smoothness. The variogram with a linear behavior at the origin is associated with a more random-looking image than are the variograms with a parabolic behavior at the origin. The image corresponding with the Gaussian model varies extremely smoothly because this variogram model is indefinitely differentiable at the origin (Fig. 2-23). The image associated with the hole-effect variogram is periodic, as expected. Finally, the linear-variogram image shows no stationarity around a mean



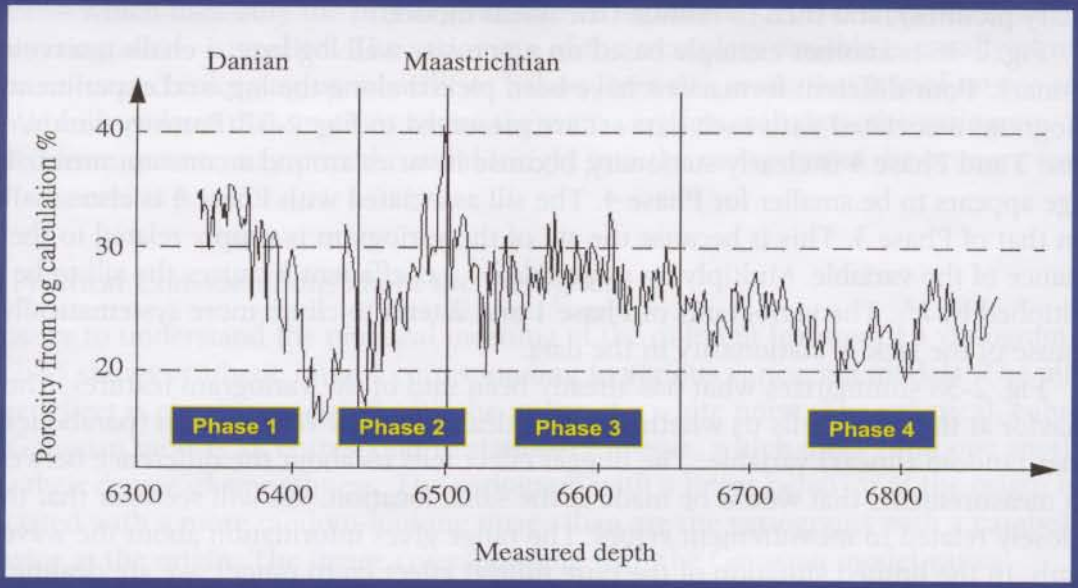
value, and hence no range. The range appears to be related to the wavelength of the oscillations around the mean, from zero with the pure nugget effect (two neighboring values have no reason to be on the same side of the mean), to medium (the three stationary pictures), and then to infinite (the linear model).

Fig. 2-36 is another example based on a porosity well log from a chalk reservoir in Denmark. Four different formations have been picked along the log, and experimental variograms associated with each data set are presented in Fig. 2-37. Porosity within Phase 3 and Phase 4 is clearly stationary, because it varies around a constant mean. The range appears to be smaller for Phase 4. The sill associated with Phase 4 is also smaller than that of Phase 3. This is because the sill of the variogram is simply related to the variance of the variable. Multiplying a variable by a coefficient k causes the sill to be multiplied by k^2 . The variograms of Phase 1 and 2 tend to climb more systematically because of the lack of stationarity in the data.

Fig. 2-38 summarizes what has already been said of the variogram features. The behavior at the origin tells us whether we are dealing with a very smooth (parabolic) or rather random (linear) variable. The nugget effect tells us about the difference between two measurements that would be made at the same location. We will see later that this is closely related to measurement errors. The range gives information about the wavelength. In the limited situation of the pure nugget effect (zero range), we are dealing with a very high-frequency white noise model. The sill itself does not have much geological meaning and is simply related to the variance. Usually, the experimental variance of the data is close to but smaller than the sill, and the difference between the two



WELL LOG INTERVALS CORRESPONDING TO DIFFERENT STATIONARY VARIOGRAM MODELS (1)

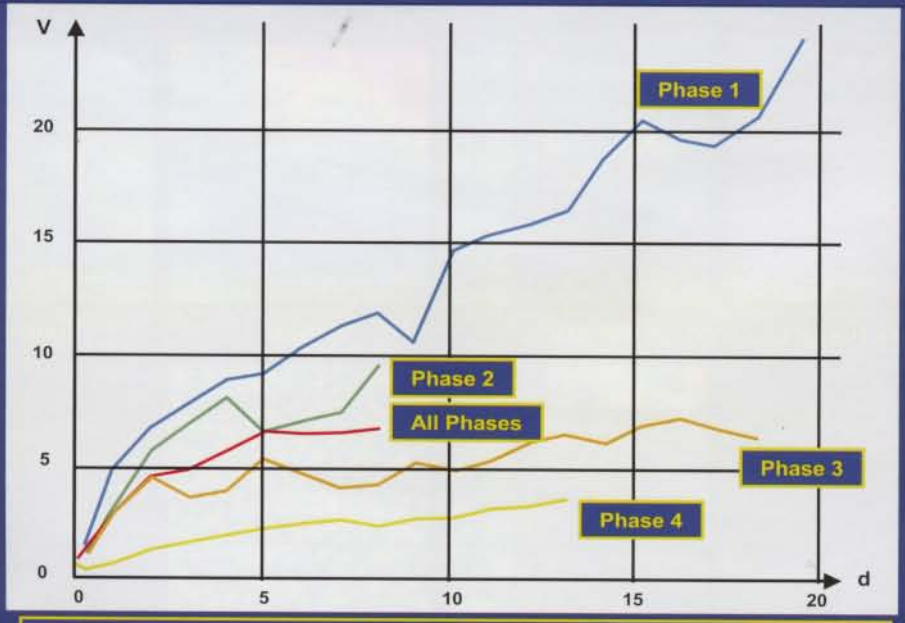


Porosity log of well M-10x in the Dan Field (Almeida and Frykman, 1994)

SEG/EAGE DISC 2003

2-36

WELL LOG INTERVALS CORRESPONDING TO DIFFERENT STATIONARY-VARIOGRAM MODELS (2)



Variograms on porosity log of well M-10x in the Dan Field

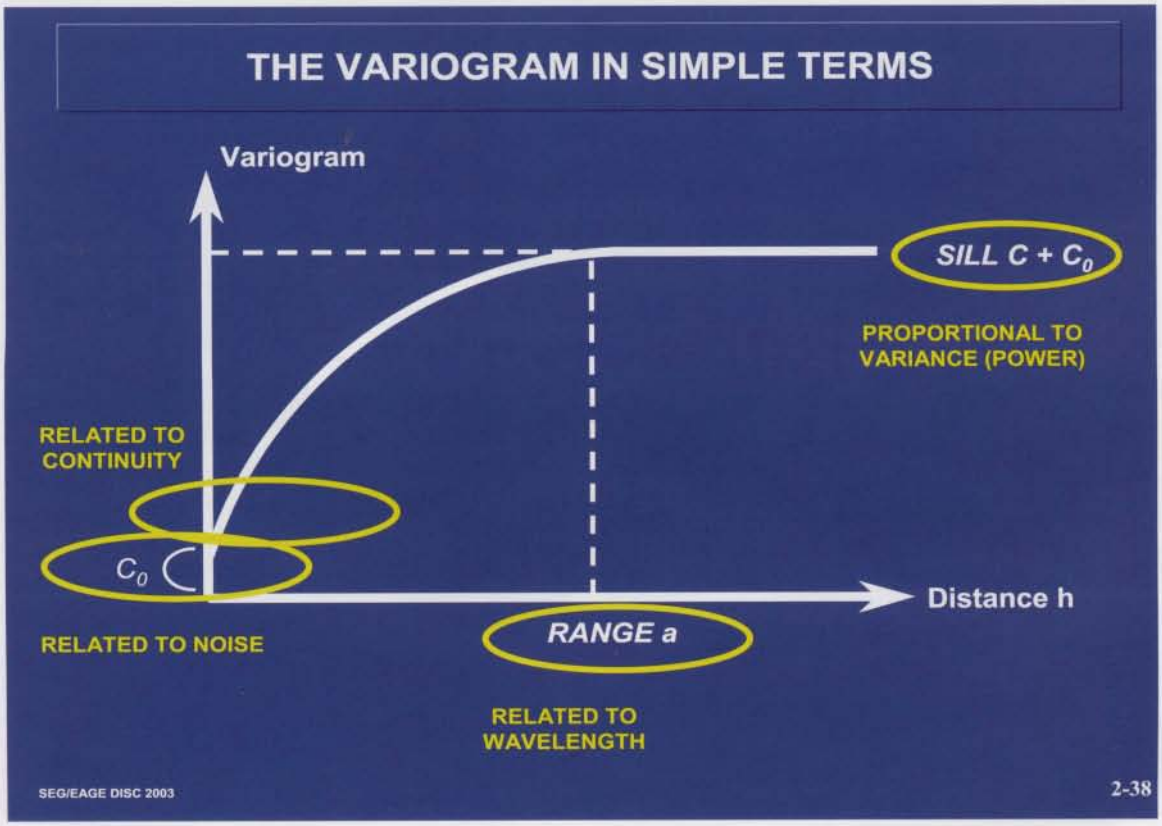
SEG/EAGE DISC 2003

2-37

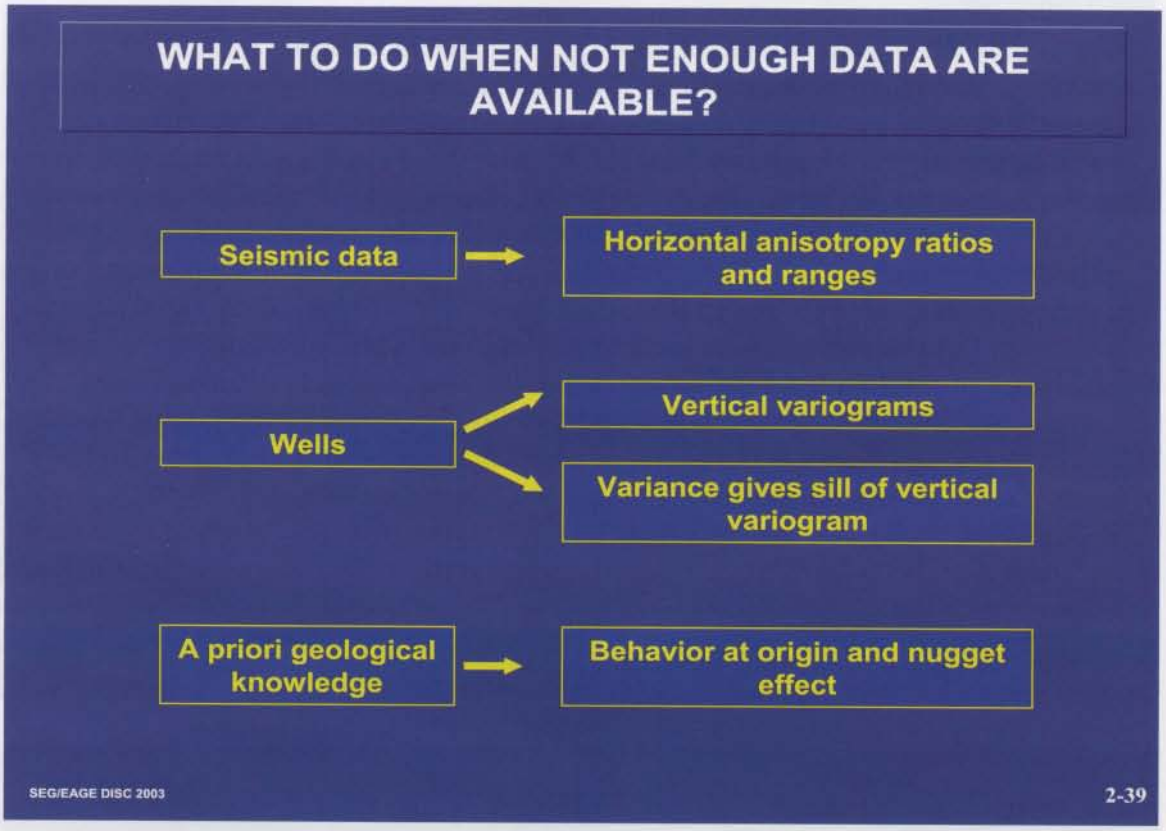
increases with the range of the variogram. So in conclusion, these four features of the variogram have a very practical interpretation. This will help us construct a variogram model when not enough data are available.

Unfortunately, in many situations, the primary variable (porosity, for instance) is known at only a few wells, whereas the secondary variable (acoustic impedance or seismic attribute, for instance) is known on the whole seismic grid. If the secondary and the primary variable are related, we can expect that anisotropies seen on seismic data will also be present on the primary variable, because these anisotropies are controlled by geology (depositional environment, diagenetic effects and the like). However, seismic data are much lower resolution than well data, and they also have different orders of magnitude (see the example of Fig. 2-19). This means that the sill of the primary-variable variogram can only be inferred from the (variance of) the well data. However, we also have a priori knowledge about each type of variable. This knowledge is the result of experience or information from analog data from other fields. For instance, we know that thickness varies more smoothly than permeability (behavior at the origin). We also know that stacking velocity is somewhat noisy. Such practical considerations can help us construct the variogram model (Fig. 2-39).

The above shows that the "geo" is at least as important as the "statistical." In most cases, because of the lack of data we do not apply a statistically rigorous approach, but we instead try to summarize all our a priori knowledge of the variable into the variogram model. It is a geoscientist's choice, not a statistician's calculation (Dubrule, 1994). This is why it is crucial to have a practical understanding of the various parame-



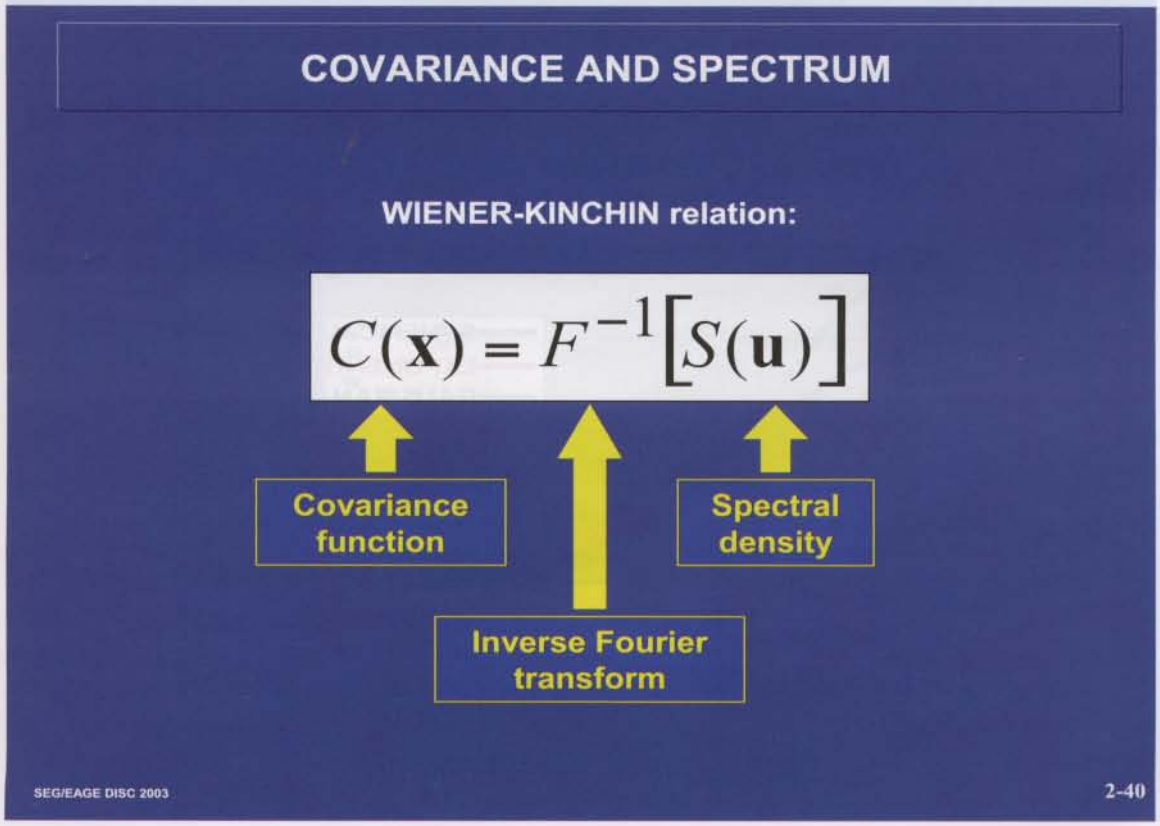
ters of a variogram model.



Covariance, Fractals, and Spectral Density

A covariance function is associated with a stationary model. Thanks to the Wiener-Kinchin relationship (Fig. 2-40), there is a relationship between the covariance function and the spectral density.

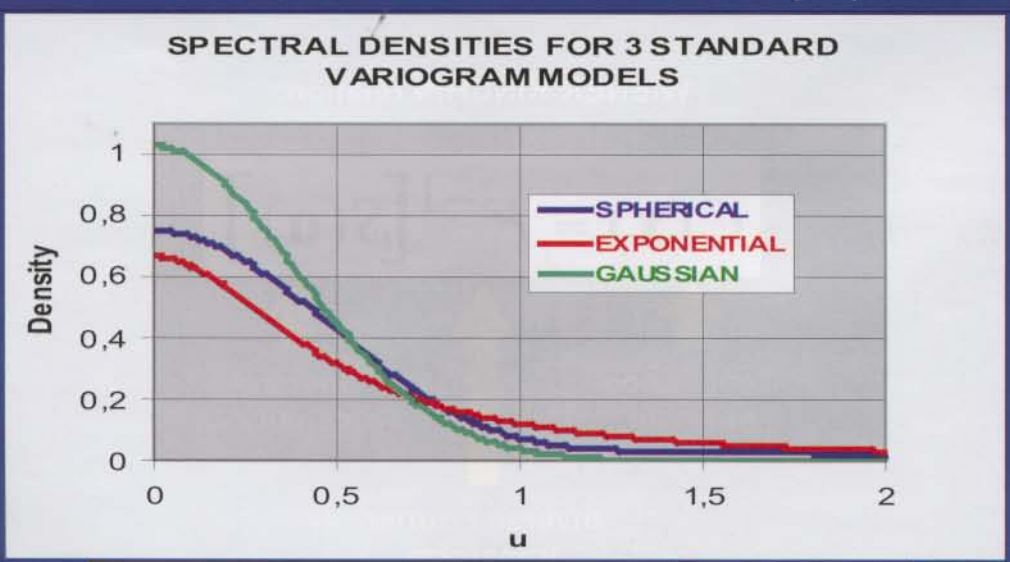
Chilès and Delfiner (1999) provide the analytical expressions of spectral densities associated with a large number of stationary variogram models. Fig. 2-41 shows the spectral densities of the spherical, exponential, and variogram models, plotted for the same value of the effective variogram range. The Gaussian model is associated with lower frequencies than is the spherical or exponential model. The exponential model, which, for the same range, has a steeper slope at the origin than does the spherical, is also associated with higher frequencies. The smoother the behavior of the covariance model at the origin, the lower the frequencies represented in the spectral density. In other words, the behavior of the covariance function at the origin translates into a behavior of the spectral density at infinity.



Maus (1999) and Maus et al. (1999) provide an interesting discussion of the value of analyzing magnetic and gravity data in the space or in the frequency domain. They also show how, in aeromagnetic applications, variogram models are derived from power spectra associated with physical models corresponding to a range of source depths. This will be discussed later.

We saw earlier that power-law variogram models are not strictly associated with stationary random functions. However, these variograms have a spectral representation. A characteristic of fractal models is that they follow a power law in their spectral density (Fig. 2-42). As already discussed in the stationary case, we see that the closer to 2 the power of the variogram is — that is, the smoother $Z(\mathbf{x})$ is — the lower are the frequencies that the spectrum carries. Fig. 2-43 shows an example of the behavior of well-log spectra modeled as fractal by Crane and Tubman (1990).

THE COVARIANCE, A FREQUENCY INTERPRETATION (1D)



The smoother the covariance at the origin, the fewer high-frequencies it carries

SEG/EAGE DISC 2003

2-41

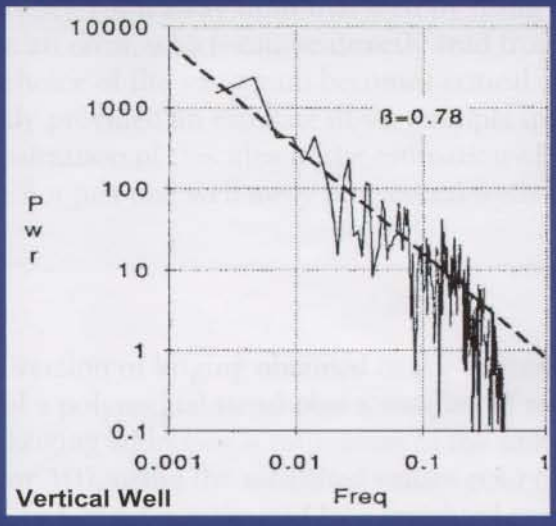
FRACTAL SPECTRAL DENSITY

A fractal follows a power law in its spectral density:

$$S(f) \propto \frac{1}{f^{\beta}}$$

β = power of the variogram + dimension of space

FRACTAL SPECTRAL DENSITY: AN EXAMPLE (CRANE AND TUBMAN, 1990)



Power spectrum of density log in a carbonate reservoir

3 Interpolation: Kriging, Cokriging, Factorial Kriging, and Splines

3.1 Introduction

In the previous two chapters, we discussed the meaning of the geostatistical model and of its parameters. We will now discuss how this model can be applied. We will start with deterministic techniques, known under the generic name of kriging. Here, "deterministic" should be understood in the sense of "providing only one solution." We will see that, although the model is probabilistic, kriging produces only one solution. Kriging covers a wide range of applications. The first one consists of interpolating one single variable in one, two, or three dimensions, and the second one consists of interpolating one variable but using the extra information provided by another variable that is related, of course, to the first one.

3.2 Kriging, an Interpolation Technique

3.2.1 Introduction

Recall the probabilistic model we defined in the previous chapter. The variable $Z(\mathbf{x})$ is interpreted as the sum of a polynomial trend, $m(\mathbf{x})$, plus a residual, $R(\mathbf{x})$, of mean zero. Under this model, universal kriging (Matheron, 1970) addresses the problem of interpolating a variable on the basis of a number of scattered data. This can be the interpolation of layer-averaged porosity from well data or the interpolation of seismic times from a 2D seismic campaign.

To understand kriging, let us consider the variogram from another perspective (Fig. 3-1). Suppose that layer-averaged porosity has been calculated at a well and that we want to estimate porosity 1 km away from that well by using the value at the well. Obviously, we will make an error, which can be directly read from the variogram model plot. This is where the choice of the variogram becomes critical. By choosing this model, we have implicitly provided an estimate of very simple interpolation errors. Kriging is just the generalization of this idea to the estimation of the value at location $\mathbf{x}$ using, this time, a value not just one well away but several wells away.

3.2.2 Universal kriging

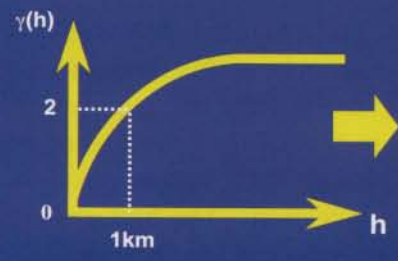
- A bit of theory

Universal kriging is the version of kriging obtained in the context of the model in which $Z(\mathbf{x})$ is the sum of a polynomial trend plus a residual of mean zero, $R(\mathbf{x})$. The problem that universal kriging addresses is estimation of the unknown value, $z(\mathbf{x}_0)$, at a location $\mathbf{x}_0$ (in 1D, 2D, or 3D), using the measured values $z(\mathbf{x}_i)$ obtained at N number of scattered data points, $(\mathbf{x}_i)$. $z(\mathbf{x}_0)$ is estimated by a weighted average of the measured values $z(\mathbf{x}_i)$. Note that we use the notation z when we are discussing a realization of the random function Z. Fig. 3-2 shows the criteria used to calculate the weighting factors. It is natural to use the standard statistical approach of calculating the unbiased minimum

KRIGING: ANOTHER USE FOR THE VARIOGRAM

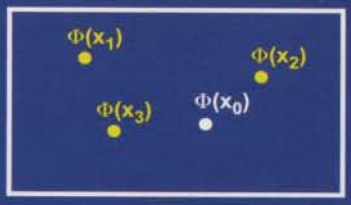
This variogram provides an estimation of the average squared error made when estimating the porosity value (in porosity units) 1km away from a well using the porosity at the well:

$\gamma(h) = 1/2\ Var[\Phi(x+h) - \Phi(x)]$



GAUSSIAN confidence interval= [Φ(well)-2x2, Φ(well)+2x2]

The variogram can also be used to estimate: $Var[\Phi(x_0) - a\,\Phi(x_1) - b\,\Phi(x_2) - c\,\Phi(x_3)]$

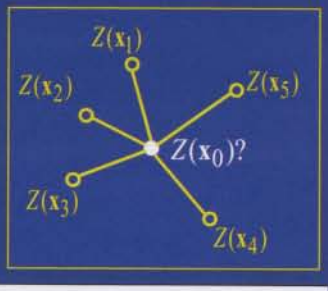


Kriging will consist of finding the weights a,b,c minimizing this variance

UNIVERSAL KRIGING

$$Z_{uk}(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i Z(\mathbf{x}_i)$$

Pb: find λ_i



The estimation variance is calculated:

$$\sigma_E^2 = Var\left[Z_{uk}(\mathbf{x}_0) - Z(\mathbf{x}_0)\right] = \sum_{i,j} \lambda_i \lambda_j C(\mathbf{x}_i, \mathbf{x}_j) - 2\sum_i \lambda_i C(\mathbf{x}_i, \mathbf{x}_0) + C(\mathbf{x}_0, \mathbf{x}_0)$$

Two conditions to satisfy:

$$E\left[Z_{uk}(\mathbf{x}_0)\right] = E\left[Z(\mathbf{x}_0)\right]$$

Unbiased

σ_E^2 Minimal

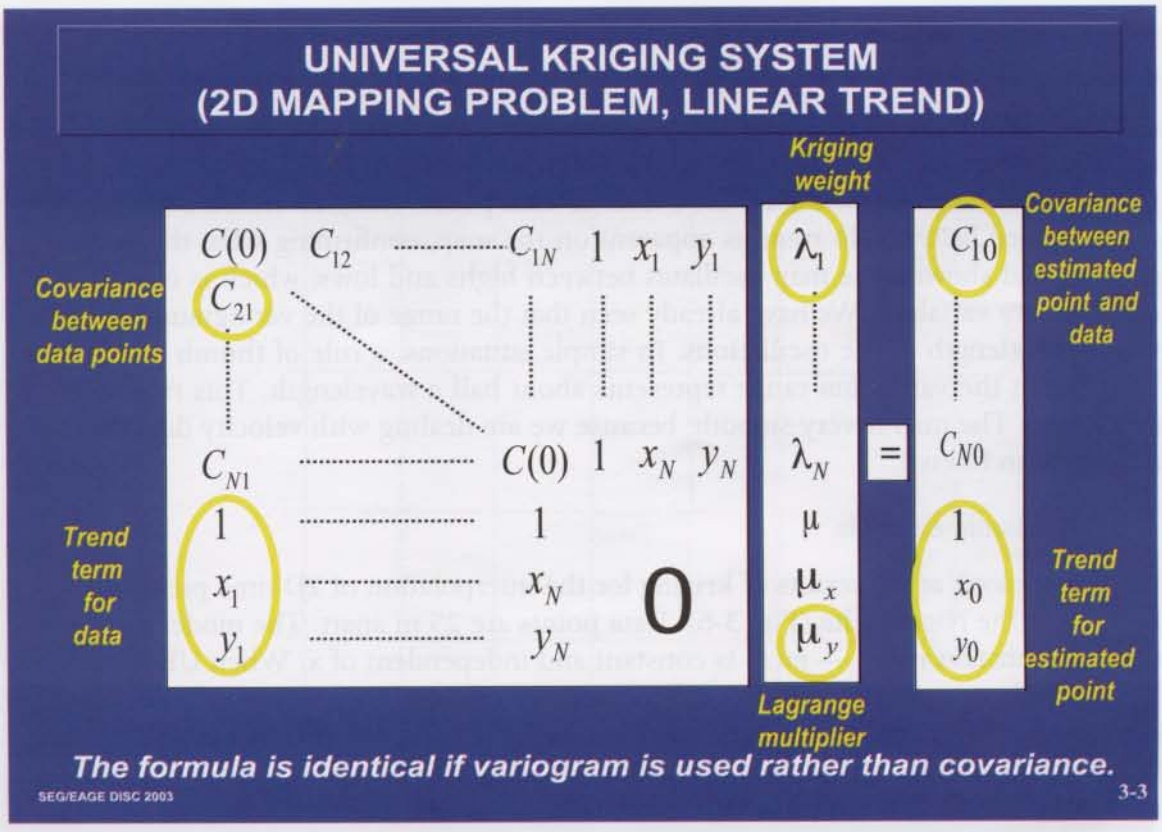
Optimal

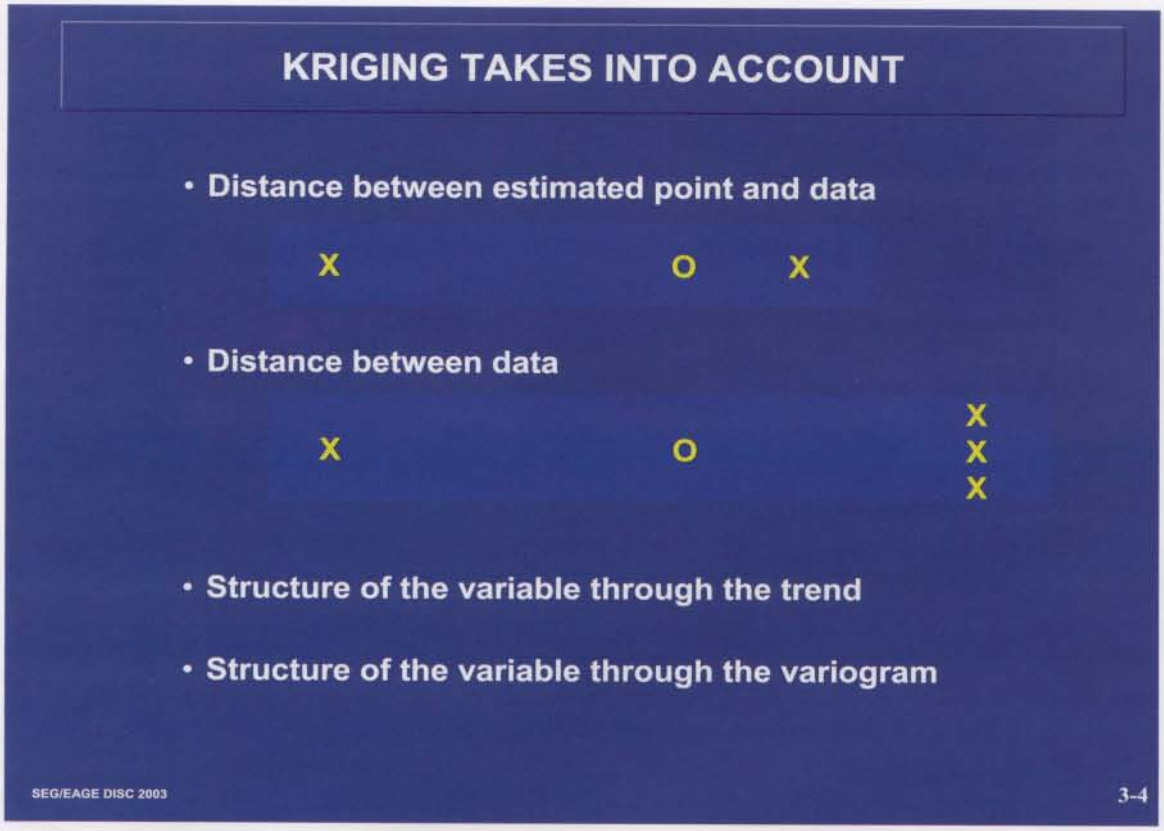
variance estimator $Z_{uk}(\mathbf{x}_0)$. Minimization of the estimation variance (using Lagrange multipliers μ to account for the unbiased condition) leads to the universal kriging (UK) system (Fig. 3-3). Note that the UK system is the same whether we use the covariance or the variogram function, because only the Lagrange multipliers change.

Obviously, the kriging system accounts for the variogram and the trend model. Changing the assumptions on any element of these two models will lead to a change in the kriging system and in the kriging weights. In the kriging system, the relative position of the data points is taken into account through the covariance function; whereas on the right-hand side, the relative locations of the data points vis à vis the estimated point are accounted for (Fig. 3-4).

Exercise 2 at the end of this book uses an example with four data points to illustrate how kriging weights change as the variogram and the data-point locations change.

A last item that is of theoretical interest at this stage is the definition of the "simple kriging" (SK) system. SK is the system of equations obtained with the covariance terms only, without any trend present. The SK system is simply derived from the equation of Fig. 3-3 by removing the last three lines and three last columns of the matrix, the Lagrange multipliers, and the three last lines of the right-hand side. With UK, the coefficients of the trend are automatically derived from the UK system. With SK, the user may inject into the system a fixed mean value. This is especially handy in situations where a correction factor is interpolated by SK, because this correction factor then becomes equal to zero far away from the well data.





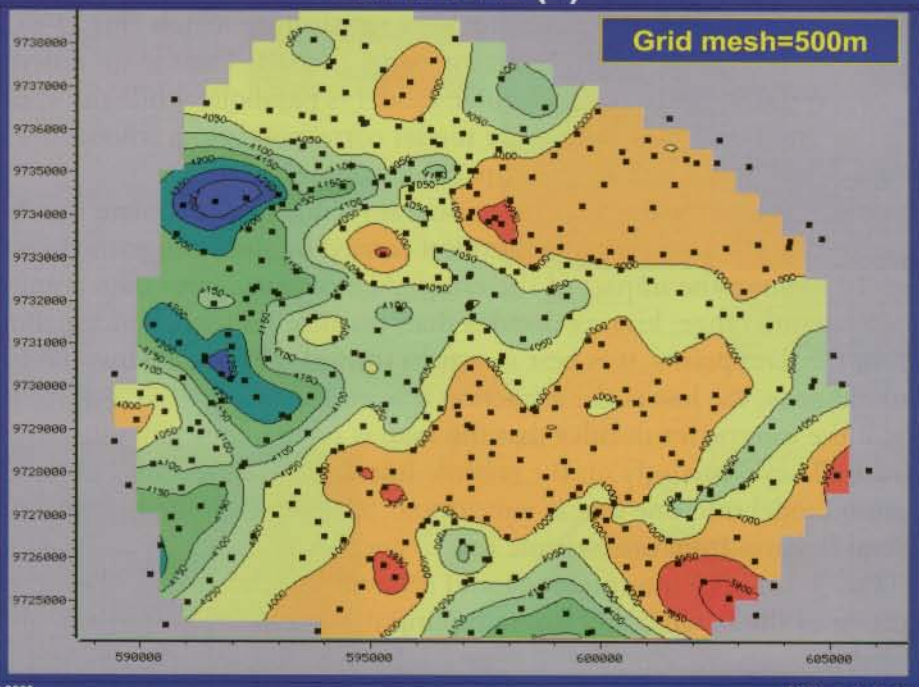
• Velocity data example

Enough of theory. Fig. 3-5 shows the kriged map calculated from the velocity data already discussed in Figs. 2-12, 2-13, and 2-15. This map is obtained with a constant trend and the variogram model of Fig. 2-13. A global neighborhood is used, which means that every grid point is estimated using all the available data points (neighborhood considerations are discussed below). No trend is apparent on the map, confirming what the variogram analysis had shown. The map oscillates between highs and lows, which is characteristic of stationary variables. We have already seen that the range of the variogram was related to the wavelength of the oscillations. In simple situations, a rule of thumb that is often used is that the variogram range represents about half a wavelength. This rule seems to apply here. The map is very smooth, because we are dealing with velocity data that have already been filtered.

• A 2D seismic example

Let us now look at the results of kriging for the interpolation of 2D time picks of a horizon in the Niger Delta (Fig. 3-6). Data points are 25 m apart. The model used is that of a constant mean — $m(\mathbf{x})$ is constant and independent of $\mathbf{x}$. When UK is performed with a constant mean, it is called ordinary kriging (OK). Because faults are present, the variogram calculation accounts for the faults, and no pairs of points are used in the calculation that have their extremities on each side of a fault. The variograms that are calculated along and perpendicular to the main structural directions (Fig. 3-7)

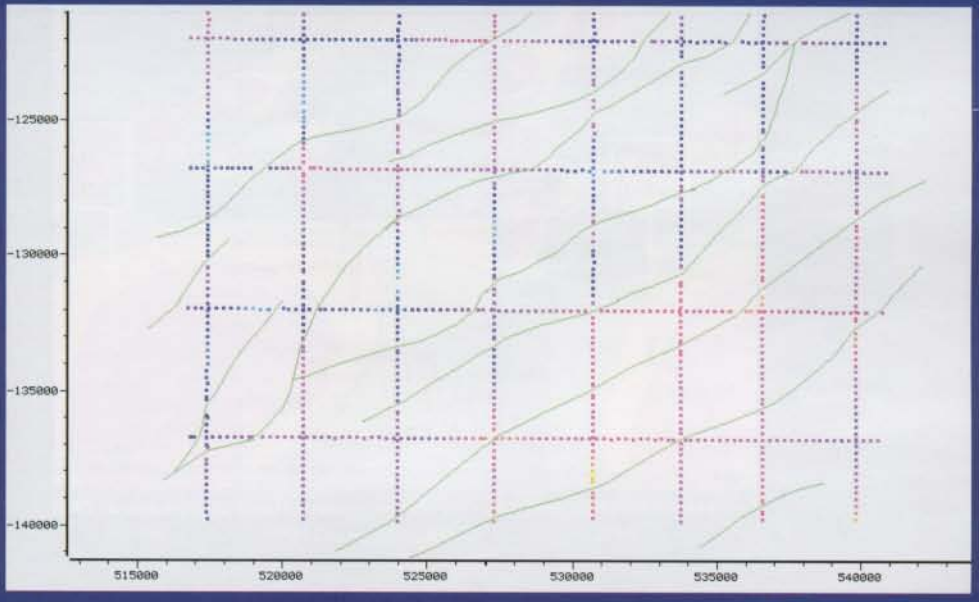
KRIGED MAP OBTAINED ON VELOCITY DATA EXAMPLE (4)



SEG/EAGE DISC 2003 V. Bigault de Cazanove 3-5

KRIGING: A 2D SEISMIC EXAMPLE

2D time data location and faults

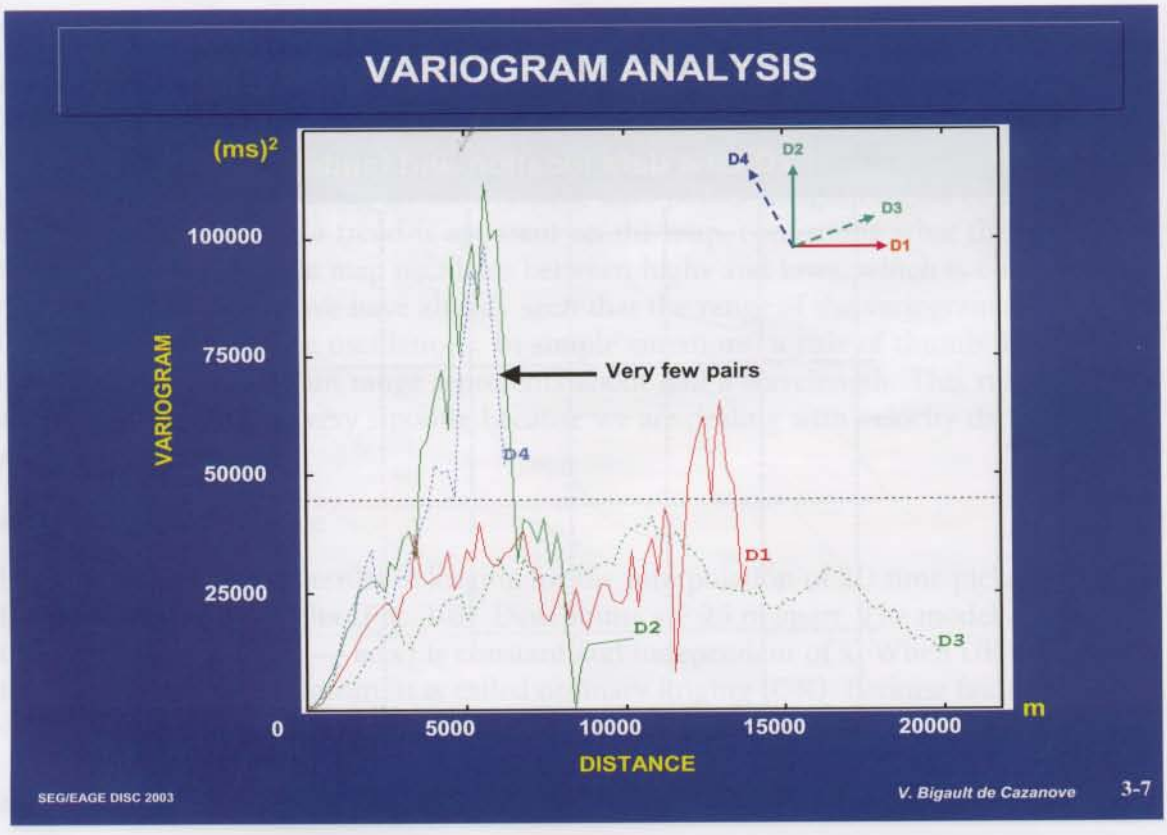


SEG/EAGE DISC 2003 V. Bigault de Cazanove 3-6

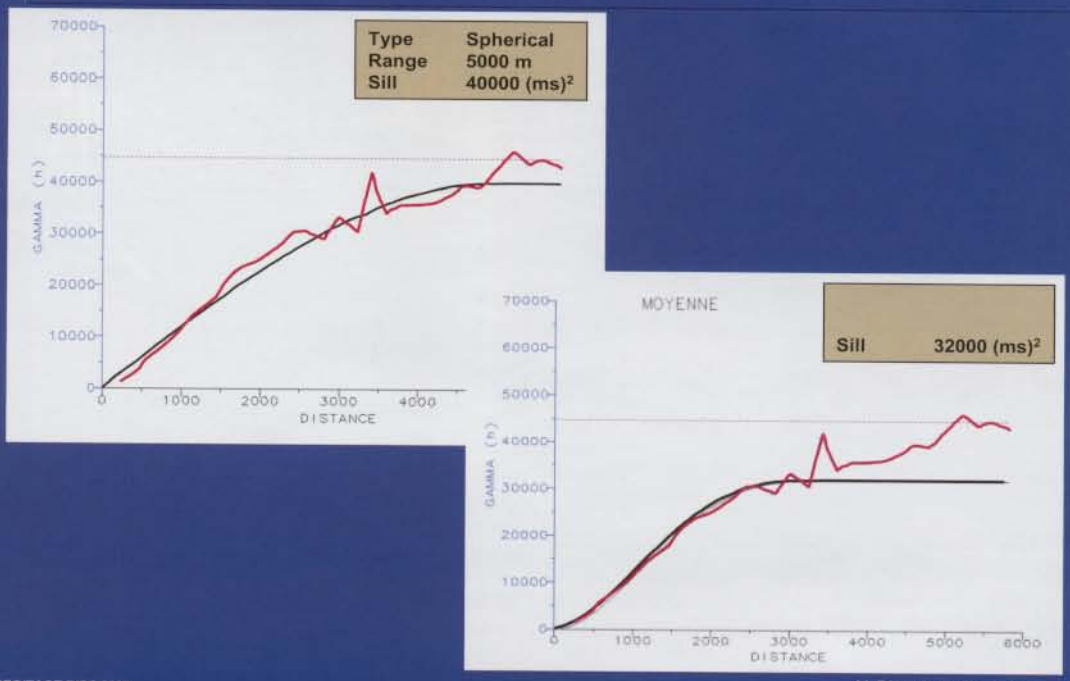
behave somewhat similarly until about 3000 m, at which point the directions D2 and D4 diverge, but the number of pairs used to calculate them is not sufficient to make the difference significant. Thus, an isotropic experimental variogram is calculated and modeled. Two variogram models appear to provide a satisfactory match (Fig. 3-8). They have no nugget effect and a slightly different range and sill. Their main difference lies in their behavior at the origin, where the cubic model is parabolic while the spherical model is linear. We know that the former model corresponds to a smoother variable than does the latter. What will be the impact on kriging?

The map associated with the cubic variogram tends to extrapolate trends more strongly because of the continuity assumption carried by the variogram (Figs. 3-9 and 3-10). Fig. 3-11 shows the impact of the range. At a distance from data points greater than the variogram's range, kriging decides that, because no useful information can be derived from the data points, it is best to return toward the mean. Thus the larger the variogram's range is, the less is the attraction toward the mean.

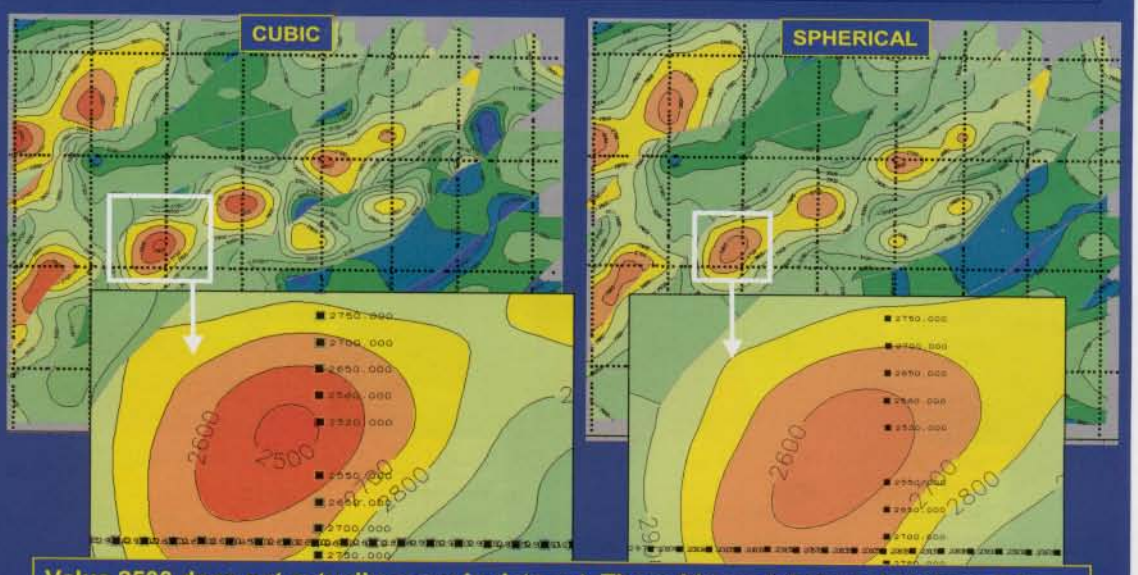
What if the interpreter decides that the map is not geologically satisfactory, because the bulls-eyes are only on the seismic lines? Then he/she may decide to change the variogram model into one that is anisotropic, which has the advantage of propagating structural features from one seismic line to another and parallel to the main fault's direction (Fig. 3-12). The interpreter should always have the last word and be able to take advantage of the flexibility of kriging to input any a priori geological information that is available.



ISOTROPIC VARIOGRAM MODEL FIT: TWO OPTIONS



COMPARE THE TWO MAPS (1)

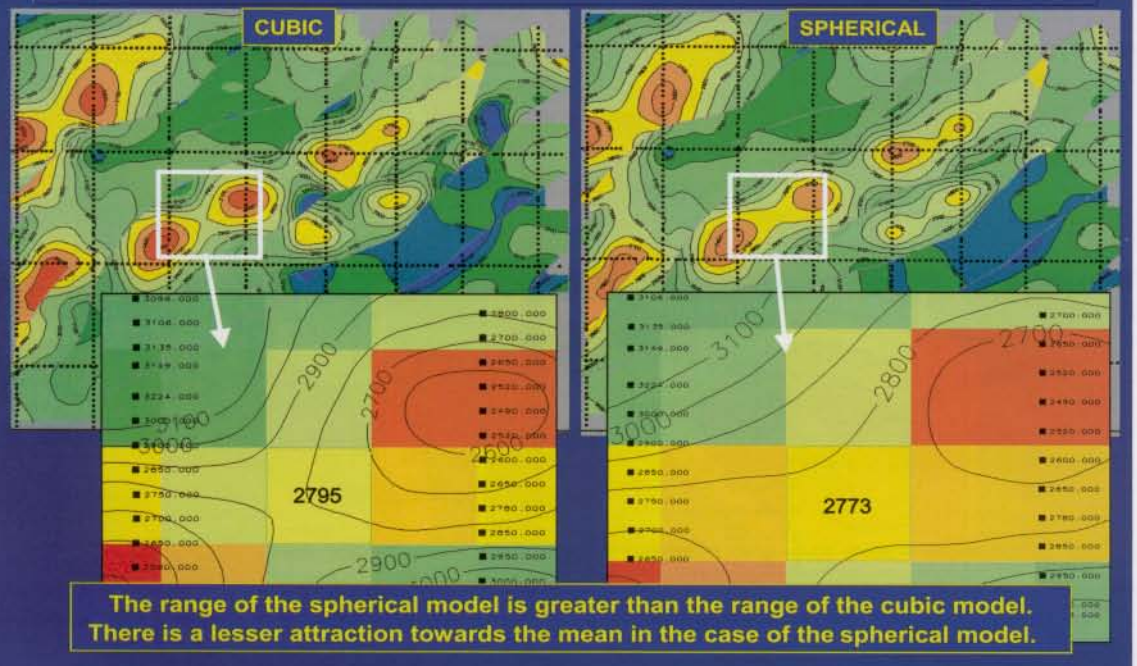


Value 2500 does not actually occur in data set. The cubic model strongly extrapolates because of its continuity assumption (behaviour at the origin). The spherical model carries a weaker continuity assumption: estimation there strictly follows the data.

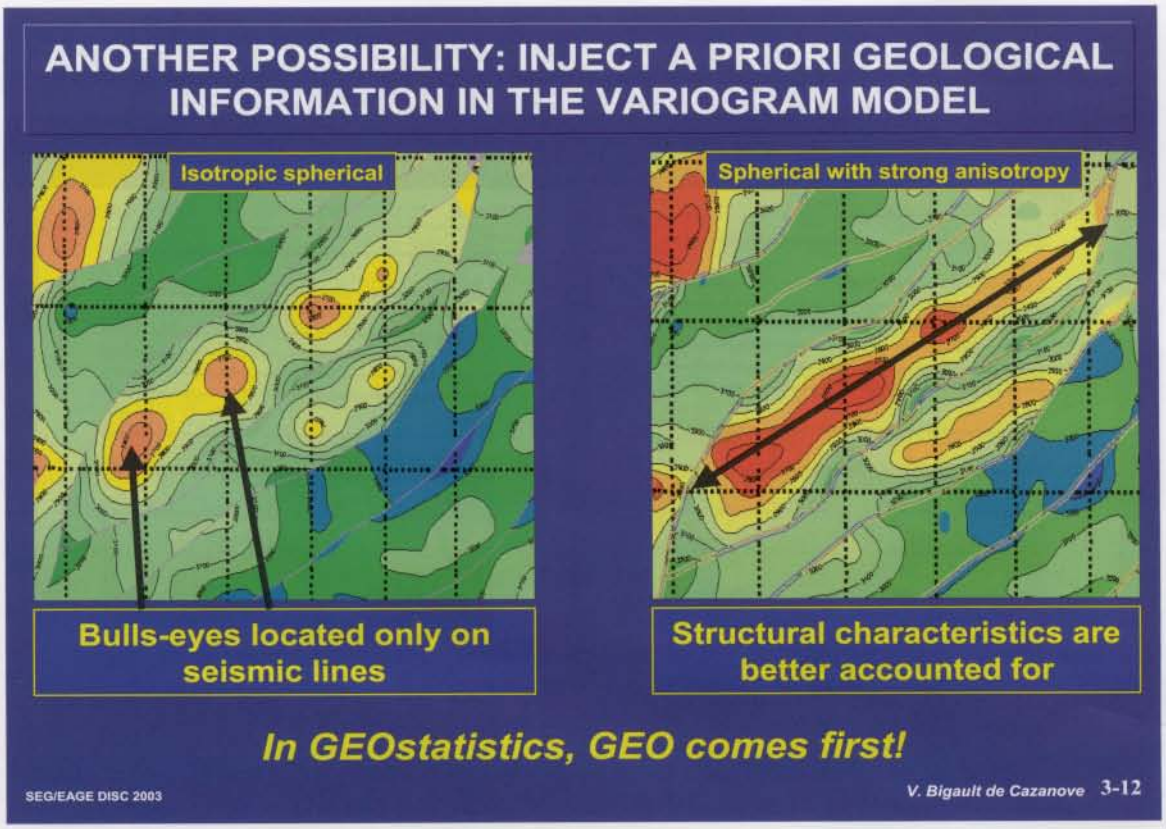
COMPARE THE TWO MAPS (2)

A nearby couple of data with great difference of value triggers a slope in the cubic model interpolation away from the data, resulting in a distal trough. The spherical model, because of its weaker continuity assumption, gives less impact to this apparent slope.

COMPARE THE TWO MAPS (3)



The range of the spherical model is greater than the range of the cubic model. There is a lesser attraction towards the mean in the case of the spherical model.



- An aeromagnetic data example

Hansen (1993) provides an interesting example of kriging of aeromagnetic data from flight data points. Instead of deriving the covariance model by geostatistical analysis, he derives it from the spectrum associated with the theoretical model of Spector and Grant (1970). He calculates the parameters of this model from a spectral analysis of the data, which have been interpolated on a regular grid using a minimum-curvature approach.

As in the previous example, Hansen observes (Fig. 3-13) that contour lines tend to close around the data lines only. By forcing an anisotropy into the covariance model, the kriging result obtained appears to be more geologically satisfactory to Hansen.

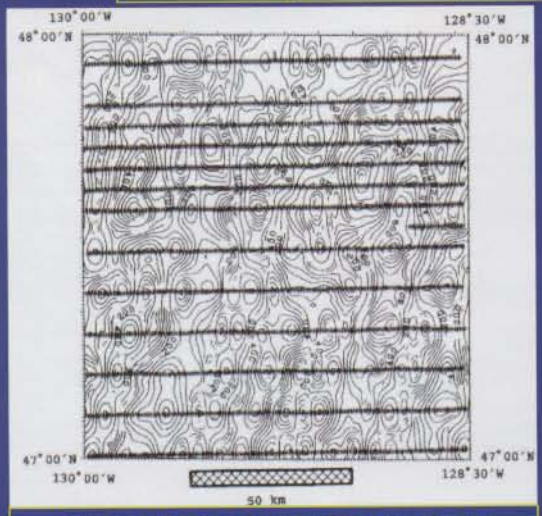
With such an example, it would have been interesting to see the result obtained from geostatistical variogram analysis followed by kriging. This would have presented the advantage of validating the covariance model derived from the Spector and Grant spectrum. It also would have simplified this approach by requiring only one interpolation. In Hansen's approach, two interpolations are required: the intermediate minimum-curvature grid (required to calculate the spectrum) and the final one.

- An example of 2D kriging from well data

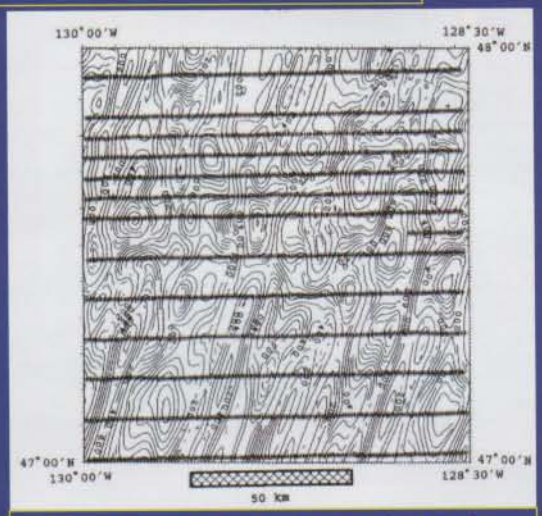
Fig. 3-14 shows different maps of average velocity obtained by kriging from 59 wells on the North Alwyn field (U.K.). As before, the trend is assumed to be constant and unknown (ordinary kriging). The variogram model is cubic and anisotropic, with a range of about 5000 m in the X direction and 10000 m in the Y direction (see Fig. 2-33).

ANOTHER EXAMPLE OF INJECTING A PRIORI GEOLOGICAL INFORMATION IN THE VARIOGRAM MODEL (HANSEN, 1993)

Aeromagnetic data (Cobb offset total field anomaly)



Kriging obtained with the covariance model associated with the Spector and Grant (1970) spectrum model.

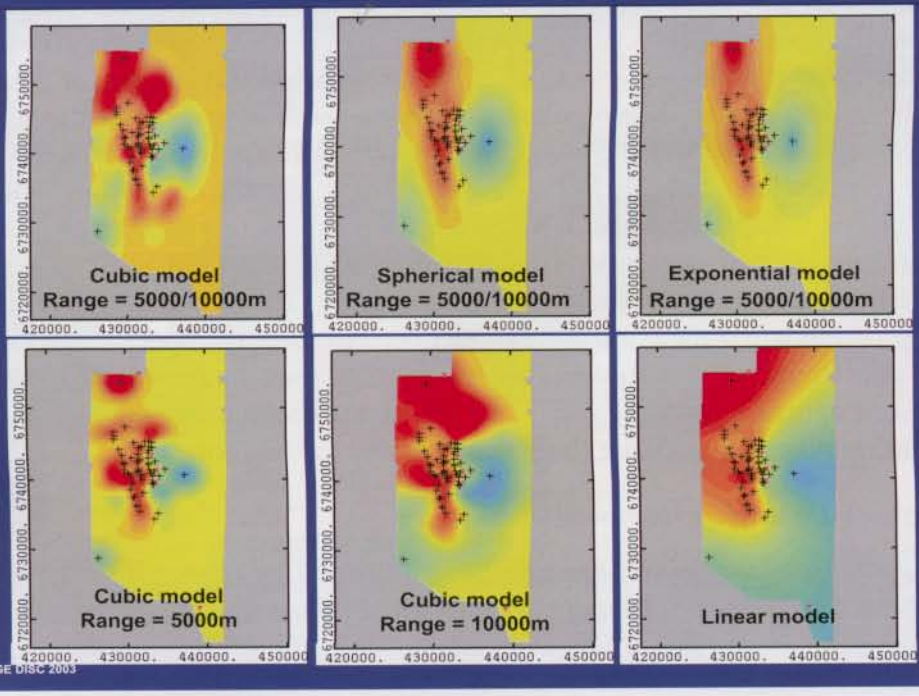


Kriging obtained with an anisotropic covariance associated with the Spector and Grant (1970) spectrum model.

SEG/EAGE DISC 2003

3-13

EXAMPLES OF AVERAGE-VELOCITY KRIGING (59 WELLS, NORTH ALWYN FIELD, NORTH SEA)



P. Delfiner/
X. Freulon

SEG/EAGE DISC 2003

3-14

The top maps show the different results obtained by keeping the same anisotropic ranges but changing the model. The exponential and the spherical models give similar results. This is not surprising, because they are both linear at the origin. The only difference is that, for the same range, the exponential model climbs faster toward the sill. This has very little impact on the maps. On the other hand, the cubic model provides a much smoother map. As mentioned earlier, parabolic behavior of the variogram at the origin leads to a model that tends to extrapolate local data trends more strongly than does a linear behavior at the origin. Note, in particular, the more pronounced extrapolations, whereas the exponential and spherical remain closer to the mean.

The bottom maps show the results of three isotropic interpolations. The impact of changing the range from 5000 m to 10000 m is clear around the isolated wells. Their high and low velocities are extrapolated farther away with the larger-range model. The last map, obtained with a linear (fractal) variogram, looks very good and is probably close to what an interpreter would have drawn by hand. Kriging with a linear variogram is closely related to the successful multiquadric mapping method of Hardy (1990).

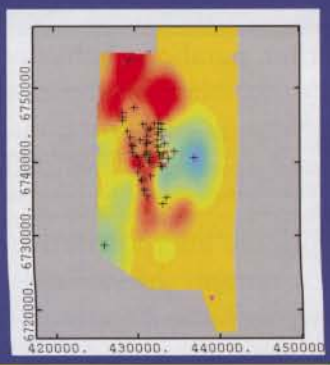
In all situations, as soon as the interpolated point becomes separated from the control points by a distance greater than the range, kriging becomes equal to the trend function. Because the variogram model says that there is no correlation between the unknown value at the interpolated point and any of the values at the data points, kriging cannot do better than use the trend value as the best estimate.

- Neighborhood considerations

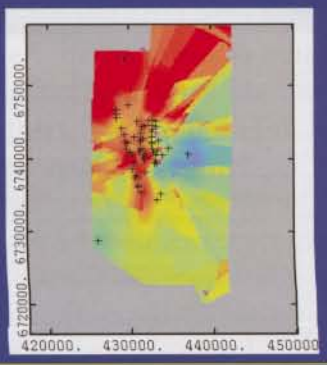
Kriging consists of estimating the value at each location using a weighted average of the data surrounding this location. In the six maps of the previous example, kriging was performed in the "global neighborhood." Because we only had 59 data points, we could use all the data points for each interpolation. The main advantage of working in the global neighborhood is that no discontinuity artifact appears in the data set. However, if we deal with the interpolation of thousands of seismic data points, it is impractical to invert the matrix associated with kriging in the global neighborhood. The required computer time and storage space would be enormous, and the inversion of the kriging system would become numerically unstable around a few thousand data points. This is why "moving neighborhood" kriging is used in such configurations.

In moving neighborhood kriging (Fig. 3-15), a limited subset of data (at least 24 for satisfactory results) is used to interpolate each grid point. A maximum search radius and a quadrant or octant selection are usually applied to ensure that data are not too far away and that they properly surround the estimated point. Nevertheless, when the moving neighborhood is too small, the resulting kriged map may look awful. Fig. 3-16 shows that this problem can be serious with 2D seismic data. Our recommendation, when dealing with fewer than 1000 data points, is to work in the global neighborhood. We will see later that this only requires the inversion of the kriging matrix once and for all. When this is not possible, it is very important to use computer packages that provide a clever neighborhood search and to use neighborhoods that are as large as possi-

THE IMPACT OF THE NEIGHBORHOOD USED (59 WELLS, NORTH ALWYN FIELD, NORTH SEA)

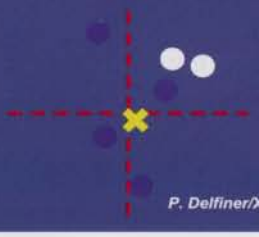


Global Neighborhood



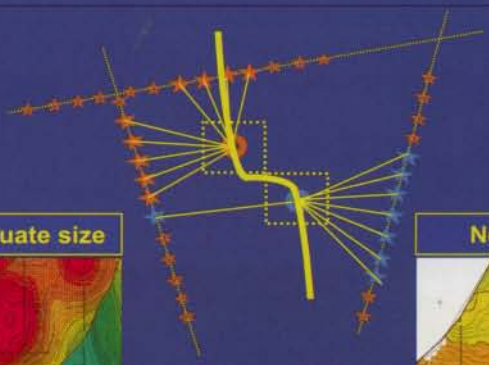
Moving Neighborhood

Data points
Point to be estimated

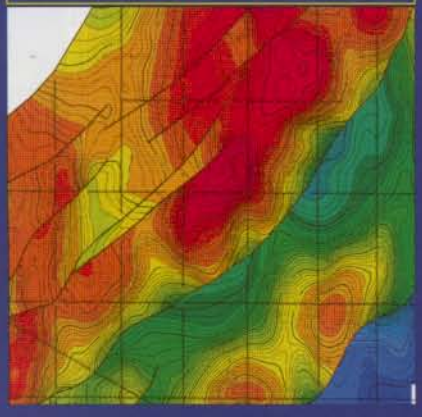


SEG/EAGE DISC 2003 P. Delfiner/X. Freulon 3-15

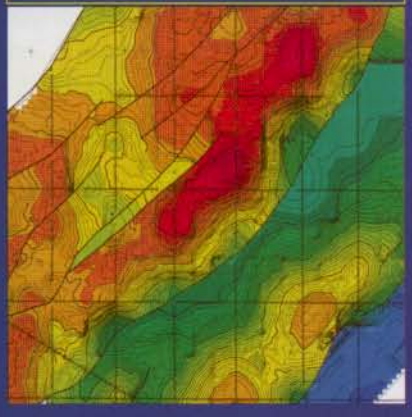
THE IMPACT OF THE NEIGHBORHOOD USED



Neighborhood of adequate size



Neighborhood too small

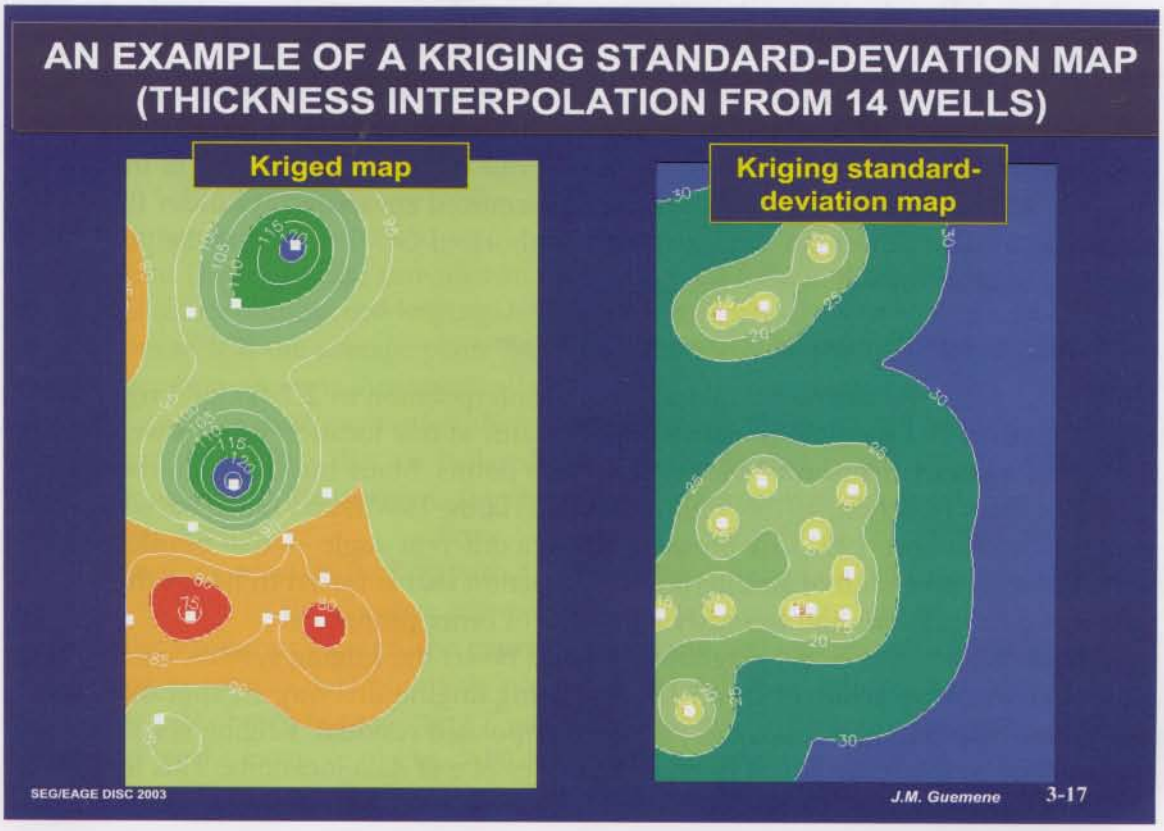


SEG/EAGE DISC 2003 V. Bigault de Cazanove 3-16

ble. For instance, a number of programs now calculate each kriging sequentially, starting with zones with no data and incorporating previously kriged values in the neighborhood of subsequently kriged points (Th. Coleou, personal communication, 2002). The extra computer time is well worth the improvement in the map.

- Kriging standard deviation

Not much has been said so far about the kriging standard deviation. Once the kriging weights have been found that minimize the estimation variance σ_E^2, this variance can be calculated at each kriged point. It is called the kriging variance σ_K^2. At the beginning of petroleum geostatistics, this property was somewhat oversold. Thanks to relationships such as that of Fig. 1-28, it was possible to translate the kriging standard deviation into a confidence interval. Assuming that kriging errors were normally distributed, the confidence interval was even equal to twice the kriging standard deviation (Fig. 1-26). Kriging was the only method that provided an estimate of the error associated with each interpolated value! Fig. 3-17 shows an example of a kriging standard-deviation map associated with a kriged thickness map. A stationary ordinary kriging is used, and, not surprisingly, the σ_K map shows bulls-eyes around the data points and slowly increases away from them. Because we are dealing with a stationary model, σ_K is constant as soon as the distance from all data points becomes greater than the variogram's range. It is easy to show that, if the variogram model is multiplied by a constant, σ_K^2 is multiplied by the same value, but the kriged map itself does not change. On the other hand, a change affecting the range or the nugget effect will affect the σ_K map. σ_K provides useful



information about the relative uncertainty affecting the estimates at various locations. We will also see that it plays a crucial role with conditional simulation.

3.2.3 *Generalized covariances of order* k

Matheron (1973) generalized the definition of stationary random functions to what he defined as "intrinsic random functions of order k" (k-IRF). This theory is more mathematically demanding than any we have seen so far, and we will simply discuss the philosophy behind it.

There are two main reasons for Matheron's generalization. The first is that, as explained before, the universal kriging model requires the decomposition of the variable $Z(\mathbf{x})$ into a trend and a stationary residual about this trend. This is satisfactory in cases where the trend is well defined over the whole area of interest, but it is more problematic when the trend cannot be modeled as a single polynomial over the entire field. Another reason to choose k-IRF is a bit more technical. We have seen that not just any function could be a covariance or a variogram and that the Fourier Transform of a covariance function must be positive. This limits our flexibility in the choice of covariance or variogram models.

Matheron (1973) realized that, in the universal kriging system, only a limited family of linear combinations of values $Z(\mathbf{x}_i)$ were considered — those "filtering" polynomial trends. This allowed Matheron to relax the constraints on stationary covariances and on variograms by defining what he called "generalized covariances of order k" (GC-k). A GC-k is such that only the variance of those weighted averages of values $Z(\mathbf{x}_i)$ that filter trends of degree k must be positive. A major advantage is that simple polynomial covariances are GC-k under certain conditions on their coefficients, which are given in Matheron (1973).

Remember that, when we introduced the variogram function, we insisted that it was associated with stationary increments of the variable $Z(\mathbf{x})$. An increment is a linear combination that filters out constant terms or terms of degree zero. Thus, in the formalism of k-IRF, variograms are associated with generalized covariances of order 0. We already know that power-law variograms are authorized GC-0 as long as the power of h is strictly smaller than 2.

3.2.4 *Kriging considered as an interpolating function*

Now, let us focus our attention on mapping, or interpolation in 2D. So far, kriging has been presented as a method for predicting the value at one location using a weighted average of measurements available at actual data points. Many interpolation methods, such as splines (Duchon, 1975), multiquadrics (Hardy, 1990), and radial basis functions (Franke and Nielson, 1991) are presented from a different angle — that of calculating the analytical expression of an interpolating function that is forced to honor the $z(\mathbf{x}_i)$ values at data locations and to satisfy a number of other properties.

If we work in the global neighborhood and invert the kriging system, we find that z_{uk} follows the relationship of Fig. 3-18. By writing kriging this way, z_{uk} appears simply as the sum of an interpolated trend plus an interpolated residual. Kriging is an exact interpolator, in the sense that it honors the values of z at data locations. This leads to as

KRIGING SEEN AS AN INTERPOLATING FUNCTION (FOR A LINEAR TREND IN 2D)

$$z_{uk}(x,y) = a_0 + a_1 x + a_2 y + \sum_{i=1,N} b^i \gamma\left[\sqrt{(x-x_i)^2 + (y-y_i)^2}\right]$$

Interpolated trend — Interpolated residual

N+3 conditions to determine the N+3 coefficients

$$\sum_{i=1,N} b^i = \sum_{i=1,N} b^i x_i = \sum_{i=1,N} b^i y_i = 0 \quad \& \quad z_{uk}(x_i, y_i) = z_i \text{ at N data points}$$

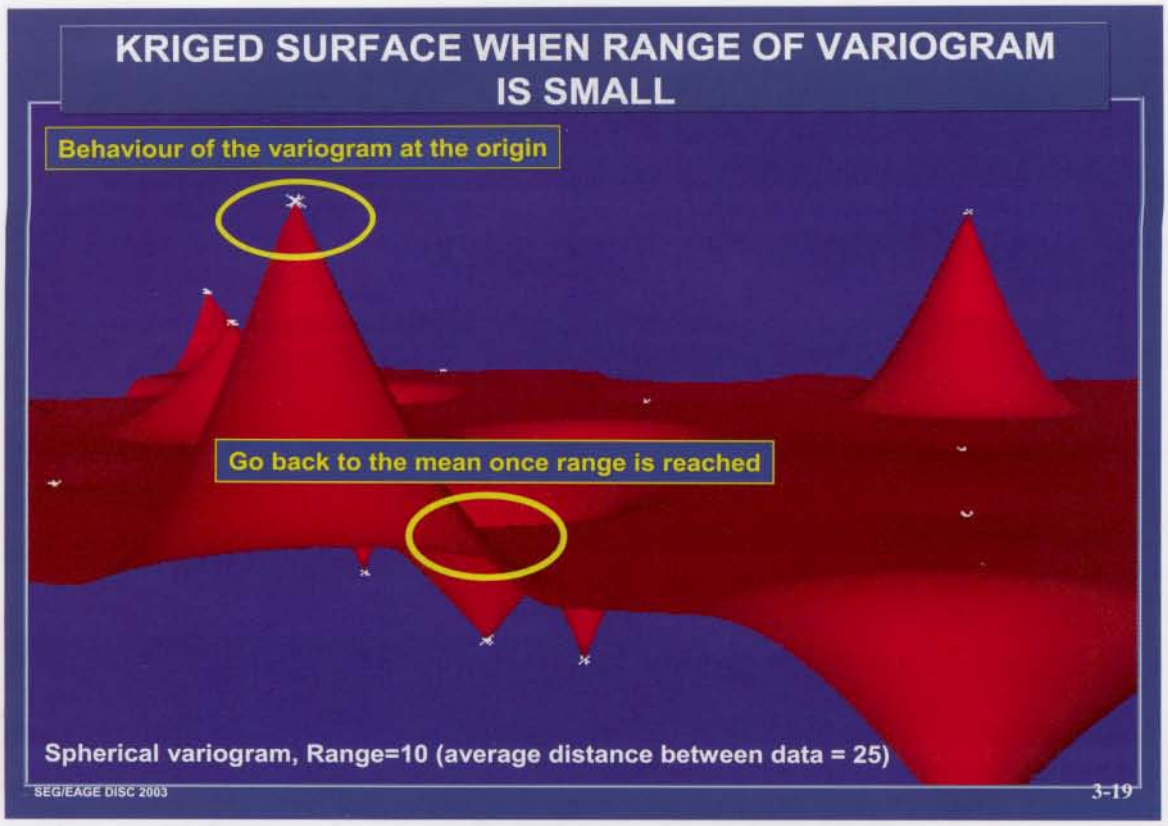
SEG/EAGE DISC 2003 3-18

many equations on the interpolating function's coefficients as there are data points. In 2D with a linear trend, there are three other equations that imply that, if the unknown parameter is itself a linear function of coordinates (a trend), kriging will be exactly equal to this function everywhere.

In the case in which the variogram is a pure nugget effect, it is easy to see that z_{uk} becomes exactly identical to the trend surface interpolation. This is expected, because the model behind trend surface analysis (Fig. 1-39) assumes that residuals from the trend are uncorrelated.

Coleou (1996, 2002a) provides an interesting discussion of the interpolated trend that is the result of universal kriging: he calls it "georegression." In the case where the trend is linear, it is the average plane fitted through the data points, taking spatial redundancy into account. Automatically, data points that are clustered — that is, are at a smaller distance from each other than the variogram's range — will be given less weight than data points that are isolated. Only when the data points are uncorrelated with each other will the georegression be equal to the standard statistical regression of $Z(x,y)$ against x and y.

What if we now are dealing with a stationary variogram model (finite range)? In this case, as soon as the distance between the interpolated point and all data points exceeds the variogram's range, the expression of z_{uk} reduces to that of the trend. This means that as soon as data points do not bring any statistical information, kriging makes the safe choice of predicting Z using the low-frequency term: the trend. This is illustrated by the example of Fig. 3-19, where the variogram range is smaller than the data spacing. This example also nicely illustrates the impact of variogram behavior at



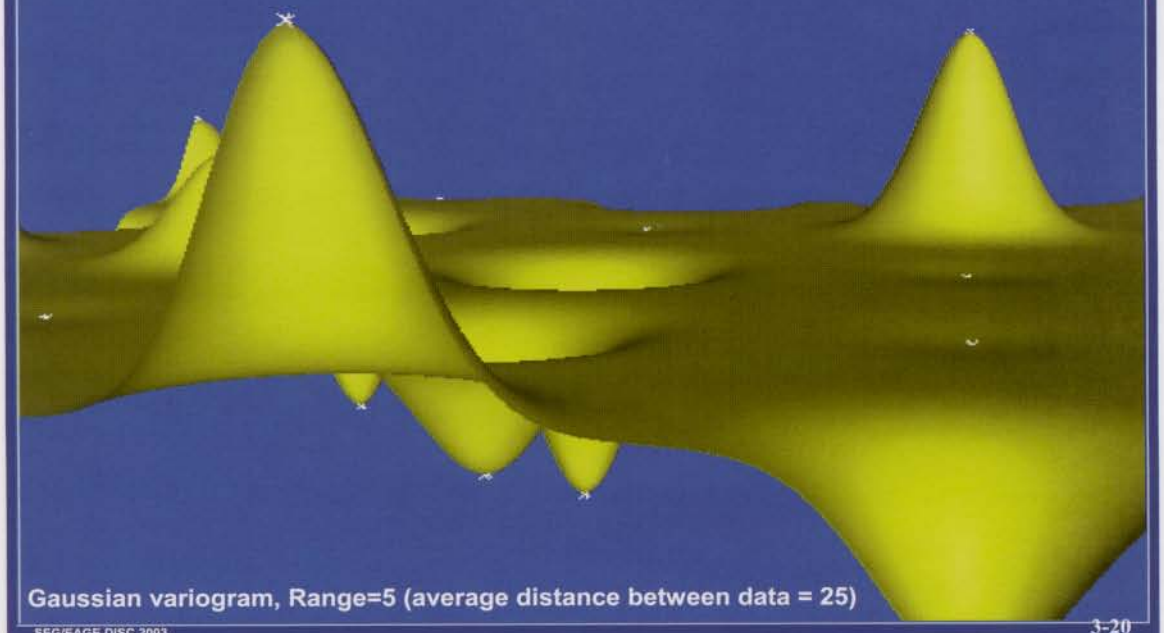
the origin. At the neighborhood of each data point, the shape of the interpolating function is nothing other than the shape of the variogram! This is confirmed by Fig. 3-20, obtained by keeping the same variogram range but by changing the variogram's behavior at the origin by that of a Gaussian model. As stressed by Oliver (1998), kriging applies a covariance-based convolution operator to the data. Kriging is a smoothing operator. Fig. 3-21 shows the impact of using a linear variogram with a constant trend. The kriging behavior close to the data points remains controlled by the linear behavior of the variogram at the origin, but there is no extrapolation toward the mean, which is not clearly defined because the model is not stationary. Note, however, that the shape of the interpolation functions is quite satisfactory and "natural," as already noted in the North Alwyn case (Fig. 3-14).

Fig. 3-22 confirms what happens when a trending variable is interpolated using the universal kriging model. On the left we see that, away from the data, extrapolation converges toward the linear trend, whereas on the right, a stationary model (with the same variogram) having a constant mean has been wrongly assumed. Away from the data points, kriging tries to come back to the constant mean. Experience shows that the main difference between kriging with a constant trend and kriging with a linear or a parabolic trend lies in extrapolating at a distance from the data points that exceeds the variogram range.

3.2.5 Cross-validation

Is there an objective means of evaluating the performance of the interpolating function?

KRIGED SURFACE WHEN RANGE OF VARIOGRAM IS SMALL

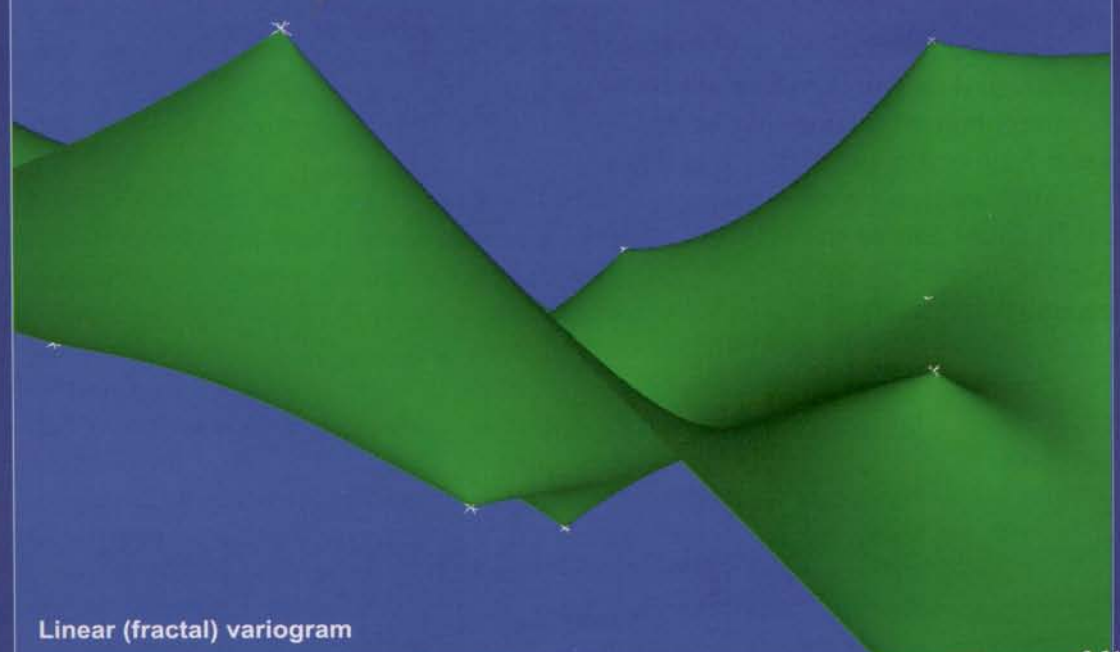


Gaussian variogram, Range=5 (average distance between data = 25)

SEG/EAGE DISC 2003

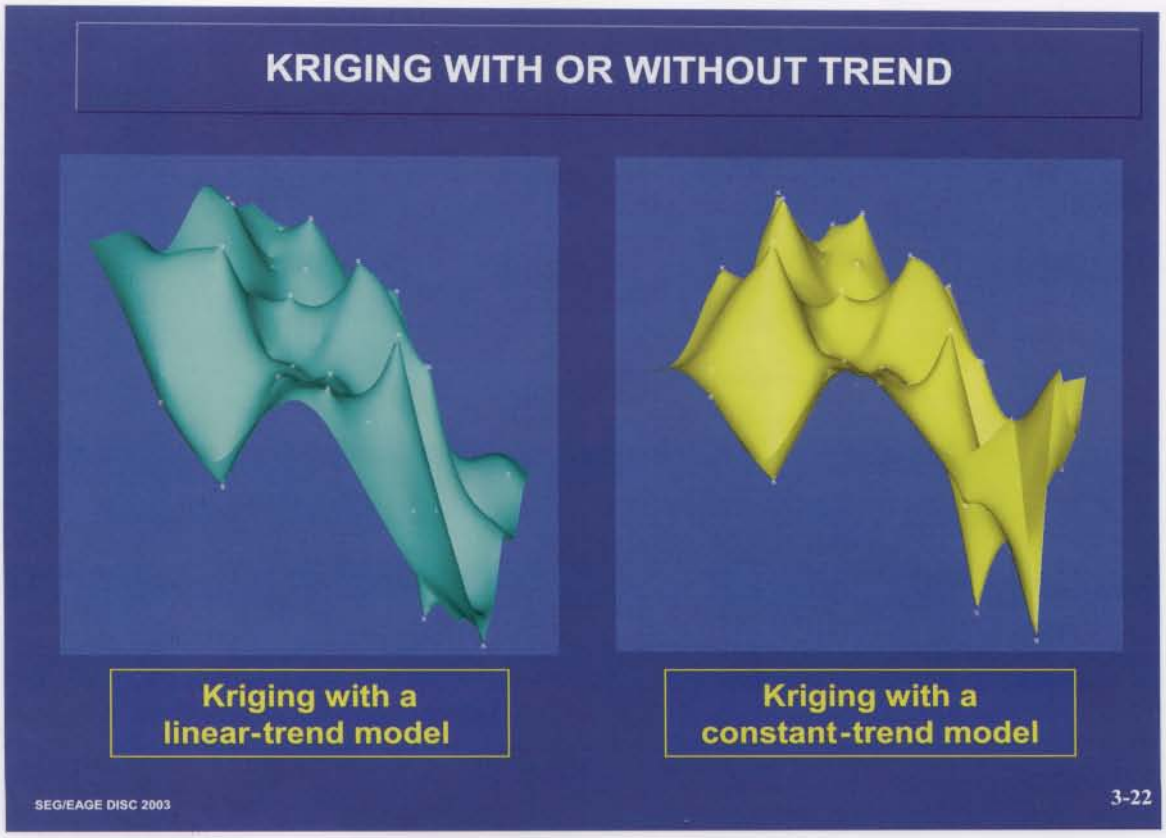
3-20

KRIGED SURFACE WHEN RANGE OF VARIOGRAM IS LARGE



Linear (fractal) variogram

3-21



The first reply could be that, if a geostatistician has fitted a trend and a variogram model to the data, and if there is a sufficient number of data points to make the model reliable, the kriging interpolator is optimal from a statistical point of view. In that case, kriging should be optimal in the least-squares sense, and the standard-deviation map should give an indication of the relative reliability of kriging at each location. However, this remains a result of the model, which may be wrong. Also, quite frequently there are not enough data points to fit a reliable trend and variogram model, and the choice of the interpolating function may be related to cosmetic considerations. The linear variogram is quite a popular choice in that situation.

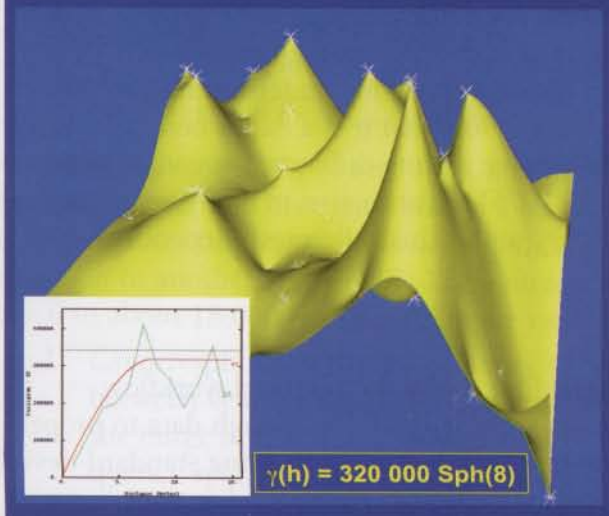
The user needs a more practical means of testing how well the interpolation is doing. The most natural test is that of comparing actual values at data points with those predicted by the model. Cross-validation (Fig. 3-23) consists of dropping each data point in succession and interpolating its value using the information at the other data points. Figs. 3-24 and 3-25 show the result in our example built from the Hohn data set. First, the best variogram fit is used, then two others are tested. In each case, we see the interpolated surface and the histogram of estimation errors obtained by successively dropping each of the 39 data points. The spherical model, as expected, performs well, and the difference with the linear model is small, whereas the Gaussian model performs more poorly.

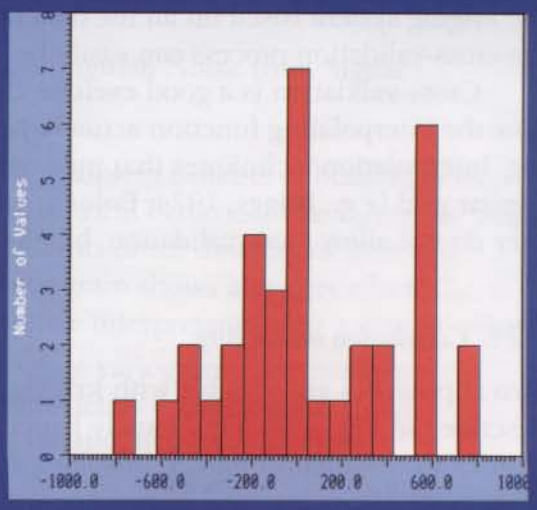
Cross-validation can be extended further. The kriging standard deviation itself can be cross-validated by comparing how closely it is related to the actual cross-validation errors. But the exercise is only meaningful in situations where the variogram model has been fitted to the data points.

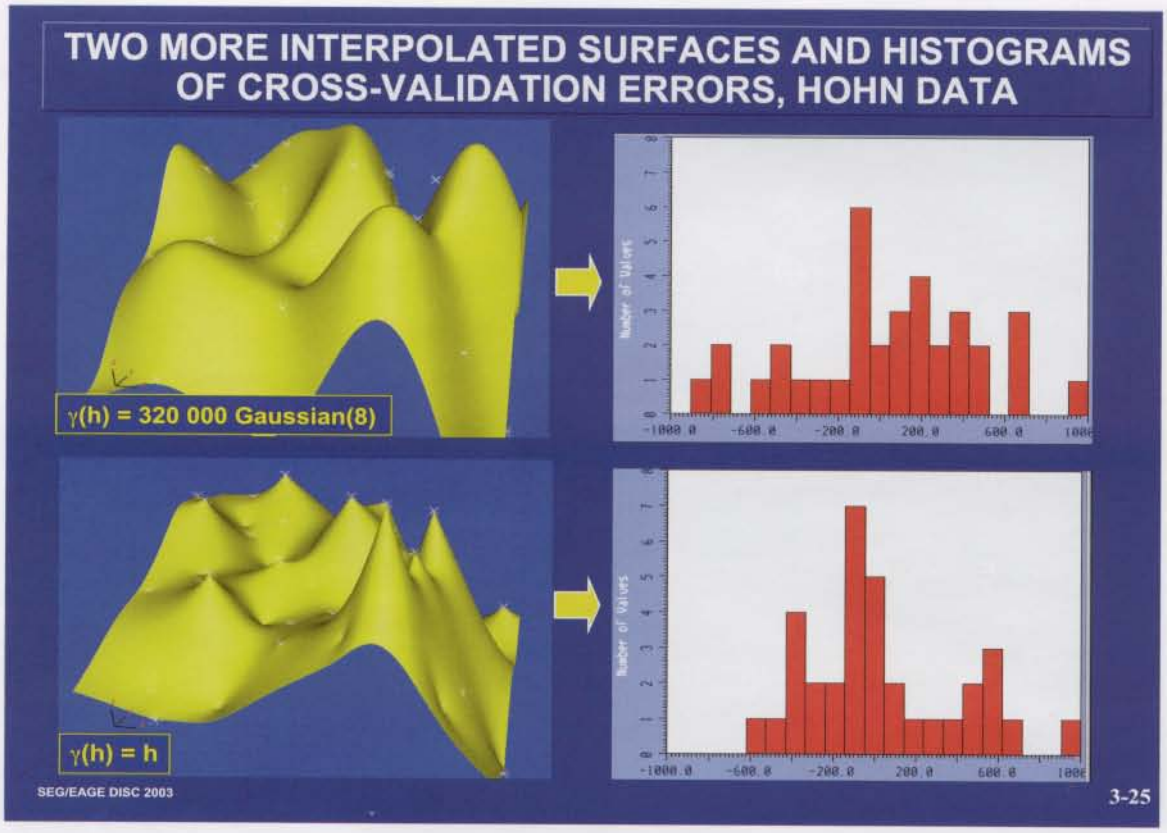
CROSS VALIDATION

- Yellow control point dropped from data set and estimated using the other control points.
- Estimated value can then be compared with actual value.
- Perform the same operation for all data points.

KRIGED SURFACE AND HISTOGRAM OF CROSS-VALIDATION ERRORS, HOHN DATA





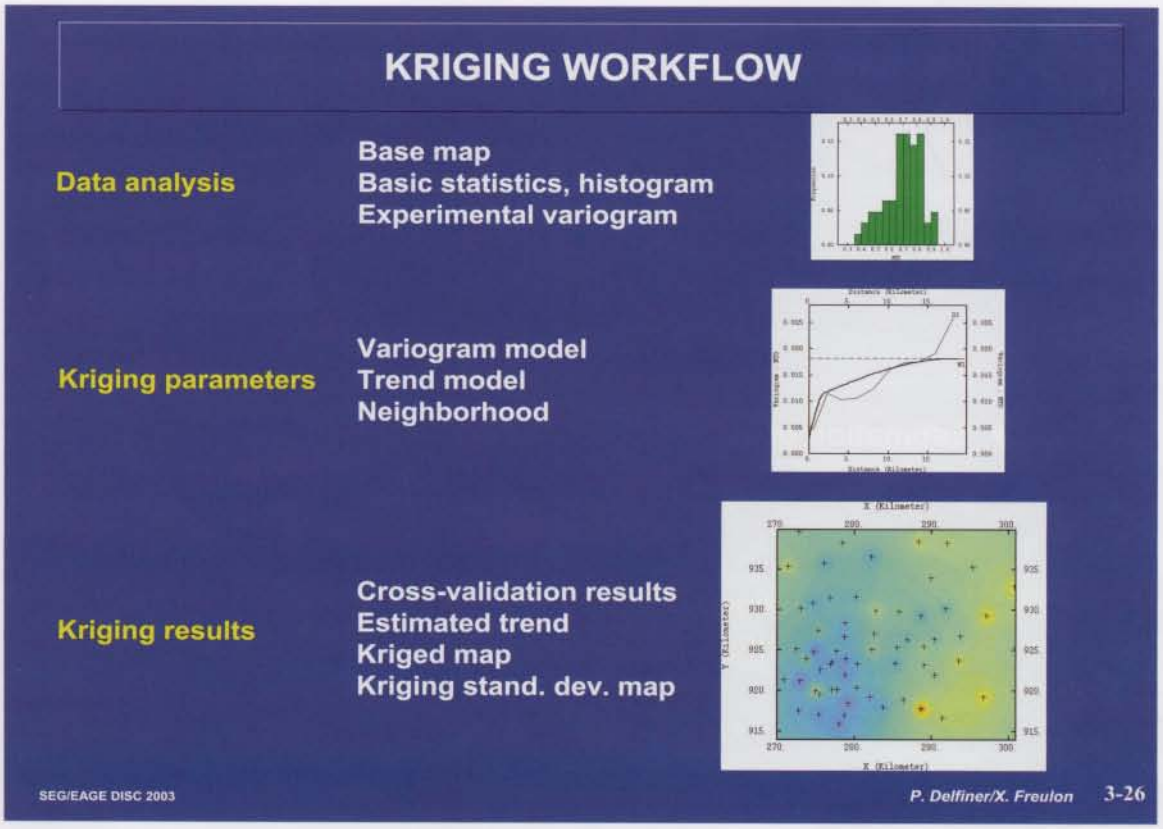


A limitation of cross-validation is that, if the number of data points is more than approximately 100 and fewer than 1000, we may want to work in the global neighborhood. As a result, re-estimating each data value from all the others requires inverting a large kriging system each time, which maybe prohibitive. Dubrule (1983) demonstrated that a more straightforward approach consists of inverting the matrix corresponding to the kriging system based on all the data points. Then, the estimate of each data point in the cross-validation process can easily be derived from this inverse matrix.

Cross-validation is a good exercise that provides precious understanding about how the interpolating function actually performs. Note that it is a major benefit of kriging. Interpolation techniques that numerically solve a finite-difference equation on a regular grid (e.g., Briggs, 1974; Bolondi et al., 1976) are attractive in many ways, but they do not allow cross-validation, because they are intrinsically mesh-dependent.

3.2.6 Conclusion on kriging

Two approaches are possible with kriging. The first one is the standard workflow described in Fig. 3.26. This usually happens as soon as there are enough data to properly infer a geostatistical model. In such a situation, results such as kriging standard-deviation maps are meaningful from a geostatistical standpoint. The second approach, usually applied when not enough data are available to build a reliable geostatistical model, consists of choosing the trend model and the covariance that appear to provide the best interpolation. "Best" can be interpreted according to various criteria, such as geological



realism, smoothness, or anisotropy. This use of kriging may be quite justified, because kriging provides a great flexibility, thanks to the choice of the trend and the covariance model. However, in this situation, the results of kriging cannot be regarded as optimal from the geostatistical point of view, and the kriging standard-deviation map is meaningless (Fig. 3-27).

3.3 Error Cokriging and Factorial Kriging to Distinguish Noise from Signal

3.3.1 Error cokriging for V_{stack} interpolation

Consider the case study of Fig. 3-28, in which the problem consists of building a map from V_{stack} data. The histogram shows a nice, symmetrical behavior. The variogram map is slightly anisotropic, which is not significant in terms of sill differences from one direction to another (Fig. 3-29). The isotropic variogram shows a nugget effect, C_0, equal to about 18,000 $(m/s)^2$. There are two possible interpretations for a nugget effect.

1. C_0 corresponds to the sill of a very short-range variogram. The distance between velocity data is too large to provide detailed information about the range of this model, which may be associated with very high-frequency spatial velocity variations (we know that no frequency information can be derived above the Nyqyst frequency). This model, where the variogram is the sum of several models of different ranges, will be discussed in the factorial kriging section.

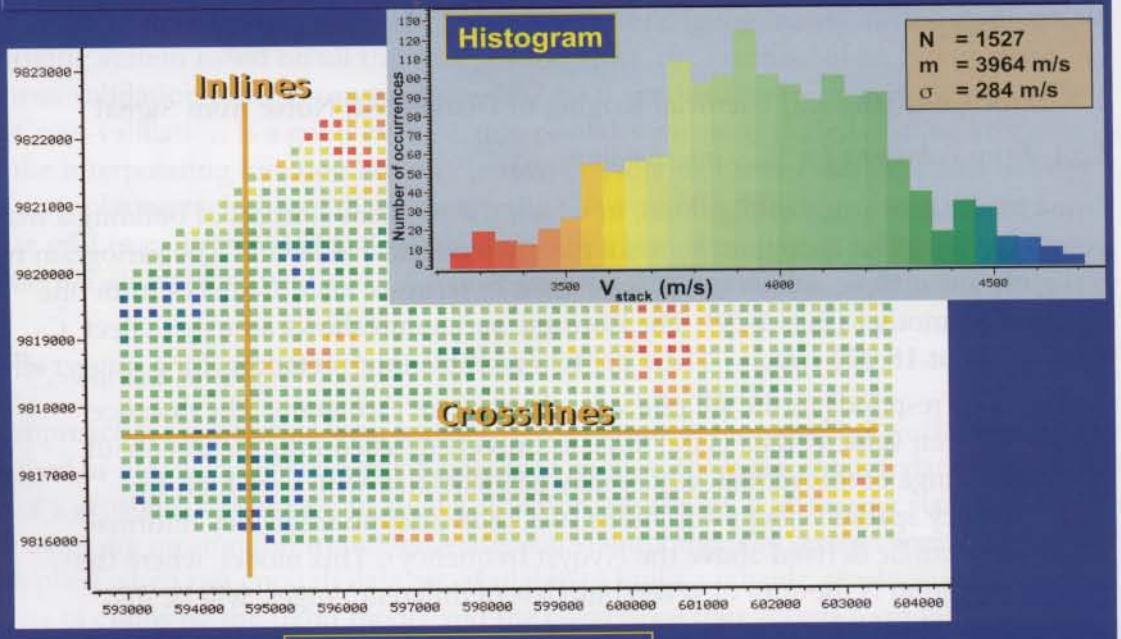
SOME CHARACTERISTICS OF KRIGING

- Honors the data points if no nugget effect $Z_k(\mathbf{x}_i) = Z(\mathbf{x}_i)$
- Estimation of standard deviation of the estimation error
- For interpolation to be optimal in the geostatistical sense, trend and variogram must be representative of the data

SEG/EAGE DISC 2003

3-27

AN EXAMPLE OF ERROR COKRIGING DATA ANALYSIS

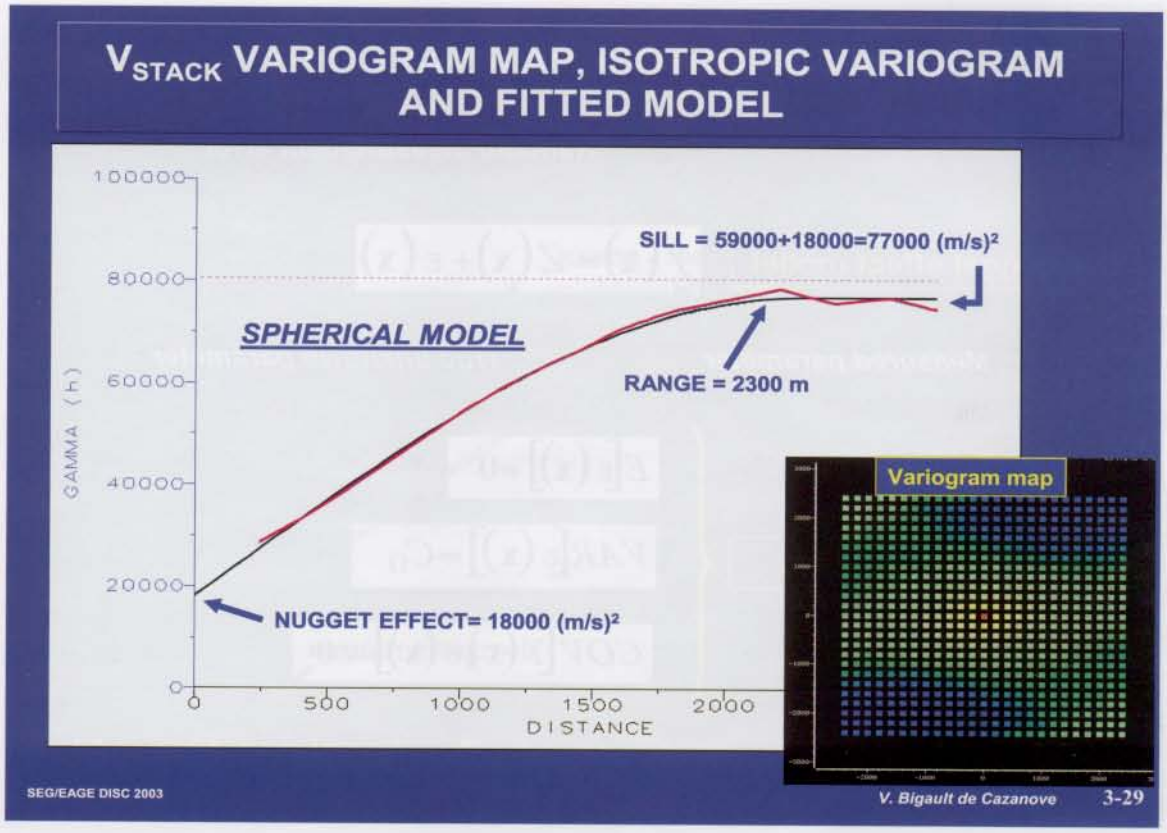


Map of Raw V_{stack} data

SEG/EAGE DISC 2003

V. Bigault de Cazanove

3-28



2. There is no short-range variation in velocity, but data are affected by a measurement error, which directly causes a discontinuity in variogram behavior at the origin. The model is described in Fig. 3-30. It is easy to show that, if the variable of interest $Z(\mathbf{x})$ is affected by a measurement error, $\varepsilon(\mathbf{x})$, of mean zero and variance C_0, spatially uncorrelated and independent of velocity, this random error translates into a C_0 nugget effect on the variogram. This is the error cokriging model, thus called because $Z(\mathbf{x}_0)$ will be interpolated using values of the "other" variable $Y(\mathbf{x}_i)$ — that is, $Z(\mathbf{x}_i)$ plus the unknown error $\varepsilon(\mathbf{x}_i)$ at data points $\mathbf{x}_i$.

With stacking velocities, which are known to be error-prone, the second interpretation usually applies. How will this interpretation affect kriging? Fig. 3-31 shows the map obtained and a cross-section associated with this map. Contrary to what would happen with a zero nugget effect, velocity data are not honored, and the map appears to be quite smooth: Error cokriging filters Y and tries to interpolate Z instead. Without dwelling on mathematics, let us simply mention that error cokriging amounts to adding to the diagonal term of the kriging matrix (Fig. 3-3) a constant term equal to the nugget effect.

What if we now use the first variogram interpretation? The variogram model shown in Fig. 3-32 is fitted. As a result, kriging considers it is dealing with a variable not affected by measurement errors. The map obtained is shown in Fig. 3-33. This time, all V_{stack} data are honored, and, as a consequence, the map is much more noisy!

As a matter of curiosity, we have also plotted the map obtained with a pure nugget

NUGGET EFFECT INTERPRETATION (1)

Random measurement error

Nugget effect model $Y(\mathbf{x}) = Z(\mathbf{x}) + \varepsilon(\mathbf{x})$

Measured parameter *True unknown parameter*

Properties of error ε

$$E[\varepsilon(\mathbf{x})] = 0$$

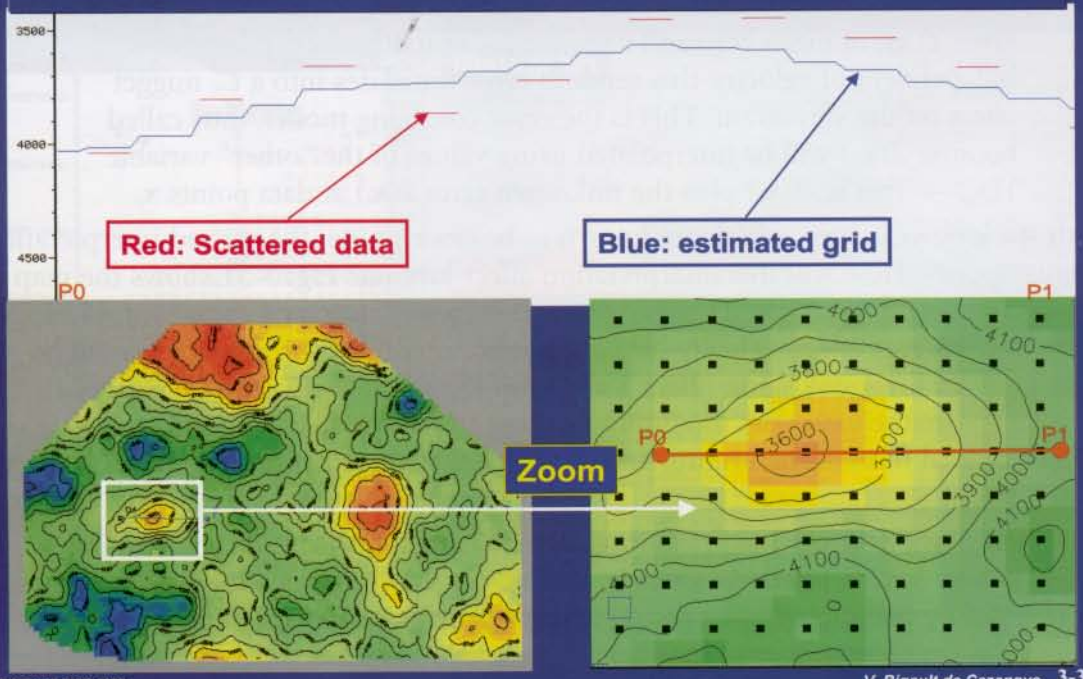
$$VAR[\varepsilon(\mathbf{x})] = C_0$$

$$COV[Y(\mathbf{x}), \varepsilon(\mathbf{x})] = 0$$

With error cokriging, we estimate Z using the values of Y!

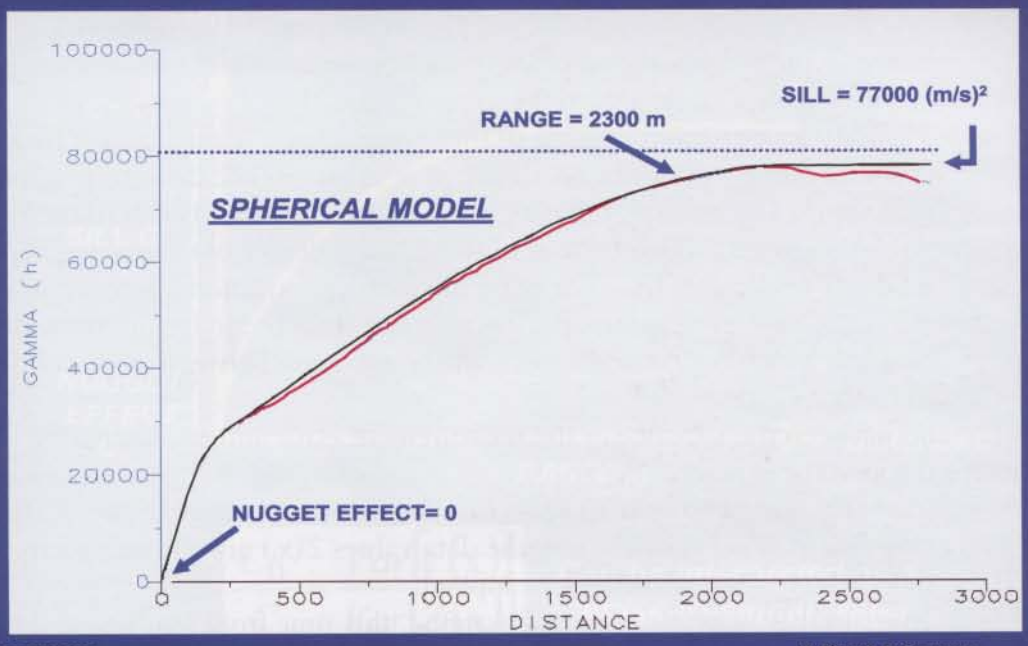
SEG/EAGE DISC 2003 3-30

RESULTS OBTAINED WITH NUGGET EFFECT SPHERICAL MODEL



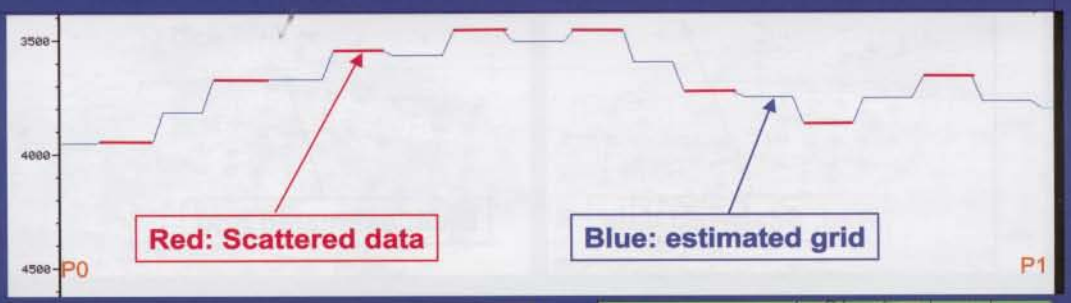
SEG/EAGE DISC 2003 V. Bigault de Cazanove 3-31

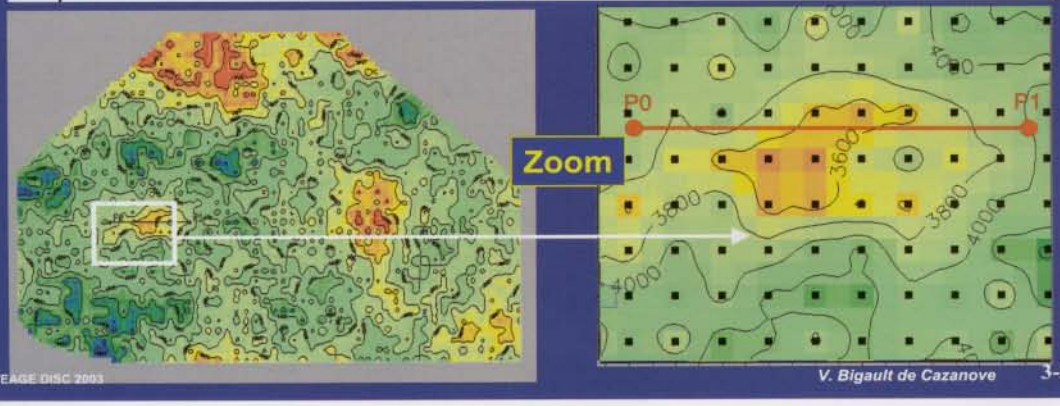
FORCING A ZERO NUGGET EFFECT ON THE VARIOGRAM MODEL



SEG/EAGE DISC 2003 V. Bigault de Cazanove 3-32

RESULTS OBTAINED WITH NO NUGGET EFFECT SPHERICAL MODEL





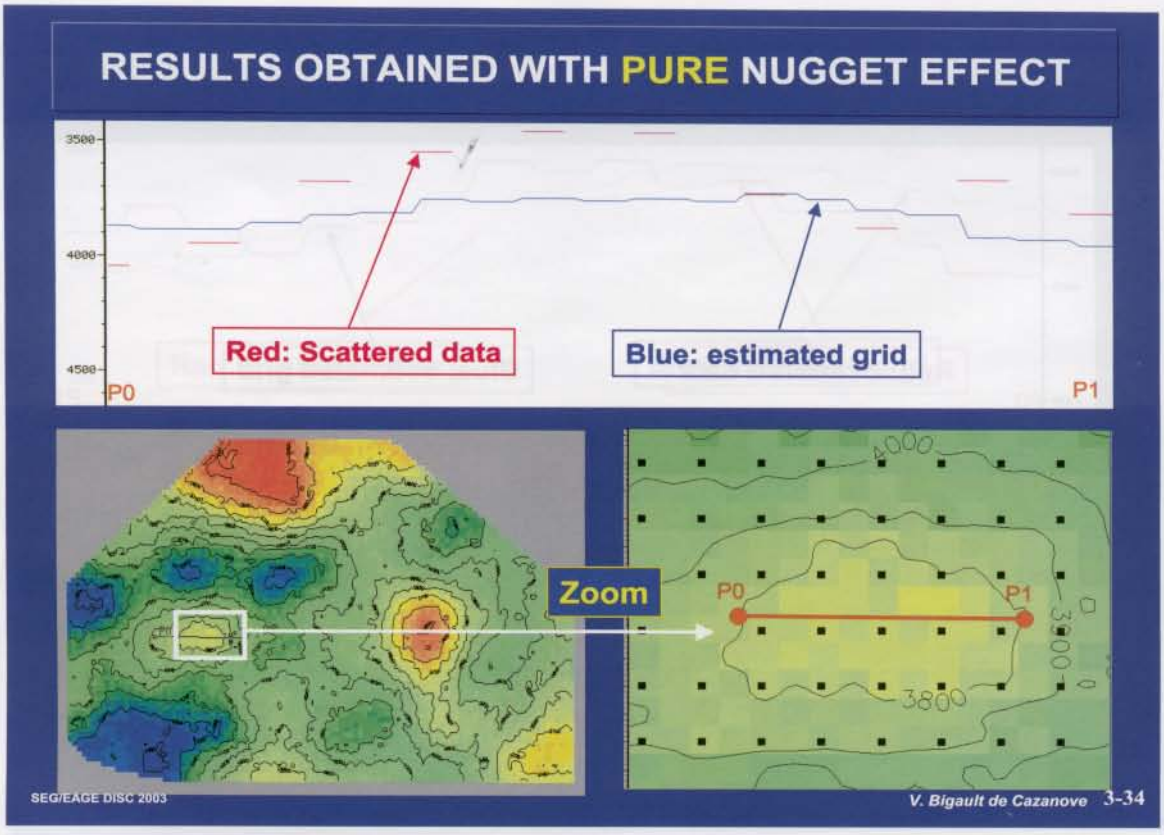
SEG/EAGE DISC 2003 V. Bigault de Cazanove 3-33

effect (Fig. 3-34). This means that there is no spatial correlation at all between V_{stack} values. A data point that is in the close neighborhood of the interpolated values has no more weight on this interpolation than a data point that is far away. The kriged value is a moving average of the data points in the kriging neighborhood. We are back to trend-surface analysis with a constant mean! The wiggles in the contour lines are due to the fact that a change in the neighborhood points translates into a significant change in the interpolated value, because all data have the same weight.

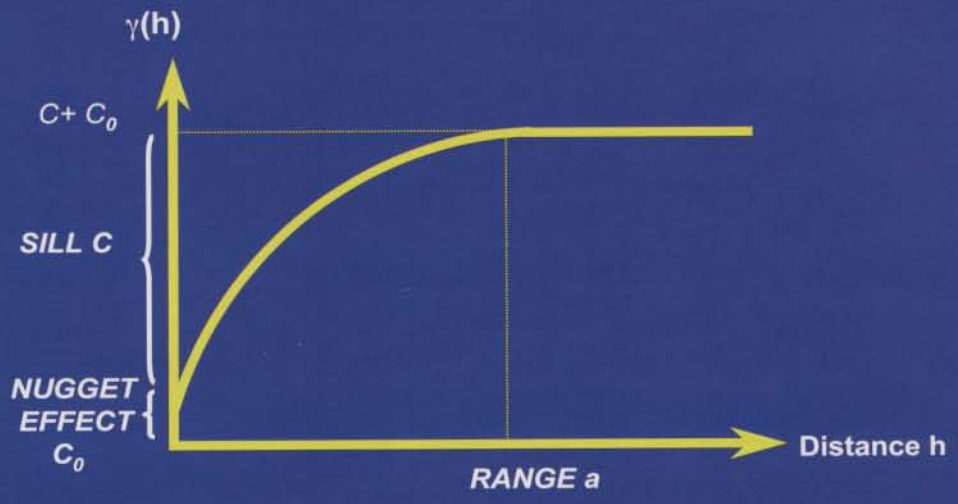
Fig. 3-35 discusses our interpretation of the nugget effect. A zero nugget means no noise, and a pure nugget effect means no signal, pure noise. Hence the interpretation of "noise-to-signal" ratio. The beauty of the approach is that this ratio is determined directly from the variogram model. Instead of fudging the degree of smoothing versus the degree of fidelity to the data (a major issue in many optimization approaches, as we will see below), error cokriging automatically calibrates this "fudge factor" from the data. This point will be addressed again when the relationship between error cokriging and smoothing splines is discussed.

The ratio between the variance of the measurement error and the variogram sill determines the amount of smoothing applied by error cokriging. If this ratio is low, the method assumes that the measurement error can be neglected in comparison with the variations of the interpolated variable, and the data values $Z(\mathbf{x}_i)$ are almost exactly honored. If the ratio is high, then smoothing is applied.

Fig. 3-36 is another example of error cokriging, this time from Mathieu and Nutt (1985). Again and again, this approach proves attractive for mapping stacking veloci-



NUGGET EFFECT INTERPRETATION (2)



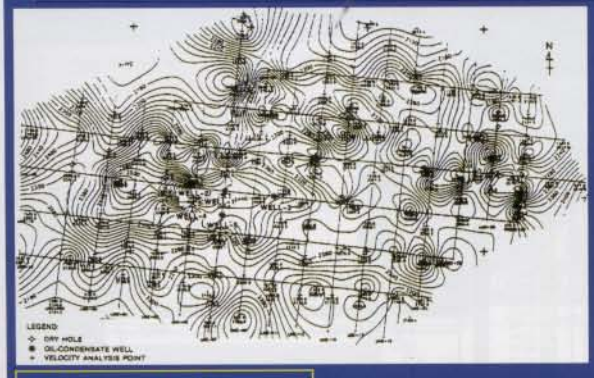
$$\frac{C_0}{C} = \frac{Var\left[\varepsilon(\mathbf{x})\right]}{Var\left[Z(\mathbf{x})\right]}$$

Can be interpreted as a "noise-to-signal" ratio

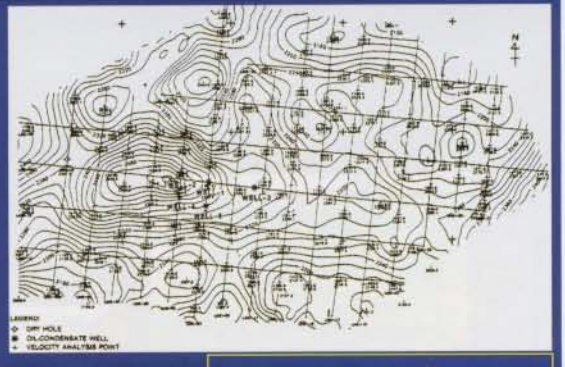
SEG/EAGE DISC 2003

3-35

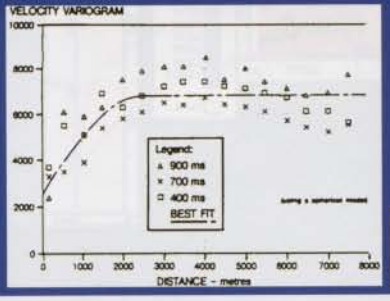
ANOTHER EXAMPLE OF ERROR COKRIGING FOR VELOCITY INTERPOLATION (MATHIEU AND NUTT, 1985)



Initial velocity map



Filtered velocity map



Stacking velocities variogram

SEG/EAGE DISC 2003

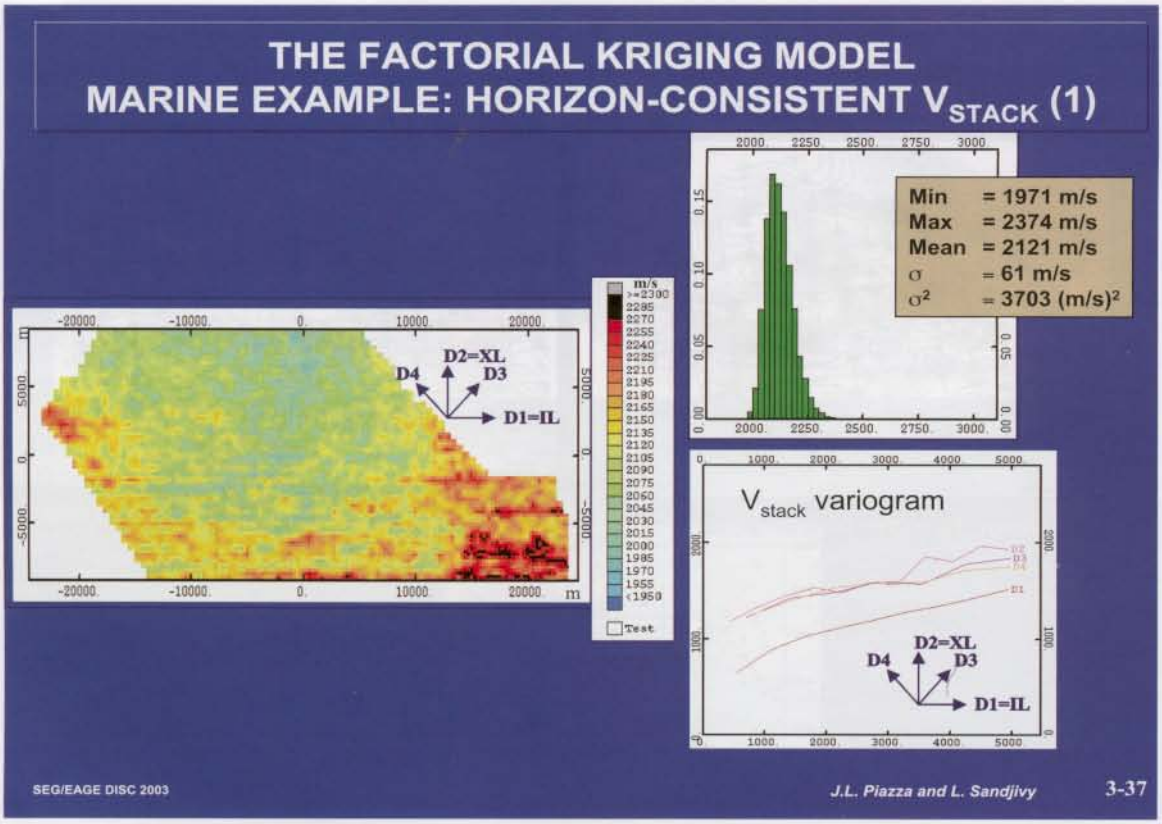
3-36

ties. An example from Dubrule and Haldorsen (1984) also shows a successful application of the technique to the joint interpolation of core- and log-derived permeability values at wells. The variance of the measurement error can also be a function of location. This can be a useful assumption when data resulting from two seismic campaigns of different quality are combined (Haas and Viallix, 1974). Other applications will certainly appear in the future.

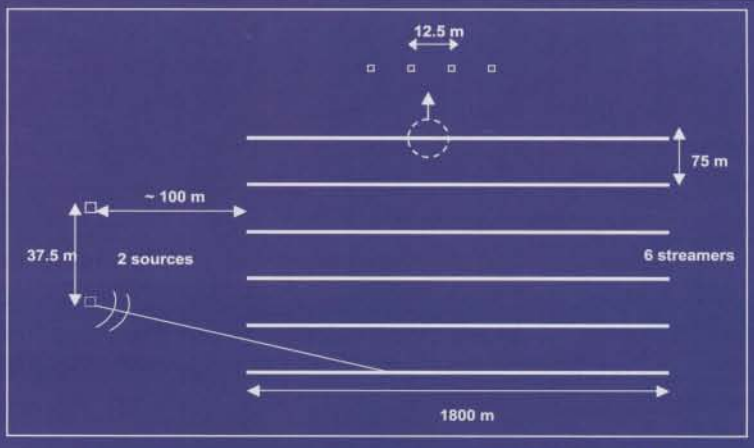
Other Fourier-based techniques allow the filtering of measurement errors. However, they usually require that the variable is first interpolated on a regular grid and then filtered. Error cokriging offers the advantage of performing the two tasks at the same time — interpolation of the data over a regular grid, and filtering of the measurement error. Error cokriging can also be applied in the context of nonstationary data.

3.3.2 *Factorial kriging*

Now let us come back to the first possible interpretation of the nugget effect, as a short-range variogram. Fig. 3-37 presents a marine example of horizon-consistent stacking-velocity data, and Fig. 3-38 presents the acquisition setup. The variogram clearly shows a different behavior between the inline and the other directions. The geostatistical analysis assumes that the variogram model corresponding to the geological signal (Fig. 3-39) is isotropic and equal to the sum of the linear model and a spherical one. The other assumption is that the difference between the experimental variograms is only due to acquisition artifacts or "inline effects." An anisotropic model is fitted to explain the difference between the inline and the three other directions.

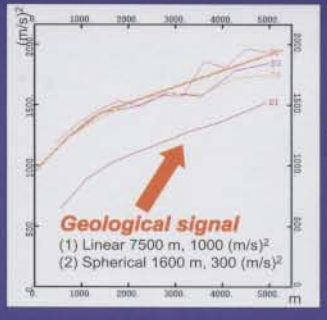


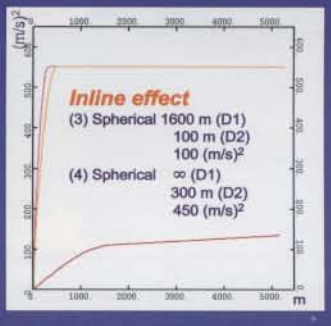
THE FACTORIAL KRIGING MODEL
MARINE EXAMPLE: HORIZON-CONSISTENT V_{STACK} (2)



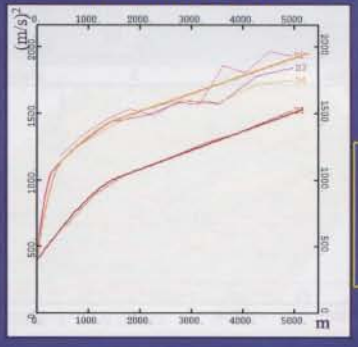
Sources number	2
Distance between sources	37.5m
Streamers	6
Groups of receptors per streamer	144
Distance between groups	12.5m
Active cable length	1800m
Distance Inline (Offset Inline)	12.5m
Distance between streamers (Offset Crossline)	75m
Azimuth Inline	140.3°

THE FACTORIAL KRIGING MODEL
MARINE EXAMPLE: HORIZON-CONSISTENT V_{STACK} (3)





Final model



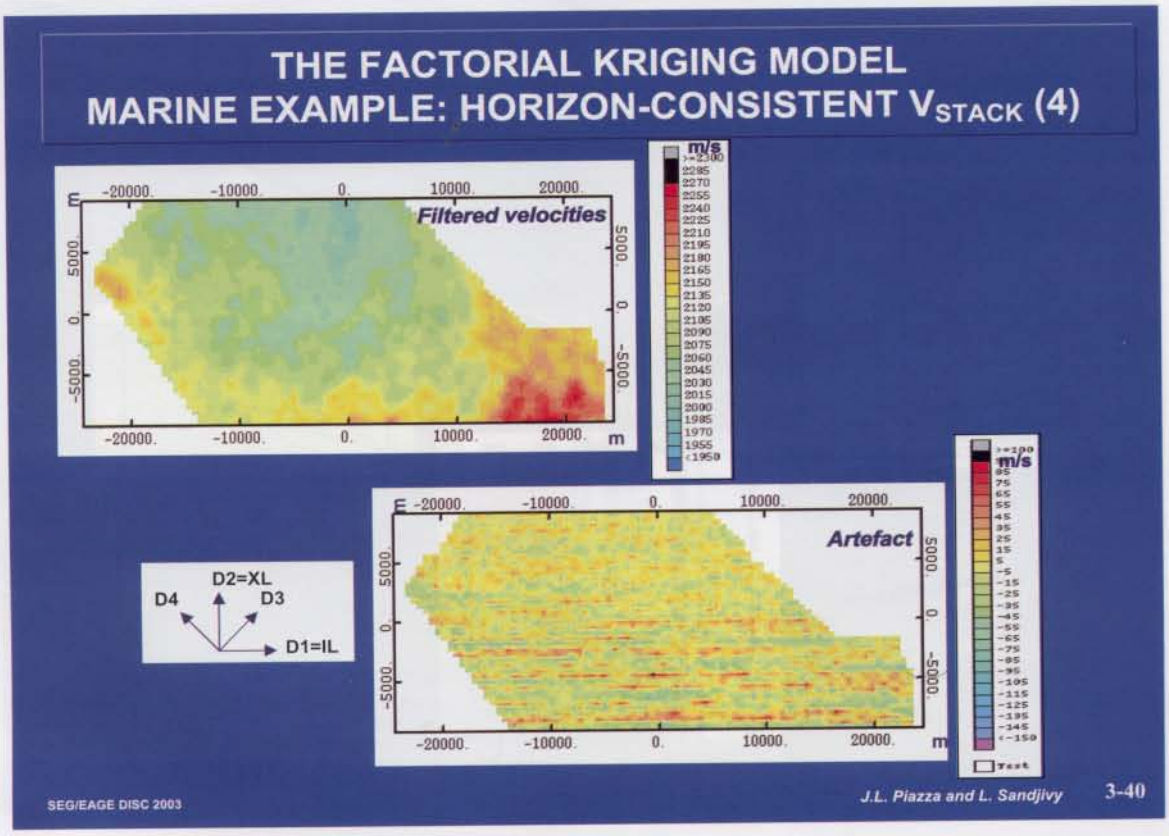
artifacts
- (1) Linear 1000 (m/s)²
- (2) Spherical 300 (m/s)²
- (3) Spherical 100 (m/s)²
- (4) Spherical 450 (m/s)²
- (5) Nugget 400 (m/s)²

The interpretation behind this model is that stacking velocity $Z(\mathbf{x})$ is the sum of two uncorrelated components: an isotropic geological signal and an anisotropic inline effect (Sandjivy, 1987; Piazza et al., 1997). As with error cokriging, we need to filter the nongeological effect out of the interpolation. The result obtained by this filtering is shown in Fig. 3-40. The validation of the two different maps will be obtained by reviewing the picked stacking velocities.

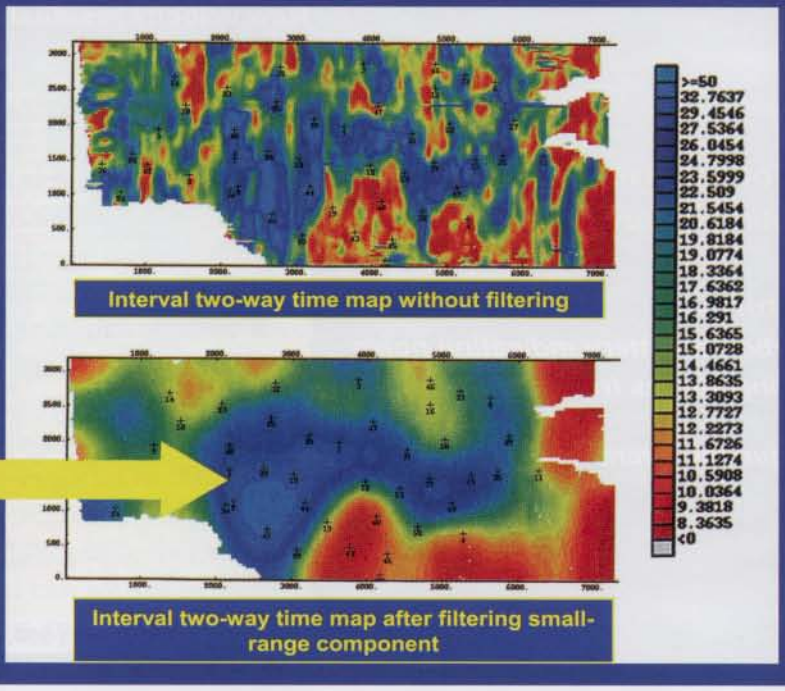
Fig. 3-41 shows another example of factorial kriging, by Mundim et al. (1999). Geological features, such as the main channel, are much more apparent after factorial kriging. A related benefit is that, after factorial kriging, there is a much better relationship between two-way time and reservoir facies thickness (Fig. 3-42).

In the example of Fig. 3-39, seismic data were affected by both a measurement error (uncorrelated noise) and correlated footprint effects. What does this actually mean? In the example of Fig. 3-43, stacking velocity is modeled as a linear function of two-way time plus a correlated residual. Fig. 3-43 is a map of this residual, which has a mean of zero. This residual also shows high-frequency noise characterized by the "checkerboard effect" and stripes associated with acquisition footprints. There is also a lower-frequency component. A moving-average filter is first applied to remove nongeological effects (Fig. 3-44). This filter is not completely successful in removing the stripes, and it may have a tendency to oversmooth the residual spatial variations.

Now the geostatistical analysis reveals a variogram map (Fig. 3-45) with a slight anisotropy. The range in the NW-SE direction is smaller than that in the NE-SW direction. An analysis of the autocorrelation function is even more interesting. For both the



GENERALIZING ERROR COKRIGING: AN APPLICATION OF THE FACTORIAL KRIGING MODEL (MUNDIM ET AL., 1999)

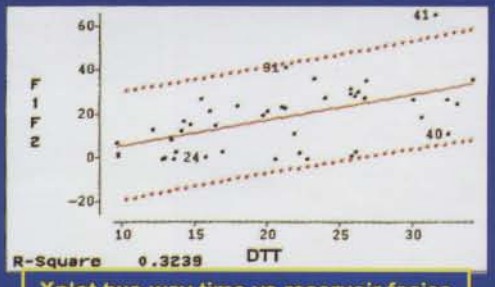


Interval two-way time map without filtering

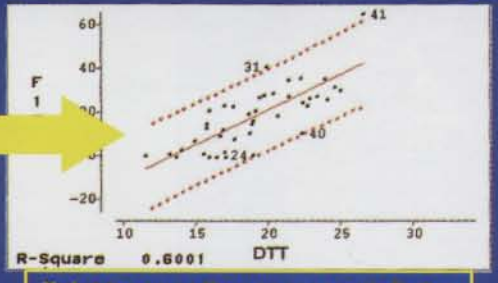
Interval two-way time map after filtering small-range component

Turbidite channel is much better imaged!

GENERALIZING ERROR COKRIGING: AN APPLICATION OF THE FACTORIAL KRIGING MODEL (MUNDIM ET AL., 1999)



Xplot two-way time vs reservoir facies thickness before filtering

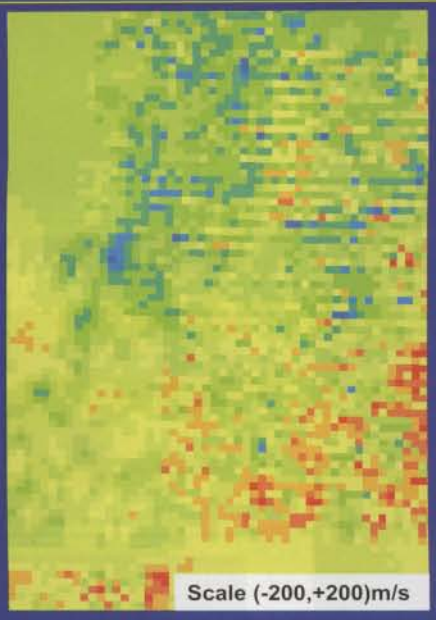


Xplot two-way time vs reservoir facies thickness after filtering

Correlation is much improved!

A VELOCITY SMOOTHING EXAMPLE (COLEOU, 2001)

Raw velocity residuals (V_{stack}-aT-b)



Compaction removed
Extrapolated to prevent edge effects

Three features:

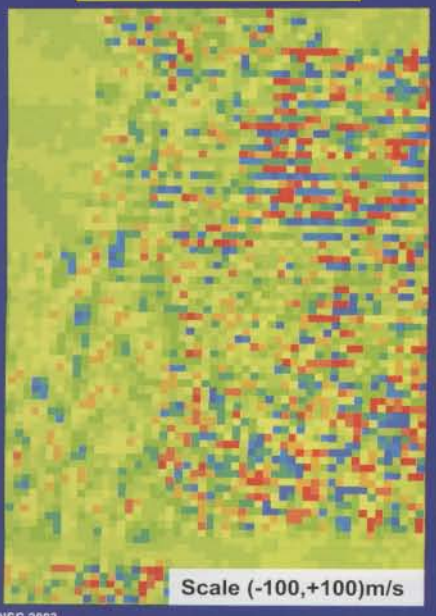
- *checkerboard pattern indicating noise*
- *alternating stripes from acquisition imprint*
- *low-frequency lateral variations*

SEG/EAGE DISC 2003

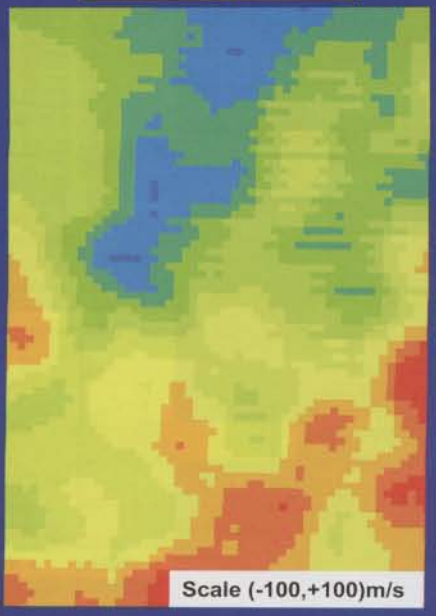
3-43

MOVING AVERAGE HIGH-LOW FREQUENCY FILTER (COLEOU, 2001)

High frequencies



Low frequencies



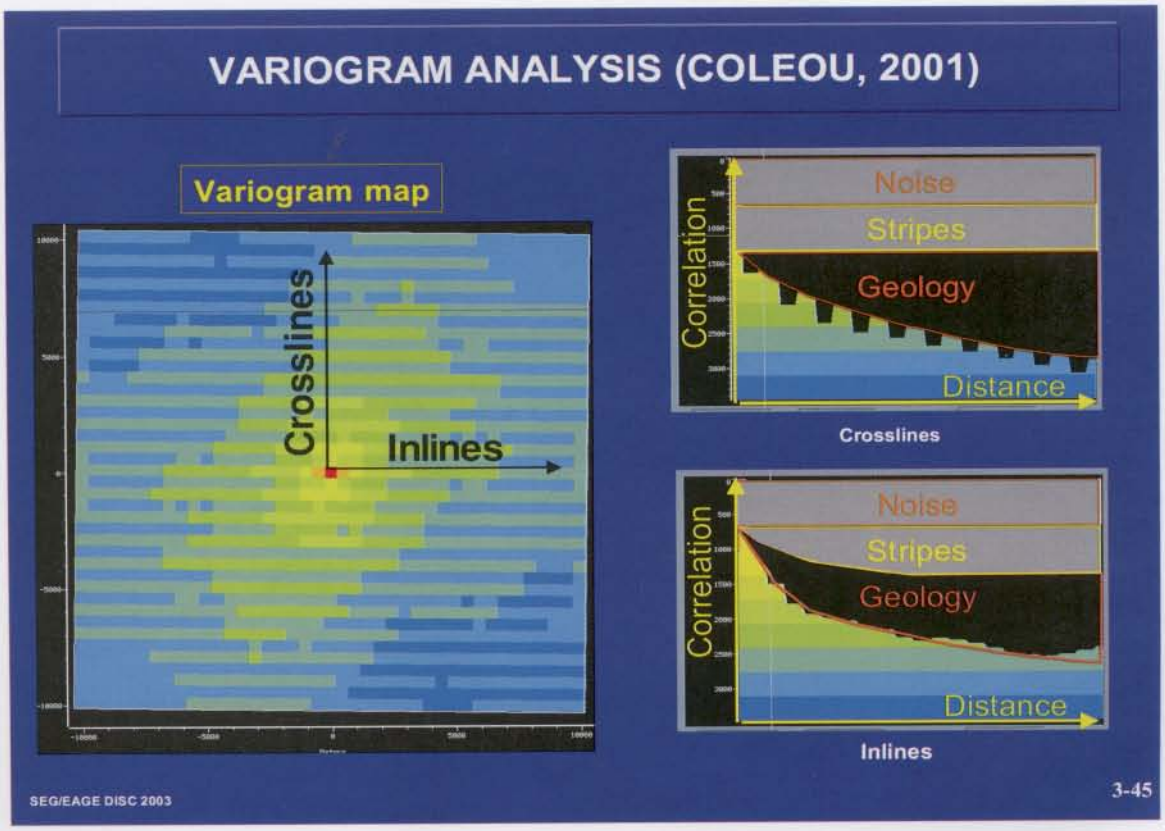
SEG/EAGE DISC 2003

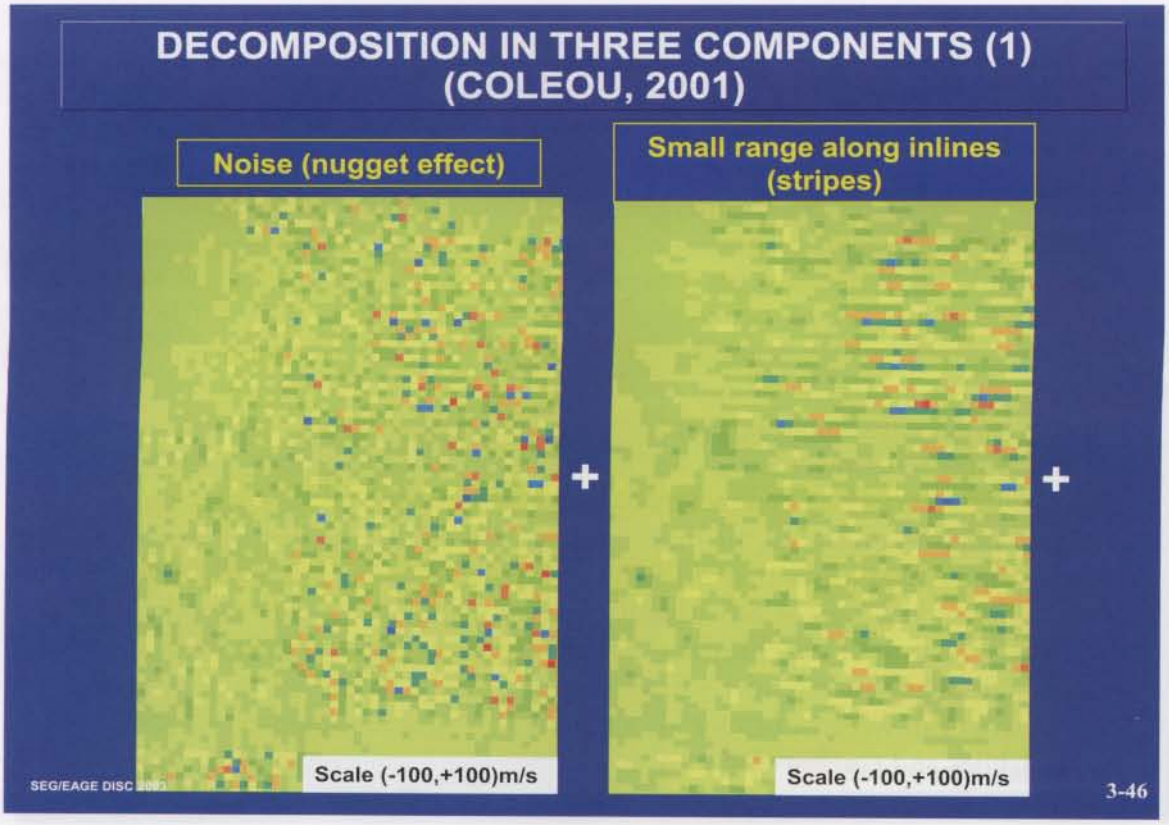
3-44

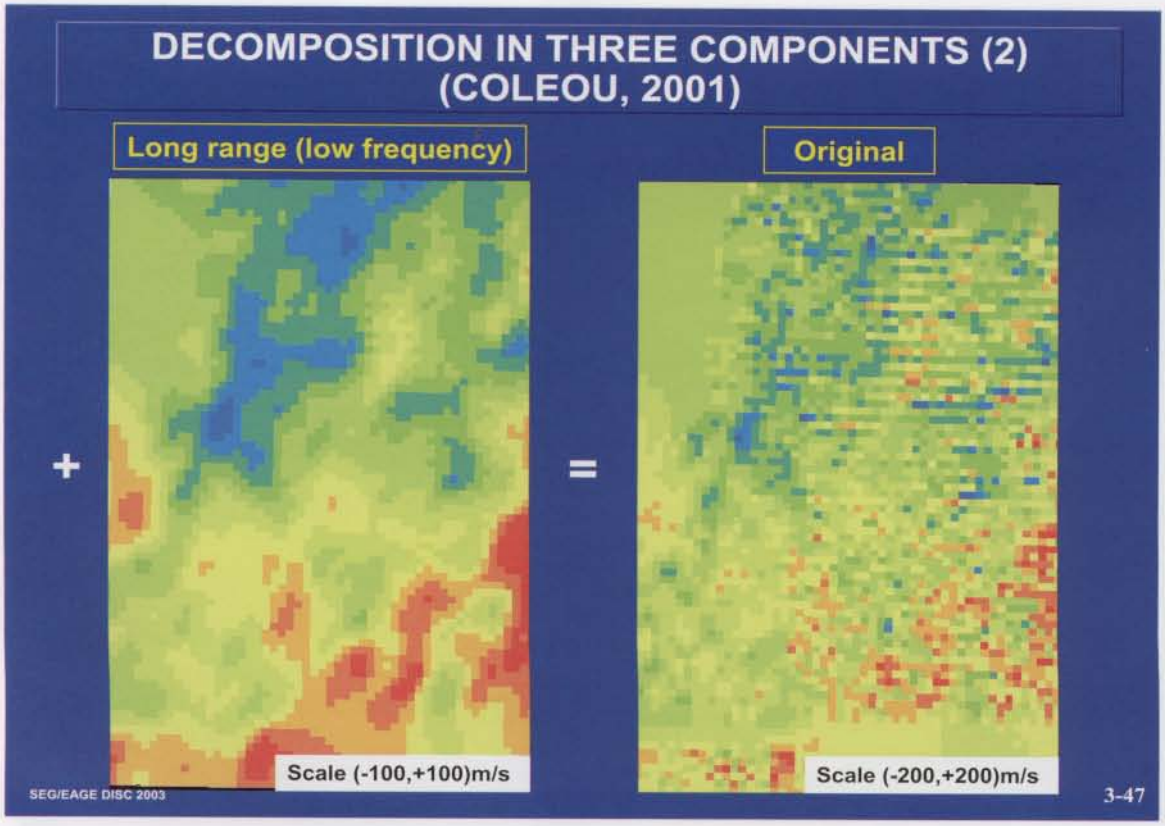
inline and crossline functions, noise characterizes itself as a nugget effect, whereas footprints have a different impact in both directions. The autocorrelation function along crosslines shows a nugget effect due to the random variations as one moves from one inline to another (acquisition footprints). The third component is related to geology. The variogram model is decomposed into these three components and then the velocity residual is reinterpolated, but this time, thanks to the variogram decomposition, the three different components are separated (Figs. 3-46 and 3-47). The low-frequency component is to be compared with the result of the moving-average filter of Fig. 3-44. The geostatistical low-frequency component contains higher frequencies than does the moving-average result. This is not the result of an arbitrary choice of the interpreter, but is the result of the geostatistical analysis, which has produced an estimate of the relative variance of noise, acquisition footprint, and low-frequency component. As a result, kriging has been able to define what is correlated and noncorrelated noise. Fig. 3-48 summarizes the differences between factorial kriging and moving-average filtering.

Fig. 3-49 gives a flavor of the mathematics behind factorial kriging. Actually, the kriging matrix is the same as that with universal kriging, because it is a function of the full covariance model. On the other hand, the right-hand side is only a function of the component of interest of the covariance. Note that the system does not incorporate any Lagrange multiplier here, because we are mapping residuals that have a zero mean.

Now what can we say of factorial kriging versus spectral approaches? This is discussed in Chilès and Guillen (1984), who compare the results obtained with factorial kriging and spectral analysis on French gravimetry data. Both variogram and frequency




DECOMPOSITION IN THREE COMPONENTS (1)
(COLEOU, 2001)
Noise (nugget effect)
Small range along inlines
(stripes)
+
+
Scale (-100,+100)m/s
Scale (-100,+100)m/s
SEG/EAGE DISC 2003
3-46


DECOMPOSITION IN THREE COMPONENTS (2)
(COLEOU, 2001)
Long range (low frequency)
Original
+
=
Scale (-100,+100)m/s
Scale (-200,+200)m/s
SEG/EAGE DISC 2003
3-47

FACTORIAL KRIGING VS MOVING AVERAGE FILTERING (COLEOU, 2001)

Moving average

- Smoothing increases when the size of the window increases
- Some criteria to choose smoothing window:
  - Maximising correlation with the well velocities
  - Visual QC of noise and smoothed grid

Factorial kriging

- Initial analysis allows identification of three features on the map
  - Noise (checkerboard pattern), through nugget effect
  - Acquisition/picking imprint. Stripes along inlines characterized by short-range spherical model
  - Low-frequency lateral variations with anisotropy N30E
- Interpolation only function of variogram model. Specific features can be filtered
  - Random noise
  - Organized noise (stripes)

FACTORIAL KRIGING EQUATIONS (SANDJIVY, 1987)

As usual, value at x_0 estimated by:

$$Z_{fk}(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i Z(\mathbf{x}_i)$$

Factorial kriging system for component C^p:

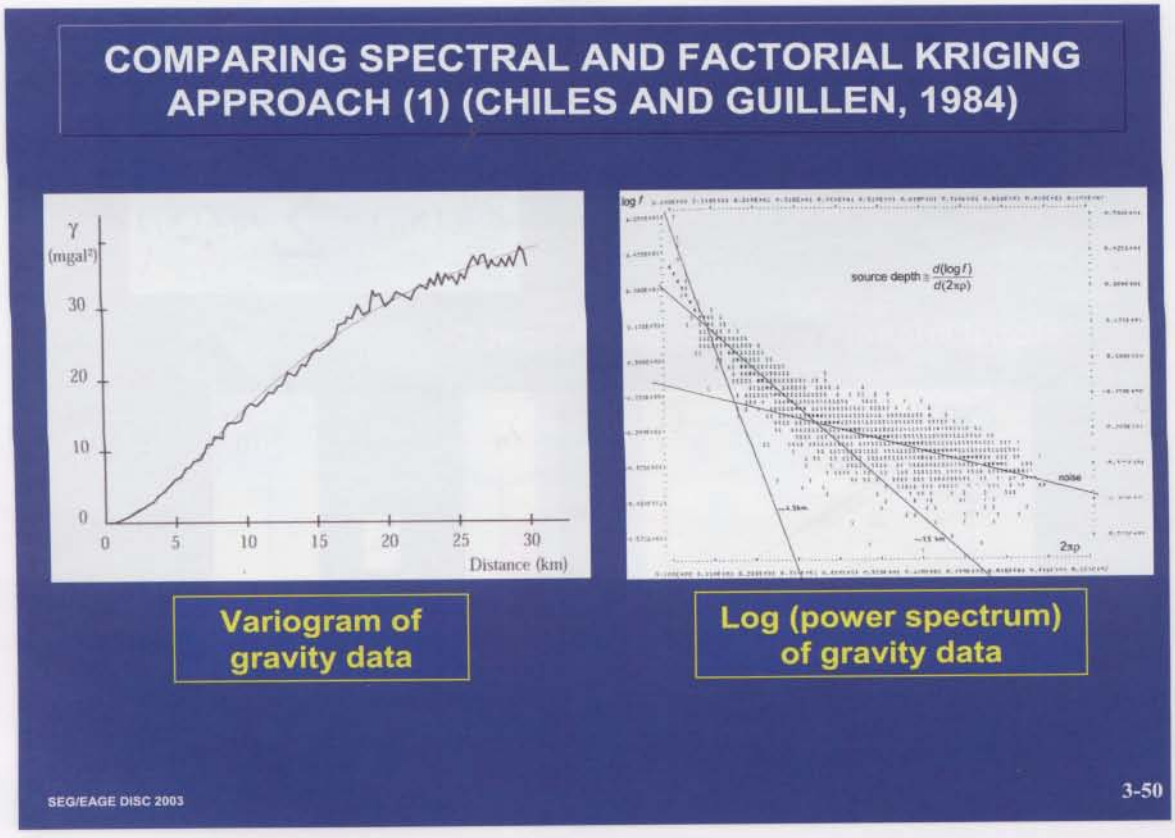
$$\begin{bmatrix} C(0) & C_{12} & \cdots & C_{1N} \\ C_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \\ C_{N1} & \cdots & & C(0) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} C_{10}^p \\ \vdots \\ C_{N0}^p \end{bmatrix}$$

Same matrix as kriging

Only component C^p

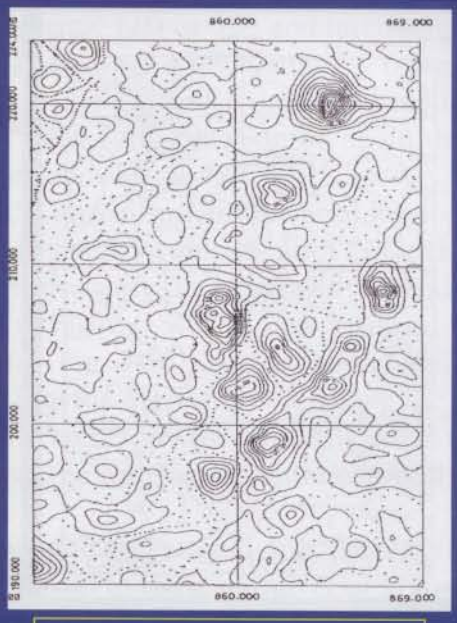
spectrum analysis lead to a split between two components (Fig. 3-50). The low- (respectively high-) frequency component is associated with the deep (respectively shallow) gravity field. Both analyses require significant interpretation that is based on a priori information about the geology. A significant difference is that variogram analysis is directly applied to the data, whereas spectral analysis is performed after a preliminary interpolation of these data on a rectangular grid. This interpolation usually introduces some smoothing and, hence, some low frequencies that were not initially in the data. In the example, the variogram that is fitted to the data is composed of the sum of two elementary Cauchy models (Fig. 2-25). The Cauchy model is closely associated with frequency spectra derived by Spector and Grant (1970) after making simplifying assumptions about the source locations. Fig. 3-51 show the results of kriging with this variogram model, and Figs. 3-52 and 3-53 compare the interpolation of the deep and shallow gravity fields obtained by each method. The two results are somewhat similar, however, the spectral method appears to have limitations that the factorial kriging approach does not have (Fig. 3-54).

As already mentioned with error cokriging, factorial kriging performs, in one single operation, the filtering and the smoothing, whereas spectral analysis must start from data already interpolated on a regular grid. As Coleou (2001) discusses, "Geostatistics, through variogram decomposition and factorial kriging, designs spatial filters very efficient for the removal of organized noise present in seismic velocities."



COMPARING SPECTRAL AND FACTORIAL KRIGING APPROACH (2) (CHILES AND GUILLEN, 1984)

Kriged gravity field

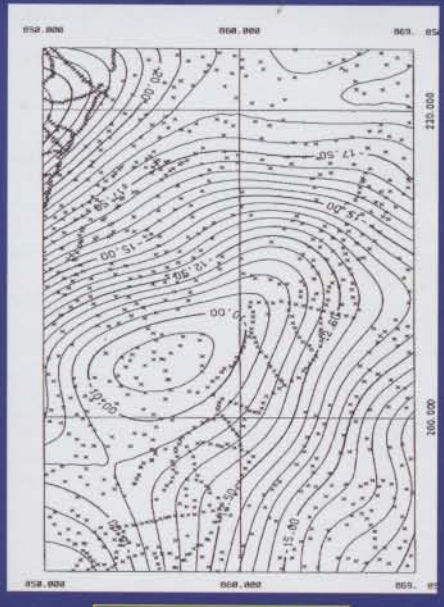


Kriging standard-deviation

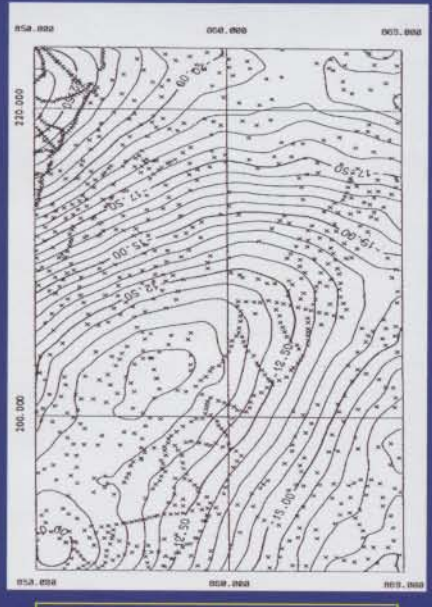
SEG/EAGE DISC 2003

3-51

COMPARING SPECTRAL AND FACTORIAL KRIGING APPROACH (3) (CHILES AND GUILLEN, 1984)



Spectral estimate



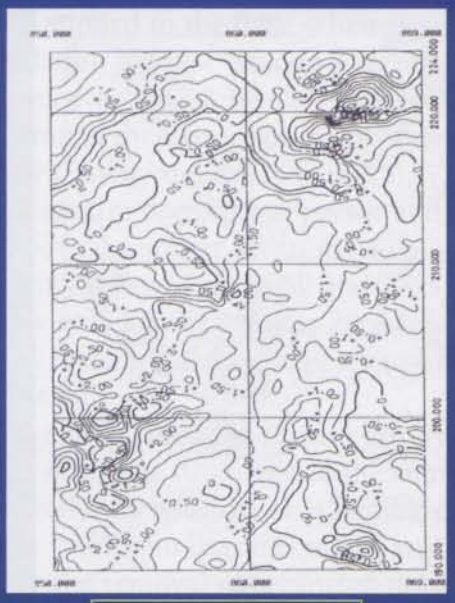
Factorial kriging estimate

Deep gravity field

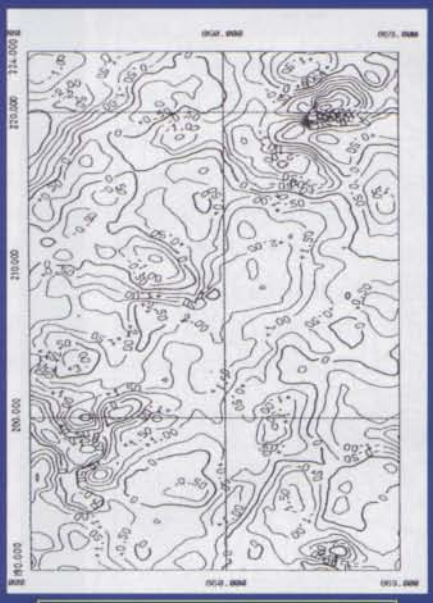
SEG/EAGE DISC 2003

3-52

COMPARING SPECTRAL AND FACTORIAL KRIGING APPROACH (4) (CHILES AND GUILLEN, 1984)



Spectral estimate



Factorial kriging estimate

Shallow gravity field

COMPARING SPECTRAL AND FACTORIAL KRIGING APPROACH (5) (CHILES AND GUILLEN, 1984)

Limitation of spectral method

- Complete rectangular grid requires interpolation + extrapolation
- Tapering or padding to make input grid periodic in both directions
- Stationarity and isotropy assumptions

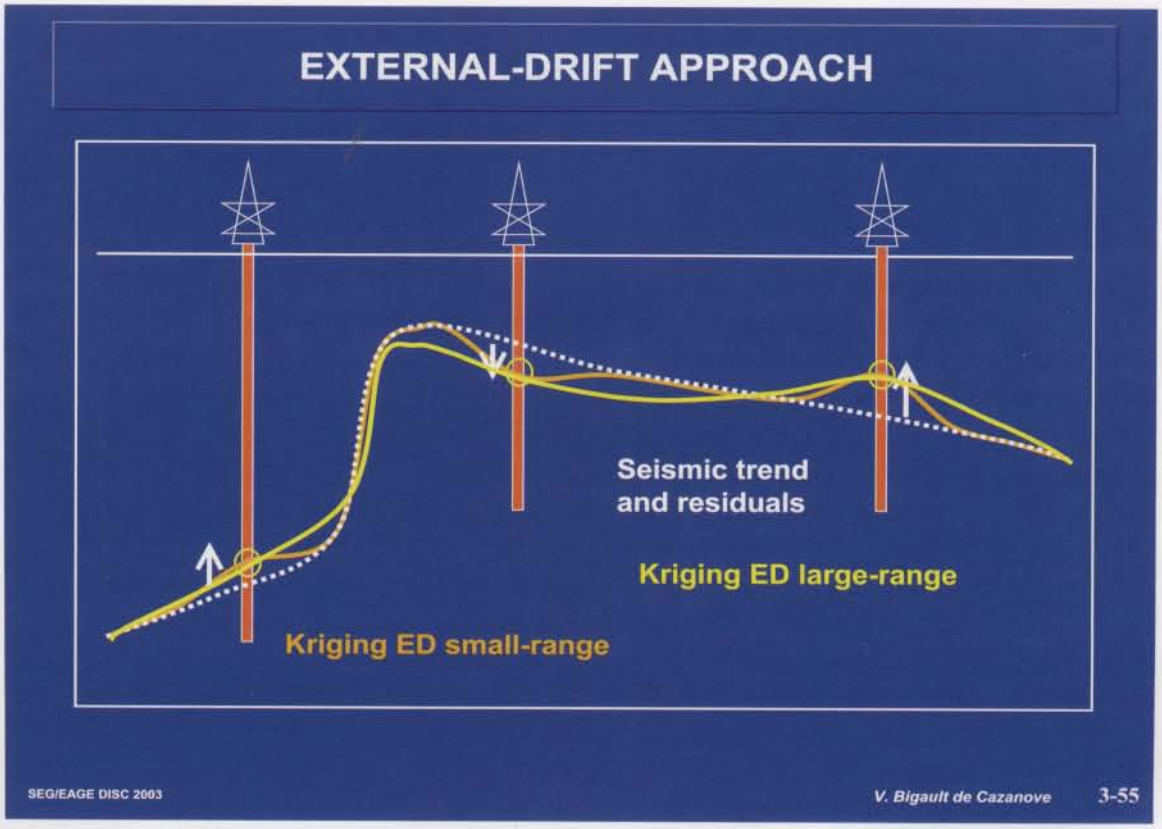
3.4 Kriging with an External Drift

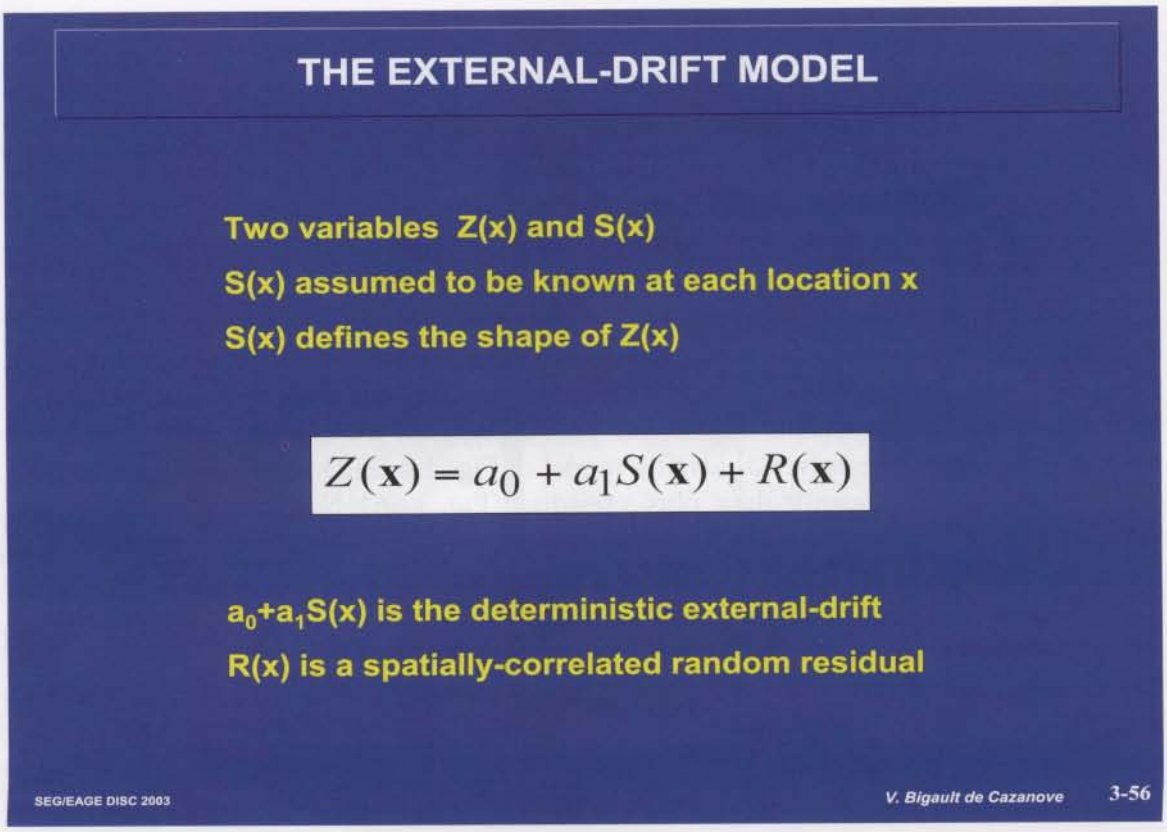
3.4.1 The external-drift model

The "external-drift" method has been specifically developed with seismic applications in mind. Take the example of a marker picked on seismic (2D or 3D) and intersected by a small number of wells. The seismic map can be considered a low-frequency representation of the actual seismic horizon. To construct this horizon, well data will need to be combined with information provided by seismic data, with some control about the range of influence of well data compared with seismic (Fig. 3-55).

How can this be modeled using geostatistical formalism? In universal kriging, the trend $m(\mathbf{x})$ can be regarded as precisely the low-frequency component of $Z(\mathbf{x})$. Then why not substitute the seismic data there? This is precisely the external drift model, easily derived from the universal kriging model of Fig. 2-6. $Z(\mathbf{x})$ is now assumed to be a linear function of the seismic data, plus a residual (Fig. 3-56). Naturally $R(\mathbf{x})$ will be close to zero away from the wells, but $R(\mathbf{x})$ will provide the "hump" allowing the model to go through the wells! Based on what we have seen before, we can easily guess that the residual's variogram range will help us control the width of the hump around the well data.

There is no need to write down the "kriging with an external drift" (KED) equations. In 2D (mapping) applications, they are identical to those of universal kriging with a linear trend, except that the two trend components x and y are replaced by the single component $S(x,y)$, the value of the seismic variable at location (x,y). An interest-





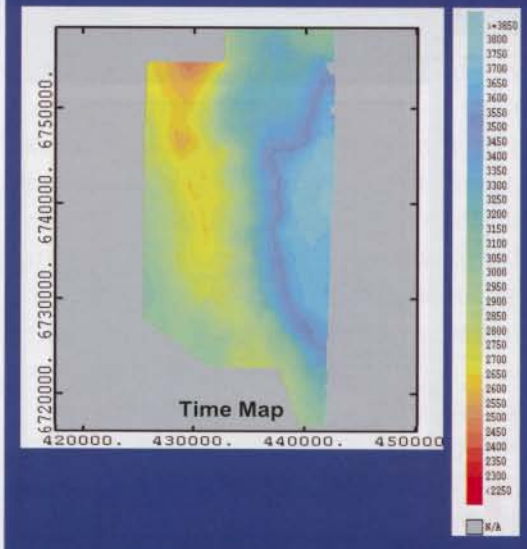
ing KED property is that, if $S(x,y)$ is multiplied by a constant, the result is left unchanged. KED is used mostly for time-to-depth conversion.

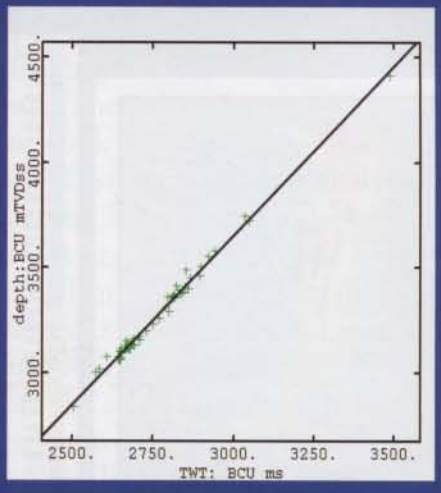
3.4.2 Examples of time-to-depth conversion using the external-drift approach

Fig. 3-57 shows a time horizon (the base Cretaceous unconformity, or BCU) from the Alwyn North example already discussed. Sixty-two wells are available, but we use the depth data from only 59 of them and save the three others to check the results. Fig. 3-57 shows that there is an excellent linear correlation between time and depth. In spite of this excellent correlation, the residuals may take values as high as 50 m (Fig. 3-58), but there is no correlation between residual and time to the BCU. The residuals' experimental variogram has a range of around 4 km, which confirms that depth can be interpreted as a sum of a linear transform of time plus a stationary residual. Note that this is different from the trend-surface-analysis model (Fig. 1-39), which assumes that the residual from the trend is spatially uncorrelated.

In this example, the inputs to KED are a time map and well depths, and the output is a depth map (Fig. 3-59). This means that KED performs, in one single step, an operation that is often performed in two steps. A common practice consists of first transforming the seismic data using a linear transform, then correcting the misfits at the wells by constructing a map of the residuals between the well data and the seismic-derived map. Here, as shown in Fig. 3-60, the KED procedure makes sure that the linear transformation and the mapping of the residual take place in one step. It is easy to

NORTH ALWYN EXAMPLE
BCU TIME-TO-DEPTH CONVERSION



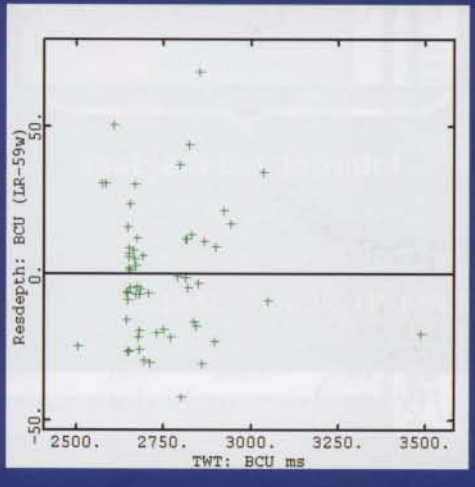


Depth predicted by a linear regression of time:

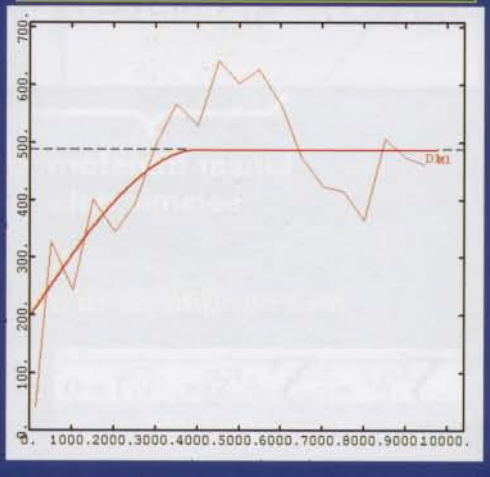
- *59 wells (on 62 available, 3 wells used for validation)*
- *TVD = 1.6*TWT-1146.4*
- $\rho = 0.996$

NORTH ALWYN EXAMPLE
BCU TIME-TO-DEPTH CONVERSION

Statistics on depth residuals

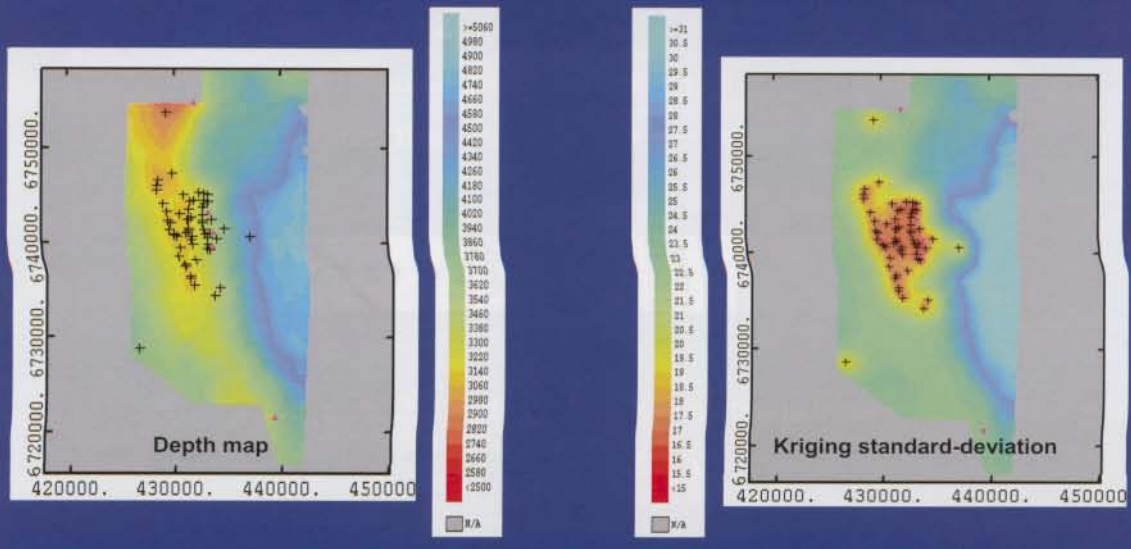


Variogram calculated on depth residuals



Variogram model:
$\gamma(h) = 200 + 288$ spherical (4000m)

NORTH ALWYN EXAMPLE: BCU TIME-TO-DEPTH CONVERSION USING TIME AS AN EXTERNAL DRIFT



SEG/EAGE DISC 2003

P. Delfiner/X. Freulon 3-59

KRIGING WITH AN EXTERNAL DRIFT SEEN AS AN INTERPOLATING FUNCTION (IN 2D)

$$Z_{ked}(x,y) = a_0 + a_1 S(x,y) + \sum_{i=1,N} b^i \gamma \left[\sqrt{(x-x_i)^2 + (y-y_i)^2} \right]$$

Linear transform of seismic data

Interpolated residual

N+2 conditions to determine the N+2 coefficients

$$\sum_{i=1,N} b^i = \sum_{i=1,N} b^i S(x_i, y_i) = 0 \quad \& \quad Z_{ked}(x_i, y_i) = z_i \text{ at N data points}$$

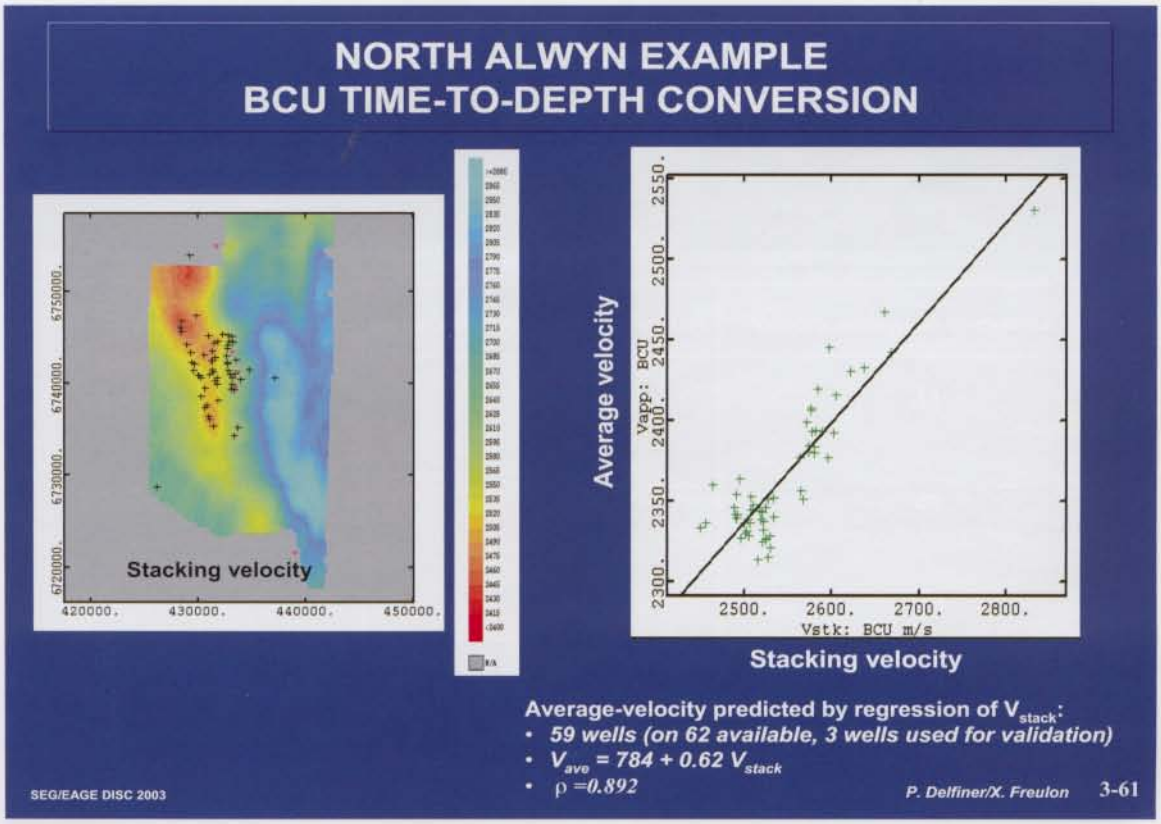
SEG/EAGE DISC 2003

3-60

check that the interpolated residual is zero, at a distance from the well data greater than the variogram range.

In our example, because the variogram has been modeled using a nugget effect, the depth values at the wells are filtered. The depth map obtained by KED is equal to the trend, that is, a linear function of time, as soon as the distance from all data points reaches 4 km. Close to the wells, however, the depth values at the wells control the interpolation. The kriging standard deviation is meaningful here, because it uses the variogram model actually fitted to the residuals' experimental variogram. Although the external-drift approach has been used a lot for time-to-depth conversion, one of its drawbacks is that it maps depth from time, without providing a detailed analysis of velocities. Another issue is that the velocity information is provided by the wells only, through their depth values. This introduces a bias if wells do not penetrate zones that are representative of the actual velocity variations. Coleou (2001) stresses that stacking velocities provide the regional information that addresses this bias issue. Let us see how such information could have been used in the North Alwyn example.

Fig. 3-61 shows a map of stacking velocities to the same horizon as before. The crossplot between stacking velocity and the 59 average well velocities confirms that stacking velocities tend to be faster than well velocities, because of nonvertical travel paths and heterogeneity due to layering. This time, the KED model will use stacking velocities as an external drift in order to interpolate seismic velocities away from the wells. Once again, velocity residuals are not correlated with time to the BCU (Fig. 3-62). An exponential variogram model is used, with a zero nugget effect. KED properly



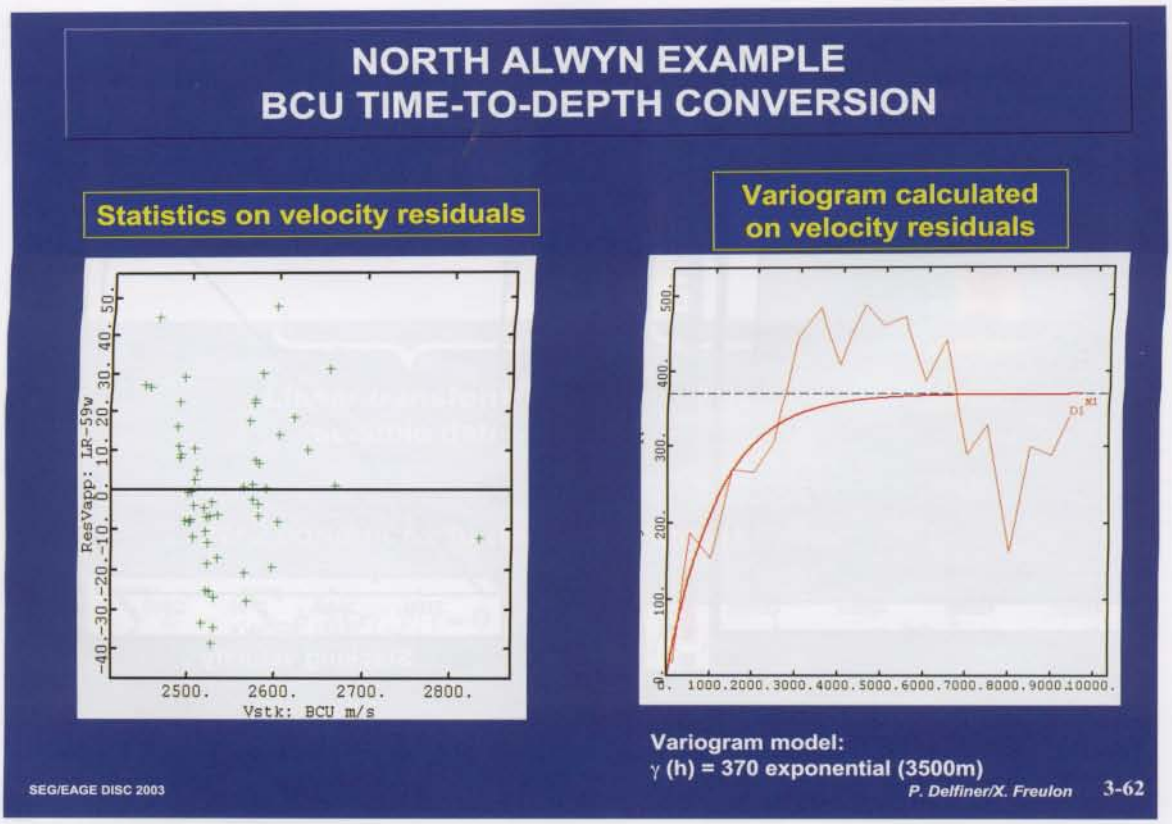
maps the large blue zone on the right, thanks to regional information provided by stacking velocities (Fig. 3-63).

Comparison with data from three wells (Fig. 3-64) shows that the method using stacking velocities seems to perform slightly better for predicting depth than does the one based only on well depths and the time map. Beyond this difference, which may be of a purely statistical nature with only three wells, geophysicists prefer working with stacking velocities in order to control the quality of the velocity map before combining it with time (Coleou, 2001).

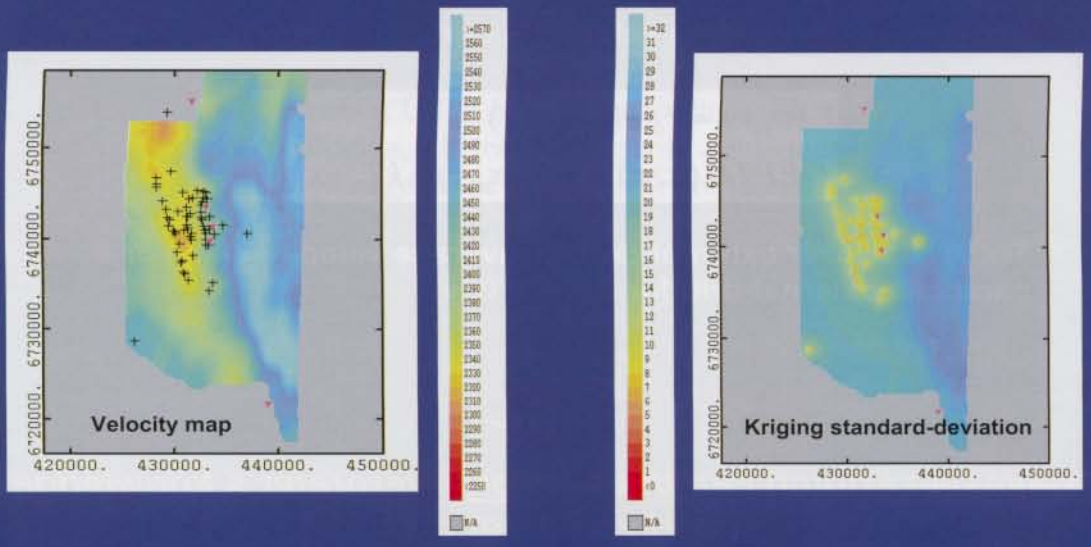
3.5 Bayesian Kriging, a Generalization of Kriging with an External Drift

Bayesian kriging (Omre, 1987) is a generalization of KED that is useful for time-to-depth conversion, when the coefficients of the external drift have a physical interpretation (Fig. 3-65). In such situations, and when the number of wells is not enough to provide a good statistical calibration, Bayesian kriging (BK) offers the possibility to specify a priori statistical constraints on these parameters. These constraints can be derived from surrounding reservoirs of the same formation that happen to be at a more mature developmental stage (Abrahamsen et al., 1991, 2000). Fig. 3-66 is a BK application in a North Sea example, using the model of Fig. 3-65. Only two wells are available, but, thanks to regional information, a priori statistical information is injected to constrain the estimation of k and V_0. This example will be used again when discussing uncertainty quantification.

Bayesian kriging can also be used in situations in which a time-to-depth conver-



NORTH ALWYN EXAMPLE: VELOCITY MAP USING V_{STACK} AS AN EXTERNAL DRIFT



NORTH ALWYN EXAMPLE COMPARISON OF ERRORS ON 3 BLIND WELLS

Well	Actual depth	Depth estimation by external drift	Velocity estimation by external drift
3/9a-N39	3524.00	16	11
3/9a-N40	3583.00	31	26
3/9a-N41	3460.00	-26	-24

BAYESIAN KRIGING, A GENERALIZATION OF THE EXTERNAL-DRIFT MODEL (ABRAHAMSEN ET AL, 2000)

Because of compaction velocity is correlated with time:

$$V(x,y) = V_0 + kT(x,y)$$
$$DEPTH(x,y) = V_0 T(x,y) + kT^2(x,y)$$

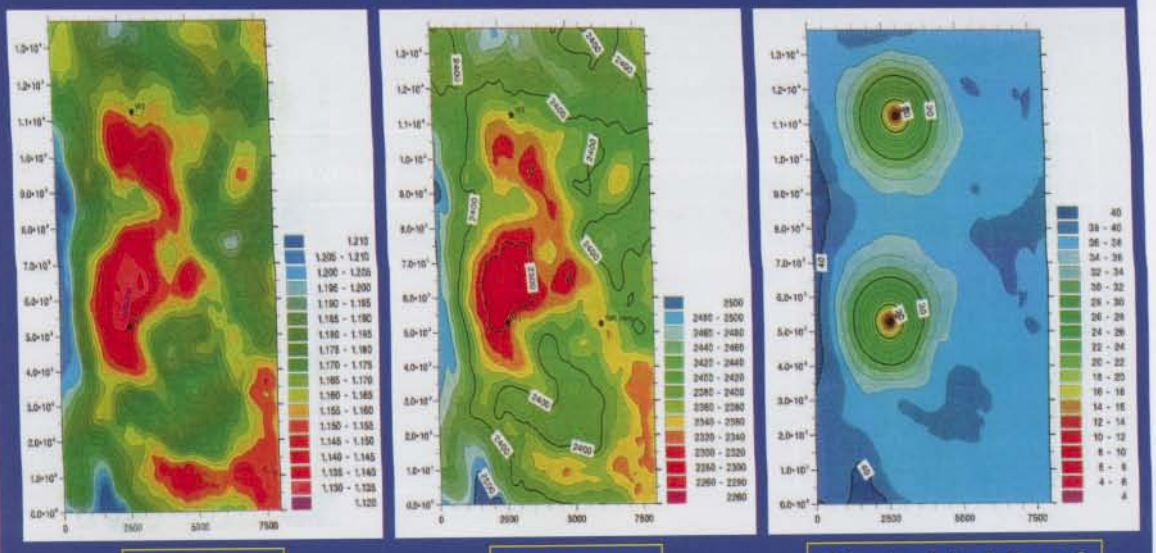
T(x,y) is a linear external-drift for average velocity V(x,y) or a parabolic external-drift for DEPTH (x,y).

Bayesian kriging can use a priori interpreter's knowledge (mean and variance) on V_0 and k

SEG/EAGE DISC 2003

3-65

BAYESIAN KRIGING EXAMPLE: NORTH SEA CASE STUDY (ABRAHAMSEN ET AL., 2000)



Time map | **Depth map** | **Standard deviation of depth map error**

SEG/EAGE DISC 2003

3-66

sion model has been constructed using a significant number of wells, say 10, and when the interpreter wants to avoid, while the model is updated by new well data, a complete change in the external drift used. Thanks to BK, the user can specify a rather small variance — hence a small uncertainty — on the coefficients of the trend, thus controlling the impact of new wells (Th. Coleou, personal communication, 2002).

The concept of BK can also be applied to universal kriging, where the trend is constant, linear, or parabolic, but is not derived from seismic as it is with KED (Omre and Halvorsen, 1989). For instance, UK or KED is equivalent to BK with no a priori constraints on the trend coefficients, whereas kriging with a zero mean is equivalent to BK with a perfect a priori knowledge of the trend coefficient (mean and standard deviation equal to zero).

3.6 Cokriging and Collocated Cokriging

3.6.1 Introduction

The KED method was a transition toward multivariate estimation techniques, because it consisted of interpolating a parameter measured at wells using information from a parameter measured on seismic, assuming that the latter could be interpreted as a trend of the former. With cokriging, we are now going to discuss the most general kriging-based approach for combining several sources of information. We will limit ourselves to bivariate cokriging, that is, the interpolation of one parameter by use of a weighted average of values of this parameter (the primary variable) at a number of locations and values of another parameter (the secondary variable) at other locations (Fig. 3-67). There is no need to mathematically develop the cokriging equations, which make use not only of the variogram of each variable but also of cross-variogram (or cross-covariance) functions already discussed in Figs. 2-33 and 2-34.

3.6.2 A cokriging example

Cokriging has been successfully used in mining applications since the 1960s. Doyen's (1988) article in *GEOPHYSICS* had a significant impact in the industry, because it showed how the technique could be used to map porosity over an Alberta oil-bearing reservoir using porosity data at the wells and acoustic impedance resulting from a seismic inversion exercise. The cross-covariance between porosity and the inverse of impedance, and the inverse impedance covariance model used by Doyen, were those of Fig. 2-34 (the assumption was made that impedance data were exact and not affected by any uncertainty, which is a simplification because inversion is known to be a nonunique process). Both models are proportional to a Gaussian covariance of range about 1 km. The porosity variogram could not be obtained from the well data (only eight wells). To solve this problem, Doyen assumed that the porosity variogram was also proportional to the Gaussian model.

Comparison of the three maps (Fig. 3-68) shows that, in areas where there is little well control (southeast part of the map), the cokriged map is controlled by the linear regression of porosity against the inverse of impedance. The map derived from linear regression does not honor the well data, but cokriging makes sure that near the wells, the map is strongly controlled by the well data. In spite of the successful application by

COKRIGING

Two variables $Z_1(x)$ and $Z_2(x)$ (such as porosity & acoustic impedance)
Use of Z_1 and Z_2 data to get a better interpolation of Z_1

$$Z_{cok}(\mathbf{x}_0) = \sum_{1 \le i \le N_{Z_1}} \lambda_i Z_1(\mathbf{x}_i) + \sum_{1 \le j \le N_{Z_2}} \mu_j Z_2(\mathbf{x}_j)$$

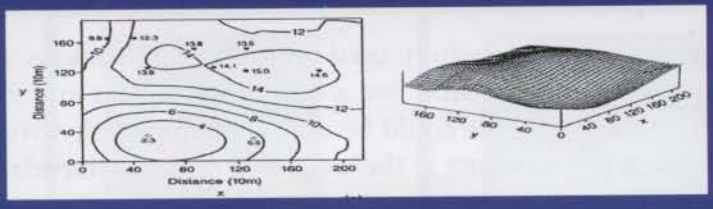
$$\sigma^2_{cok}(\mathbf{x}_0) = Var\left[Z(\mathbf{x}_0) - Z_{cok}(\mathbf{x}_0)\right]$$

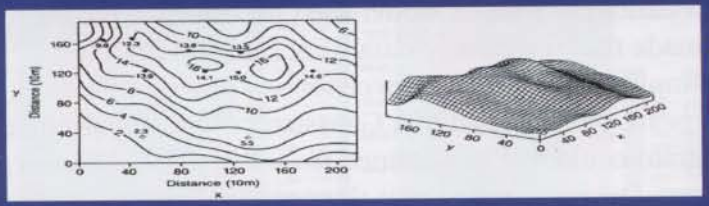
Thanks to the extra information coming from Z_2 (x):

$$0 \le \sigma^2_{cok}(\mathbf{x}_0) \le \sigma^2_k(\mathbf{x}_0)$$

COKRIGING (DOYEN, 1988)

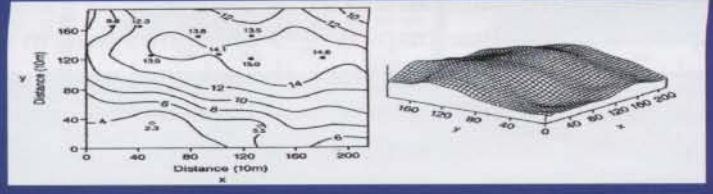
Kriged porosity





Porosity by regression from 1/(acoustic impedance)

Porosity by cokriging



Doyen, the cokriging system and the input needed to be simplified. Xu et al. (1992) changed this situation by proposing a clever simplification of the cokriging system.

3.6.3 Collocated cokriging

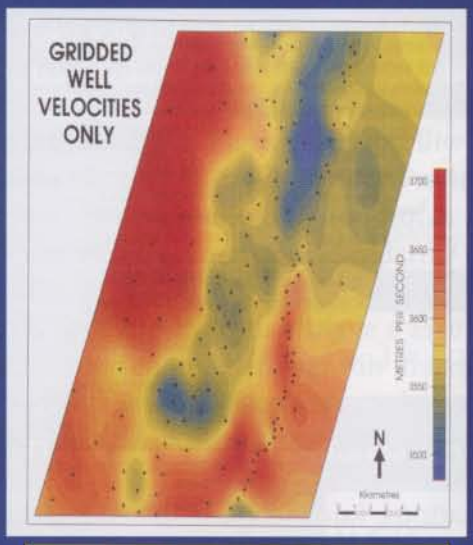
In petroleum applications, the primary variable is usually known at wells, whereas the secondary variable is derived from seismic data. When we are predicting a value away from the wells, the knowledge of the secondary variable at that location will be crucial. However, because of the smoothness of seismic data, the knowledge of seismic values at other locations close to the interpolated point will not bring much extra information, once the seismic value at the interpolated location has been used. This is the philosophy behind collocated cokriging (CCK). Xu et al. proposed to retain, at each interpolated location $\mathbf{x}_0$, only the value of the secondary variable Z_2 at $\mathbf{x}_0$ itself. This assumption also had the advantage of greatly simplifying the resolution of the system by making the matrix inversion more stable. Coleou (2002a) uses a slightly different definition of CCK, because he incorporates in the kriging system not only the value of the secondary variable at the estimated location $\mathbf{x}_0$, but also its values at the data points $\mathbf{x}_i$. Chilès and Delfiner (1999) call this other approach "multi-collocated cokriging" (MCCK), and compare the relative merits of CCK and MCCK.

It is easy to confirm that, in the CCK equations, knowledge of the variogram of the secondary variable is not necessary, except for the variance itself. Because only the secondary variable value at $\mathbf{x}_0$ is used, this value needs only to be cross-correlated with itself or with the values of the primary variable. In a way, this is a pity, because the variogram of the secondary parameter is usually the one that is best known because of the large number of seismic data available! Nevertheless, in practice few wells are available, which means that, as explained in Fig. 2-39, inferences about the primary-variable variogram tend to be made using information derived from the seismic-data variogram. Realizing that variogram and cross-variogram models were usually oversimplified because of well data scarcity, Xu et al. (1992) proposed a further approximation to CCK, related this time to the choice of the cross-covariance model. They called this the Markov model, which assumes that the cross-covariance function is proportional to the covariance of the primary variable. Thus, if the Markov model is used, all that is required as input is the knowledge of (1) the correlation coefficient between the primary and secondary variable, (2) one single model assumed to be representative of both the primary variable covariance and the cross-covariance between the primary and secondary variable, and (3) the variances of the two variables.

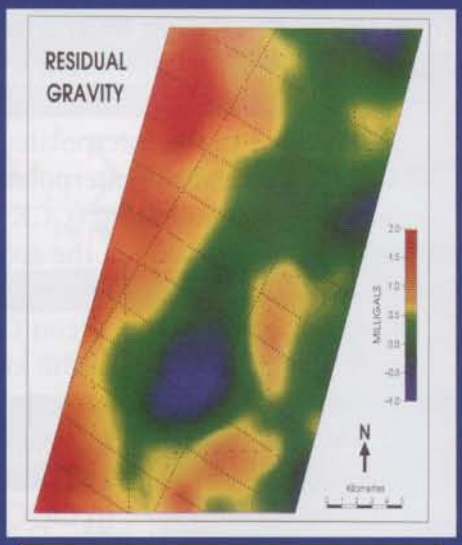
A good example from the giant Ghawar field (Saudi Arabia) is presented by Jeffery et al. (1996). Although 150 wells were available for predicting average velocity, this proved insufficient in terms of lateral coverage. Jeffery et al. used residual gravity as a secondary variable to provide velocity information between wells. Fig. 3-69 shows the maps of each variable. Well velocities are simply equal to vertical depth of the reservoir divided by one-way seismic traveltime. Residual gravity data were interpolated over a regular 500-m grid, after a number of filtering steps. Using the variogram of residual gravity data, Jeffery et al. applied ordinary kriging to estimate their values at the well location. The crossplot between these kriged values and well velocities showed that the correlation was excellent (Fig. 3-70). Predicting velocity away from the wells by using

COLLOCATED COKRIGING (JEFFERY ET AL., 1996)

Problem: predict average velocity from wells and residual gravity data

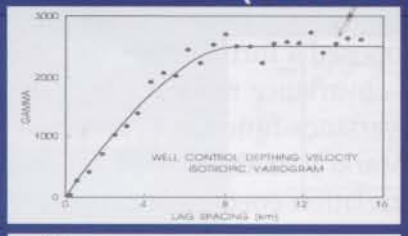


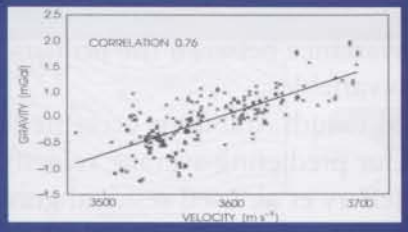
Kriged well velocities

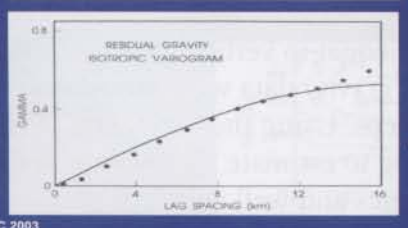


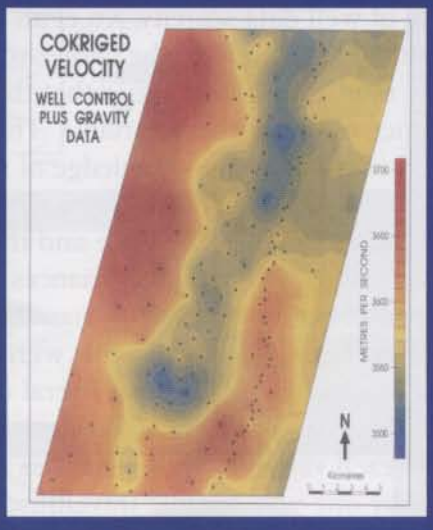
Kriged residual gravity

COLLOCATED COKRIGING (JEFFERY ET AL., 1996)







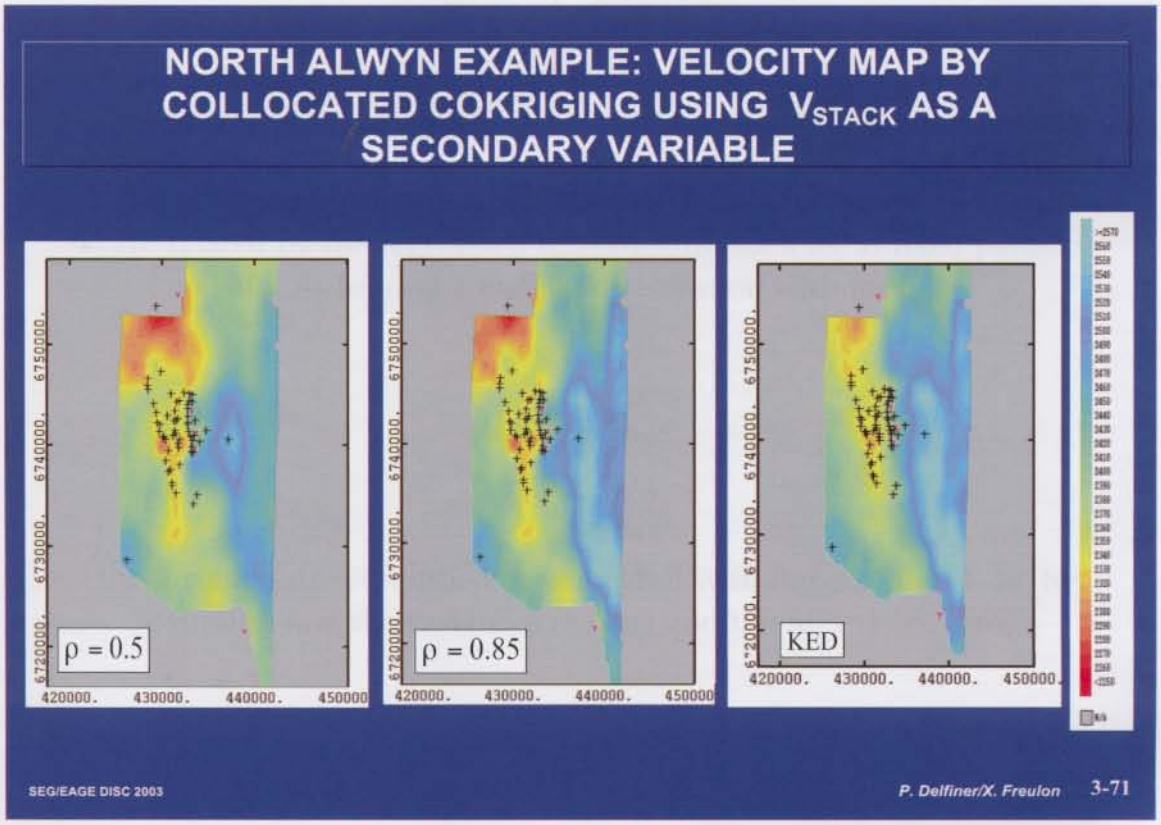


Just the variance of residual gravity is used, not the whole variogram!

just a linear transform of kriged residual gravity would already provide good results, but the well velocity data would not be honored. Thus, Jeffery et al. applied CCK. They made the Markov assumption, meaning that the input to the exercise was limited to the variogram of velocity data, the correlation coefficient, and the variance of residual-gravity data. This led to the velocity map of Fig. 3-70. Cross-validation showed that, thanks to CCK, an improvement of more than 25% was obtained in the mean absolute error of velocity estimates, which decreased from 22 m/s with standard mapping techniques to 15.5 m/s with CCK.

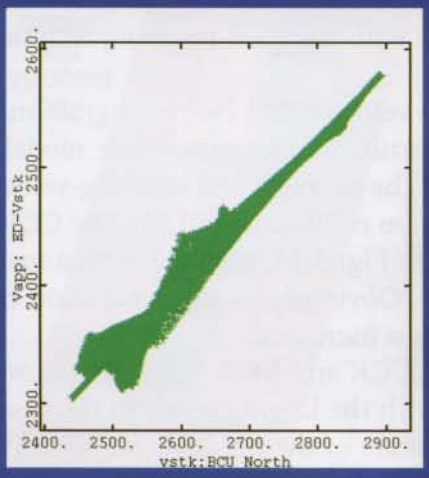
Let us now come back to the North Alwyn velocity data (see variogram models in Fig. 2-33). All three variogram models are proportional to an anisotropic model. Together with this theoretical model, the sills of the average- and stacking-velocity variograms determine the variance, and the correlation coefficient is 0.85. The CCK map obtained with these input parameters is shown in Fig. 3-71, together with a map obtained using a different correlation coefficient. Obviously, as the correlation coefficient increases, the weight of the V_{stack} information increases.

Fig. 3-72 illustrates the difference between CCK and KED. V_{stack} (the secondary variable) tends to have a stronger relationship with the kriging result in the case of KED. This may be due to the fact that the variogram of the residuals for KED (Fig. 3-62) has a smaller range than does the variogram of the primary variable for CCK (Fig. 2-33). As a result, with CCK, the final interpolation remains under the control of the well data for a longer distance.

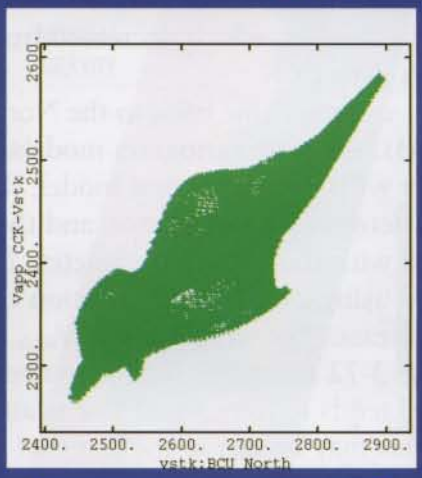


NORTH ALWYN EXAMPLE
TIME-TO-DEPTH CONVERSION

Relationship between secondary variable and kriging result



External drift



Collocated cokriging

SEG/EAGE DISC 2003

P. Delfiner/X. Freulon 3-72

A Few Words on Bayesian Statistics

Let us discuss Bayesian statistics using a simple example, inspired from Doyen et al. (1996). This is the situation where, at one location $\mathbf{x}_0$ in 2D or in 3D, we want to predict, say, porosity, using two kinds of information: that coming from the other wells and that coming from the value of a seismic-derived parameter — or attribute — at the same location $\mathbf{x}_0$. Two kinds of information will help us achieve this: the knowledge of the seismic A_0 attribute at location $\mathbf{x}_0$ and the knowledge of porosity values at the other wells.

We also assume that porosity is normally distributed. Then, the best estimate of Φ_0 from the wells only will be kriging Φ_k, with a kriging standard deviation equal to σ_k. Under the Gaussian assumption, we can even say that Φ_0 follows a normal distribution $p(\Phi_0/\Phi_k)$ with mean Φ_k and standard deviation σ_k. This is the *a priori* information we have on the unknown porosity, on the basis of the other wells and the geological knowledge quantified by the variogram.

Now, we also have a measure of the seismic attribute A_0 at location $\mathbf{x}_0$. We assume that there is a linear correlation between porosity and seismic attribute, quantified by the correlation coefficient ρ. A_0 and Φ_0 are related by the regression relationship, such as, at fixed porosity Φ_0, A_0 follows a normal distribution with mean

$$A_0 = \rho\Phi_0$$

and variance:

$$(1 - \rho^2)$$

(the formulas are simplified as compared with those of Fig. 1-35, by assuming that the variances are one and the means are zero).

Bayes's relationship provides a way to combine these two types of information, that given by the wells, and that given by the seismic attribute. We use the following formula giving the probability of finding Φ_0 given both the porosity values at the wells and the seismic attribute A_0 at location $\mathbf{x}_0$:

$$p\,(\Phi_0 / A_0, \Phi_1,\ldots, \Phi_N) \propto f\,(A_0 / \Phi_0)\; p\,(\Phi_0 / \Phi_1,\ldots, \Phi_N)$$

The second term of the right-hand side is the *a priori* pdf of Φ_0, given the values of porosity at the other wells. We saw that this was a Gaussian pdf of mean equal to the kriging estimate Φ_k and of standard deviation equal to the kriging standard deviation σ_k. This pdf represents the knowledge we have of porosity Φ_0 before using the seismic attribute, just on the basis of geological knowledge (quantified by the variogram) and wells.

The first term of the right-hand side is the likelihood function for Φ_0. In a way, for each possible value of Φ_0 at $\mathbf{x}_0$, it measures how compatible the value of the measured seismic attribute A_0 is with this value Φ_0.

Under the Gaussian assumption, the right-hand side is the product of two exponential functions,

$$\exp\left[-\frac{(A_0 - \rho\Phi_0)^2}{2(1-\rho^2)}\right] \bullet \exp\left[\frac{(\Phi_0 - \Phi_k)^2}{2\sigma_k^2}\right].$$

This is the posterior pdf of Φ_0, combining the porosity information from the wells with the information derived from the amplitude A_0 about porosity Φ_0. Simple calculations show that, as a result

of the Bayes relationship, this is itself a normal distribution, with a mean equal to:

$$\frac{\rho\sigma_k^2 A_0 + (1-\rho^2)\Phi_k}{\rho^2(\sigma_k^2 - 1) + 1}$$

This mean is simply a weighted average of the two estimates of Φ_0 at $\mathbf{x}_0$: the estimate based on regression against the seismic attribute and the estimate based on kriging. Each estimate is weighted by the inverse of its estimation variance. The variance of this pdf is equal to the harmonic average of the variances of each estimate, and it can be written:

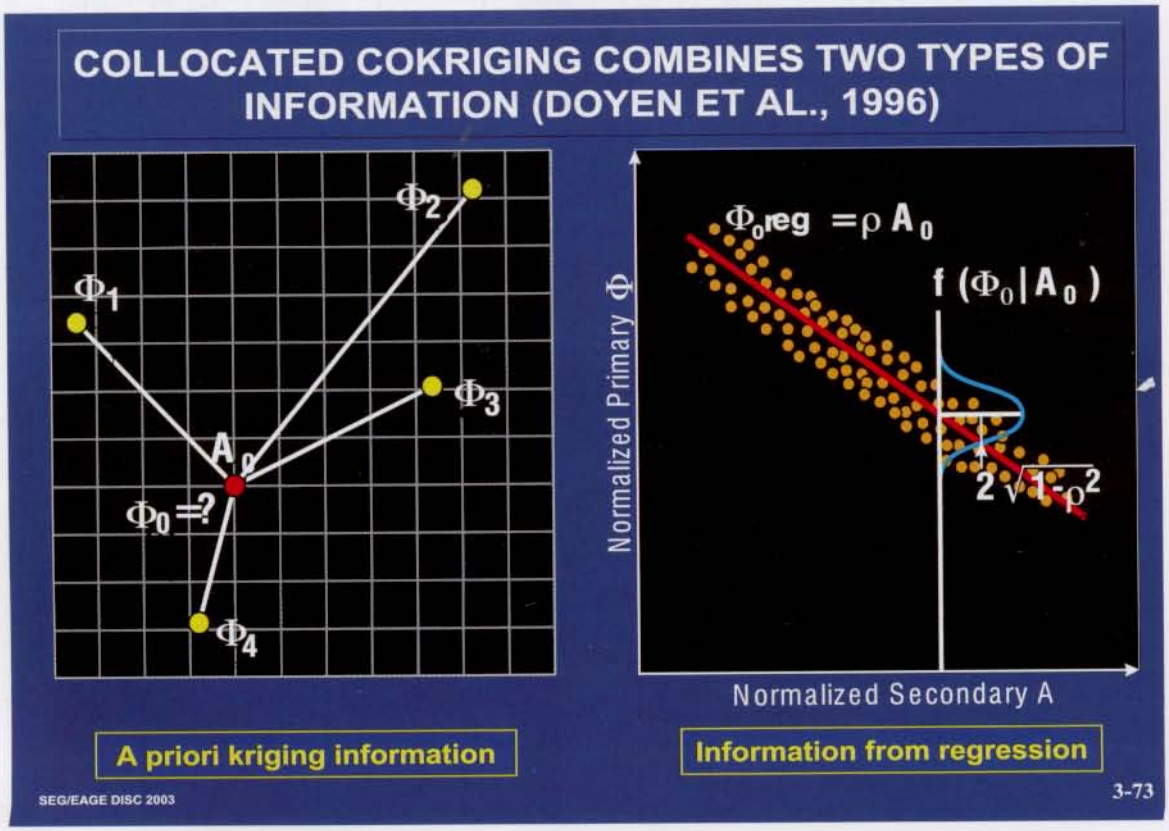
$$\frac{\sigma_k^2 (1-\rho^2)}{\rho^2(\sigma_k^2 - 1) + 1}$$

The beauty of this result is that it is exactly equal to the result of CCK! Thus, this result not only relates the world of kriging to that of Bayesian analysis, it also shows that the CCK estimate can be simply calculated as a weighted average of two estimates. The weight of each estimate is a function of the relative magnitude of the correlation coefficient versus the kriging standard deviation. This is discussed more graphically in the following section.

3.6.4 *Revisiting collocated cokriging*

Doyen et al. (1996) provided new insight on CCK by reformulating it in the context of Bayesian analysis, as mathematically developed in the previous section. They used the example where the primary and secondary variables are porosity at wells and acoustic impedance from seismic. They considered that kriging at one location, x_0, could be interpreted as the result of using the *a priori* geostatistical model. Based on the well data only and on the variogram of the primary variable, an a priori kriging estimate and its associated variance are calculated. Then this estimate is updated thanks to new information derived from a second source, which is acoustic impedance (Fig. 3-73). The a priori model is combined with the likelihood function to provide the posterior distribution (Fig. 3-74). This leads to a beautifully simple expression, showing that collocated cokriging is a weighted average of kriging and of the estimate derived from regression from seismic (Fig. 3-75). The weighting factor is simply related to the correlation coefficient between porosity and acoustic impedance. If the coefficient is zero, only kriging remains, whereas if it is one, only the regression estimate is used. Also, the higher the kriging variance, the higher is the weight of the regression estimate. This decoupling of the two sources of information also helps solve the collocated cokriging system very easily.

Doyen et al. (1996) provided an application from the Ekofisk field in Norway. Porosity is predicted from an inversion-derived impedance map (Fig. 3-76), which is assumed to be exact information, because no uncertainty is associated with this map. The correlation coefficient between acoustic impedance and porosity is equal to –0.77 (Fig. 3-77), and the result of CCK is the map of Figs. 3-78 and 3-79. In addition to this,



COLLOCATED COKRIGING BY BAYESIAN UPDATING OF KRIGING (DOYEN ET AL., 1996)

Cokriging Gaussian Distribution

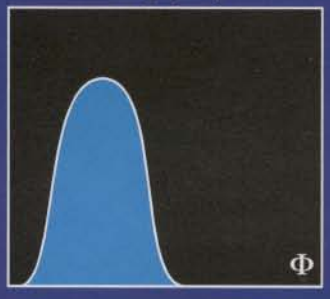
$p(\Phi_0 | A_0, \Phi_1, \dots, \Phi_4)$

Posterior

=

Secondary Gaussian Likelihood

$f(A_0 | \Phi_0)$

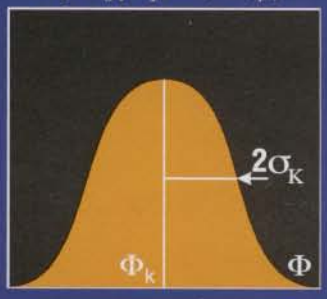


Likelihood

X

Kriging Gaussian Distribution

$P(\Phi_0 | \Phi_1, \dots, \Phi_4)$

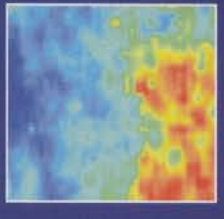


A priori

COLLOCATED COKRIGING BY BAYESIAN UPDATING OF KRIGING (DOYEN ET AL., 1996)

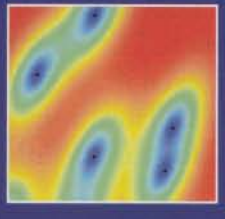
ρ

Attribute A_0



X

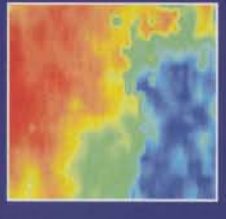
Kriging variance σ_K^2



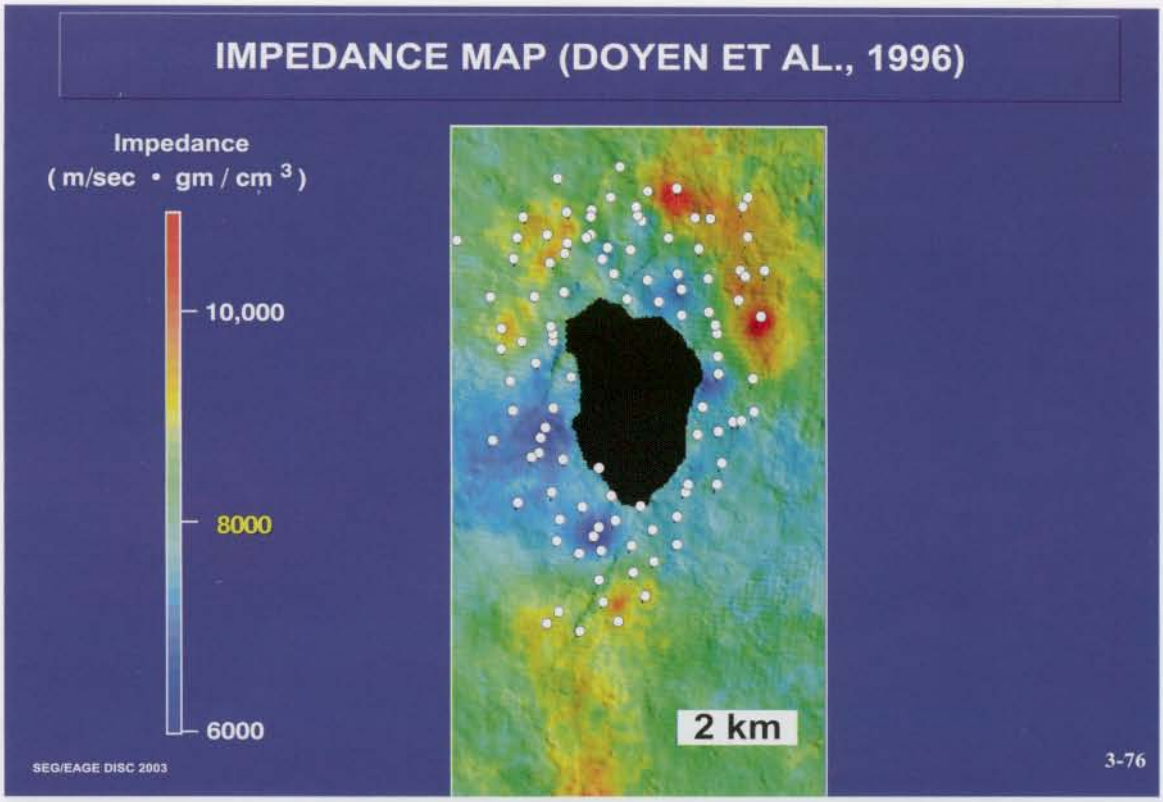
$+(1-\rho^2)$

Kriging Φ_K

$\propto$



Collocated Cokriging

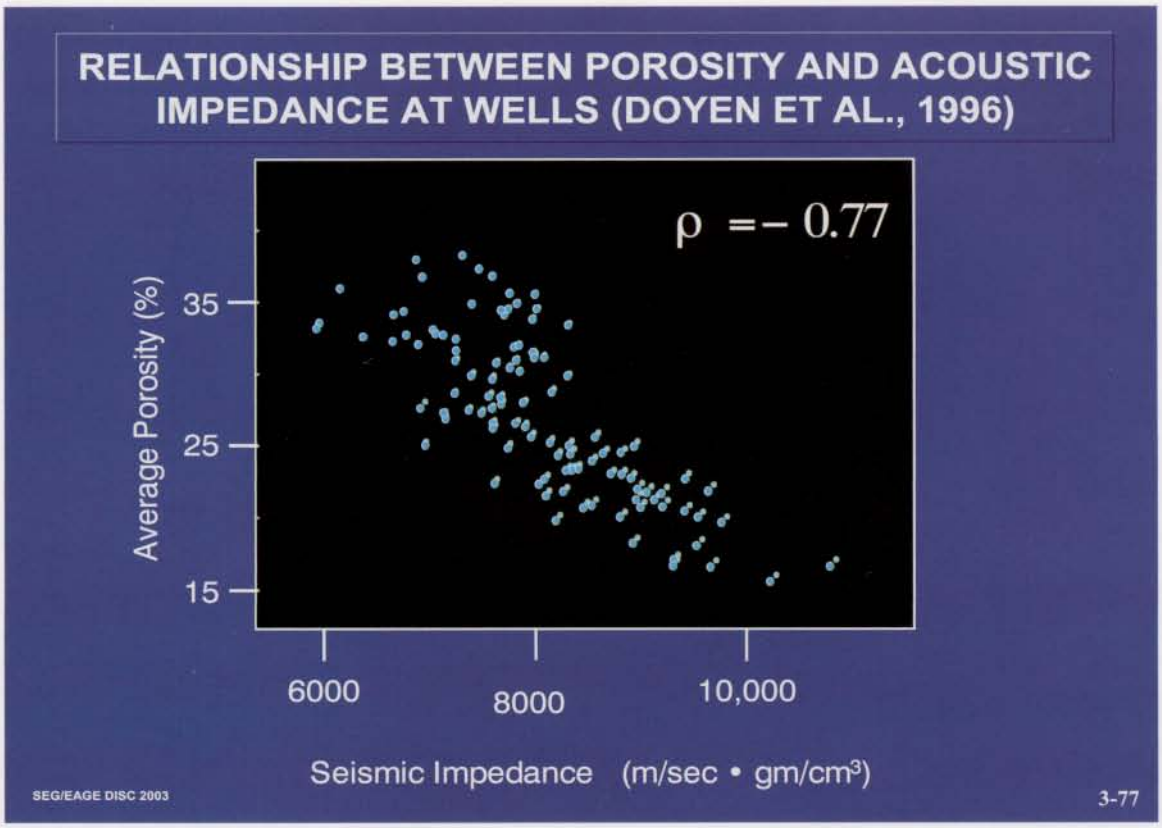


Doyen et al. ran a blind-wells exercise. The well data set was divided in two groups: wells drilled before and after 1989 (Fig. 3-80). Comparison of actual versus predicted values at the wells drilled after 1989 showed that there is a significant improvement resulting from the incorporation of seismic data in the estimation exercise (Fig. 3-81).

3.6.5 Collocated cokriging versus external drift

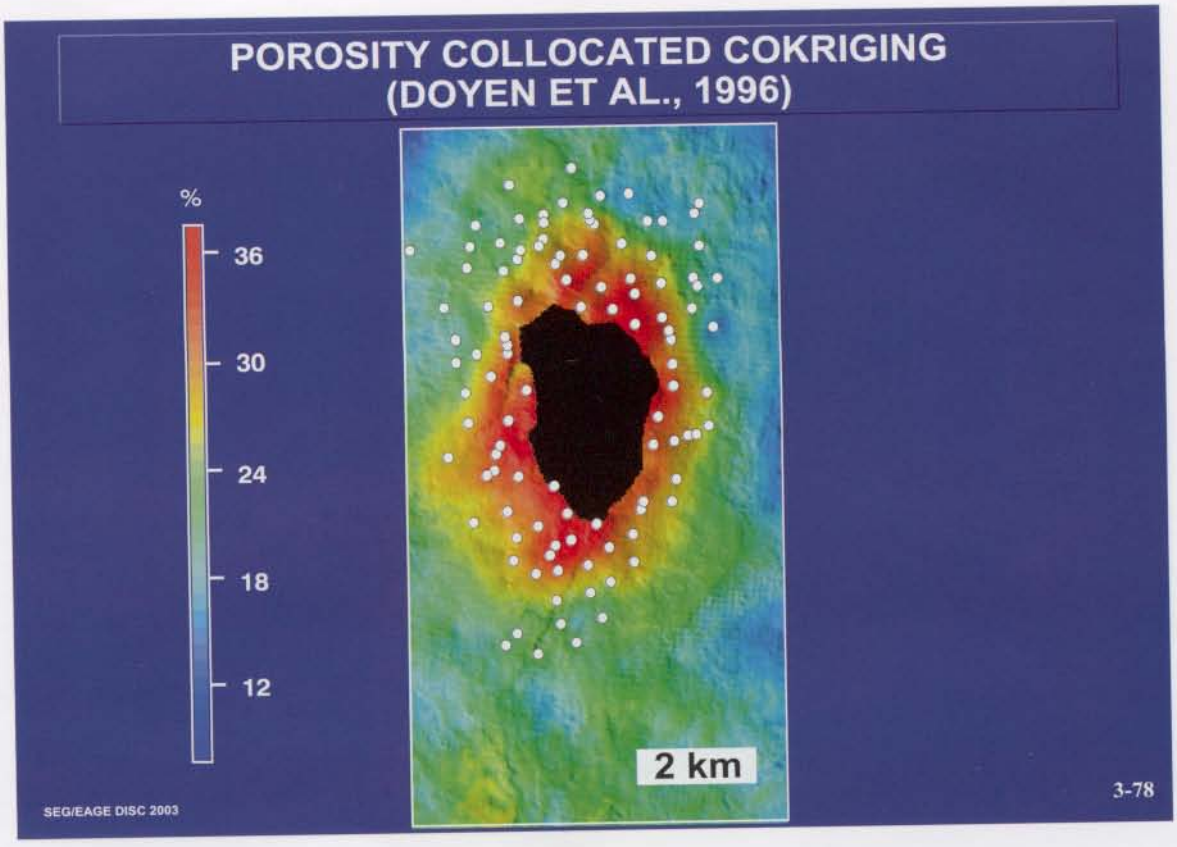
CCK and KED are the preferred industry solutions for combining well and seismic-derived data. The discussion of the North Alwyn case and the examples used showed that there could be a strong similarity between a map obtained with CCK and a map obtained with KED. A simple mathematical development (Fig. 3-82) confirms that the analytical expressions are very similar, even if the system of equations that leads to this expression is different for CCK and KED. However, the main difference lies in the assumptions that each method makes and that definitely impact one's choice of the variogram model.

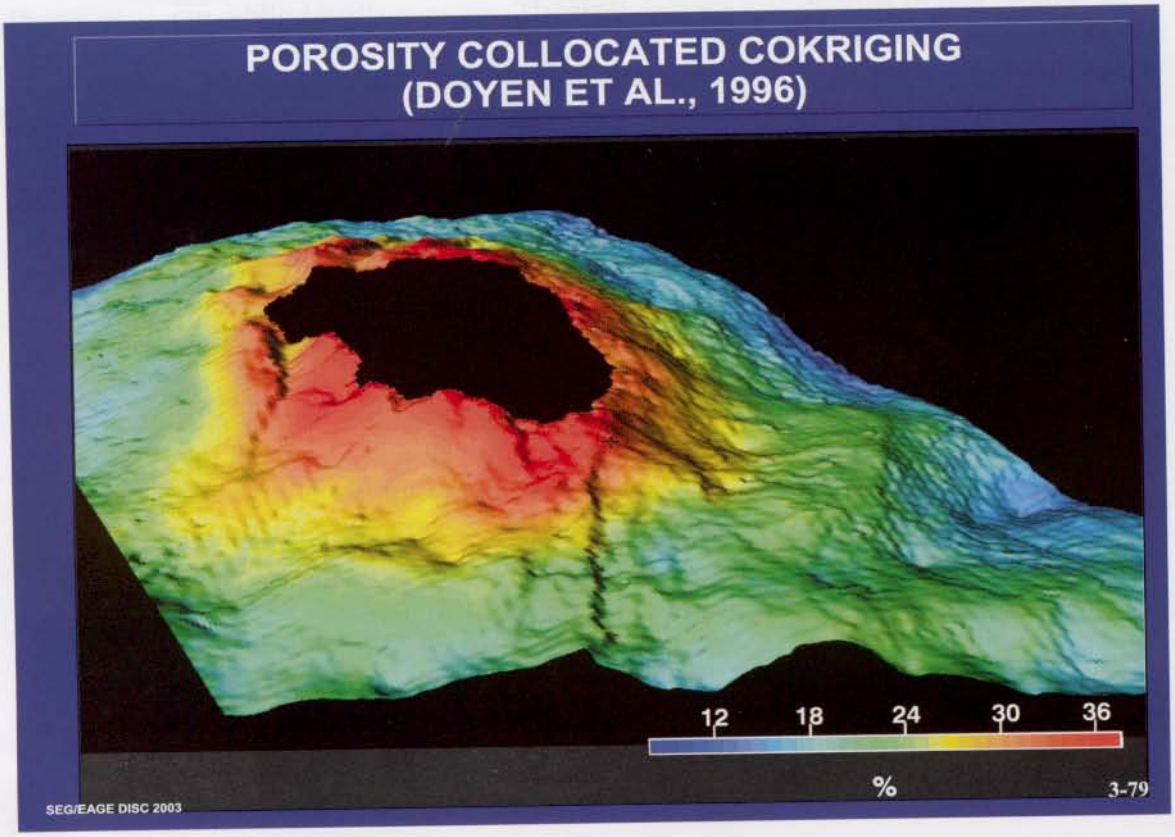
With KED, we assume that the secondary variable provides low-frequency information about the primary one. The variable of interest, $Z(\mathbf{x})$, is modeled as the sum of a linear function of the external drift plus a stationary residual of mean zero. As a result, the KED interpolation can be decomposed into the sum of a linear function of the external drift — which does not have to be stationary — plus a residual that tends toward zero at large distances. But how confident are we about this external drift? Calculation of the estimation error variance attached to KED shows that the external drift is assumed to be perfectly known.



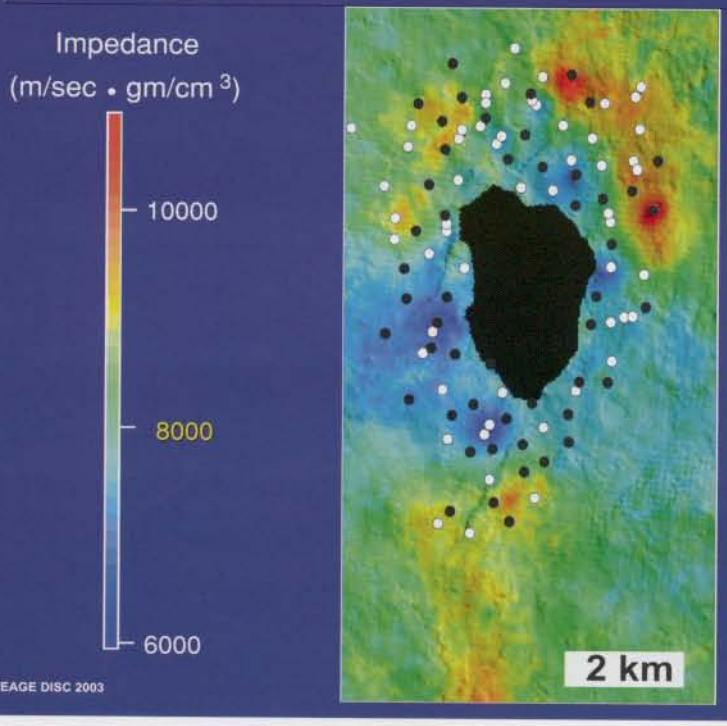
Only Bayesian kriging generalizes KED to the situation where the coefficients of the external drift itself are uncertain. This is a limitation of KED, considering typical situations that are encountered today — a few wells, often fewer than 10, and an interpreted seismic map. With KED, the only calibration of the wells against the seismic consists, after applying a linear transformation to the seismic data, of calculating the variogram of the residuals. But is this a reliable thing to do from just a few wells? It seems that a case where the use of KED appears defendable is time-to-depth conversion, where we are dealing with depth at the wells and time horizons, or average velocities and stacking velocities. In this situation, there is a physical reason why the seismic map should be related with the well data. Also, the map derived from seismic is often smooth (assuming, in the case of V_{stack}, that error cokriging has been applied in a previous step). When KED is applied, the variogram to use should be that of the residuals from the external drift fitted to the well data, as in the examples of Figs. 3-58 and 3-62. This variogram, which controls through its range how soon extrapolation will become equal to the external drift (Fig. 3-55), should not be the model directly derived from that of the primary or of the secondary variable.

With CCK, the approach is different and is based on a crossplot. A calibration step is performed, before interpolation, through the calculation of the correlation coefficient. The variances of the secondary and primary variables, and their relative values, also play a role in the shape of the interpolating function. As clearly shown by Doyen et al. (1996), the value of the correlation coefficient calibrates the relative weight of kriging


POROSITY COLLOCATED COKRIGING
(DOYEN ET AL., 1996)
%
36
30
24
18
12
2 km
SEG/EAGE DISC 2003
3-78


POROSITY COLLOCATED COKRIGING
(DOYEN ET AL., 1996)
12
18
24
30
36
%
SEG/EAGE DISC 2003
3-79

BLIND WELLS EXERCISE (DOYEN ET AL., 1996)



- 50 hidden wells completed after 1989

o 57 control wells completed before 1989

SEG/EAGE DISC 2003

3-80

COMPARISON OF PREDICTION ERROR (DOYEN ET AL., 1996)

SEG/EAGE DISC 2003

3-81

COMMON EXPRESSION FOR CCK AND KED

- **Both KED and CCK can be written in the interpolation formalism (using a 2D isotropic assumption):**

$$Z^*(x,y) = a_0 + a_1 S(x,y) + \sum_{i=1}^{N} b^i C\left[\sqrt{(x-x_i)^2 + (y-y_i)^2}\right]$$

- **Both techniques result in the sum of a linear function of the seismic-derived parameter and a residual that tends to zero away from the data!**

SEG/EAGE DISC 2003

3-82

versus information derived from seismic. This means that the secondary variable does not have to be smooth, as in the case of KED. The estimation variance produced by CCK accounts for the variations of the secondary variable through its variance. This is why many CCK applications consist of predicting a petrophysical parameter from a seismic attribute (sometimes quite random), while KED tends to be more suitable for structural applications.

When CCK is applied under the Markov hypothesis, only the variogram of the primary variable is needed, because the cross-variogram is assumed to be proportional to it. If not enough well data are available to derive the primary-variable variogram, the secondary-variable variogram can be used to help define some of its parameters (Fig. 2-39).

Again, CCK should be preferred when we are mapping attributes, whereas KED should be preferred when mapping structural parameters. CCK also provides more flexibility, thanks to the choice of the correlation coefficient, which means that, if no KED algorithm is available, one can obtain an interpolation close to KED using CCK. Coleou (2002a) demonstrated that, in the situation where all covariances are proportional, MCCK is equivalent to KED. This assumption can even be relaxed. Haas et al. (1998) show that, if there is proportionality between the cross-covariance model and the covariance of the secondary variable, then MCCK and KED are identical. Haas calls this situation "geostatistical regression." KED is optimal when the residual is uncorrelated with the secondary variable. Rivoirard (2002) has recently shown that this corresponds

to a situation where the cross-covariance is proportional to that of the secondary variable. Haas et al. (1998) call this situation "geostatistical regression," whereas Journel (1999) calls it "Markov Model 2" and Chilès and Delfiner (1999) call it "Reverse Markov."

3.6.6 Factorial cokriging

Coleou (2002b) recently presented a new application of cokriging, this time in combination with factorial kriging. This application may prove useful in the context of time-lapse seismic. Assume that two seismic surveys have been obtained on the same area. The model of factorial kriging is recalled in Fig. 3-83 on a synthetic example. Now, suppose we have the two surveys only, and we want to extract some reservoir information from these surveys. One approach would be to apply FK independently to each of them. Coleou proposes to improve the consistency by applying factorial cokriging simultaneously to both surveys, in order to obtain what he calls the "common part" of these two surveys (Fig. 3-84). The equations are not presented here, but they are a simple generalization of those of cokriging and FK.

Fig. 3-85 shows a promising application of factorial cokriging on rms amplitude in a horizon-controlled window above the target and on one of the time horizons. The continuity outside the 4D signature is reduced, thereby providing a better delineation of the 4D effect. Organized differences are only found at and below the fluid changes, where they are expected. Although it is a bit early to tell, factorial cokriging appears to have some promising properties for use in multisurvey filtering (Fig. 3-86).

3.7 Some Relationships between the Kriging Techniques

Fig. 3.87 lists a number of relationships between all the kriging methods presented so far. We will not discuss the mathematics behind these different relationships, which should appear logical after what we have discussed. Coleou (2002a) discusses the mathematics.

3.8 Kriging Versus Other Interpolation Techniques

3.8.1 Introduction

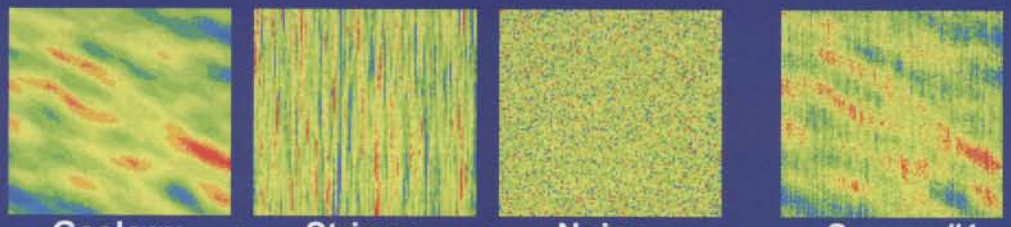
In section 3.2.4, we presented kriging from another perspective, that of an interpolating function. In this chapter, we will limit our discussion to mapping problems and consider kriging as a function $z_{uk}(x,y)$ of the coordinates (x,y) of any 2D location. We will show that the expression of kriging given in Fig. 3-18 clarifies the relationships between kriging and many other mapping techniques, especially splines. First, let us explain splines.

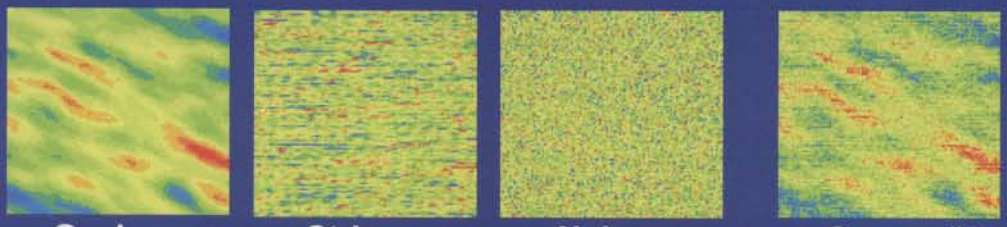
3.8.2 Splines

- Biharmonic (thin-plate) and harmonic (membrane) splines

With thin-plate splines, the idea is to calculate an interpolating function that mimics the shape of an elastic plate that would be forced to honor the data points. It can be

4D COMMON PART EXTRACTION: THE MODEL (COLEOU, 2002b)

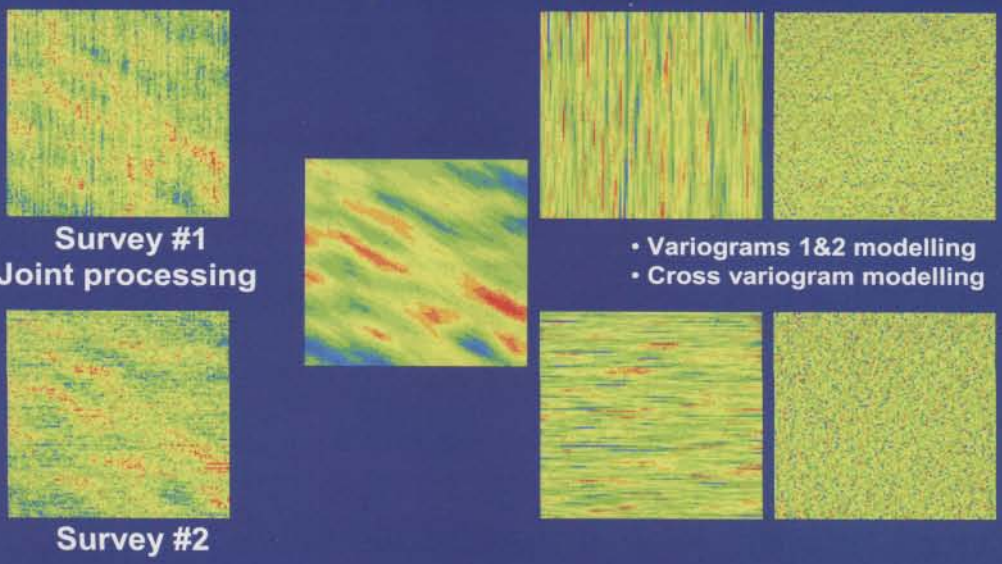




SEG/EAGE DISC 2003

3-83

FOR BETTER CONSISTENCY: JOINTLY APPLY FACTORIAL COKRIGING (COLEOU, 2002b)



Efficient removal of acquisition imprints

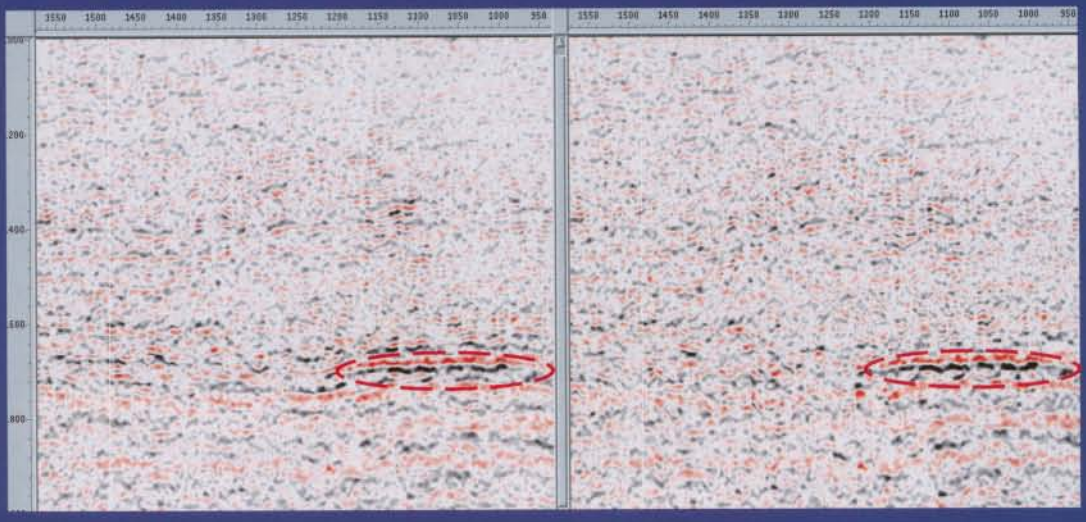
SEG/EAGE DISC 2003

3-84

WHAT HAPPENS TO THE 4D SIGNAL? (COLEOU, 2002b)

Difference before factorial cokriging

Difference after factorial cokriging



SEG/EAGE DISC 2003

3-85

FACTORIAL COKRIGING APPLIED TO 4D SEISMIC POTENTIAL BENEFITS (COLEOU, 2002b)

Balancing surveys

- Extraction of a common part, a better candidate for a reference than any of the surveys
- Removing spatially uncorrelated variations

Reducing random noise and organised noise level

- Removing spatially organised but uncorrelated variations
- Providing spatial repeatability measurement

SEG/EAGE DISC 2003

3-86

shown (Briggs, 1974) that the function minimizes the integral given in Fig. 3-88, which is representative of the bending energy of an elastic thin plate. The surface displayed in Fig. 3-88 is based on the same data points as those of Figs. 3-19 to 3-21. Biharmonic spline interpolation is smooth and tends to overshoot when there are significant variations of the interpolated variable between closely spaced data points. This is not surprising, because one of the assumptions of the method is that the interpolated surface is twice differentiable. This feature may prove unacceptable for some applications.

Another "natural" shape that comes to mind is that of a membrane forced to honor the data points (Fig. 3-89), which correspond to harmonic splines. The map obtained with harmonic splines appears unsatisfactory for mapping applications, because it tends to be very smooth between the data points and to "cone" too much at the data point locations. To our knowledge, this "membrane" interpolating function has found no petroleum application. However, we will see below that it is used to describe many physical phenomena and that it has interesting relationships with fractal models.

Duchon (1975) showed that the harmonic and biharmonic spline functions are solutions of the equations shown in Fig. 3-90. Both harmonic and biharmonic splines converge in extrapolation toward their polynomial term — which is constant for harmonic splines and linear for biharmonic splines.

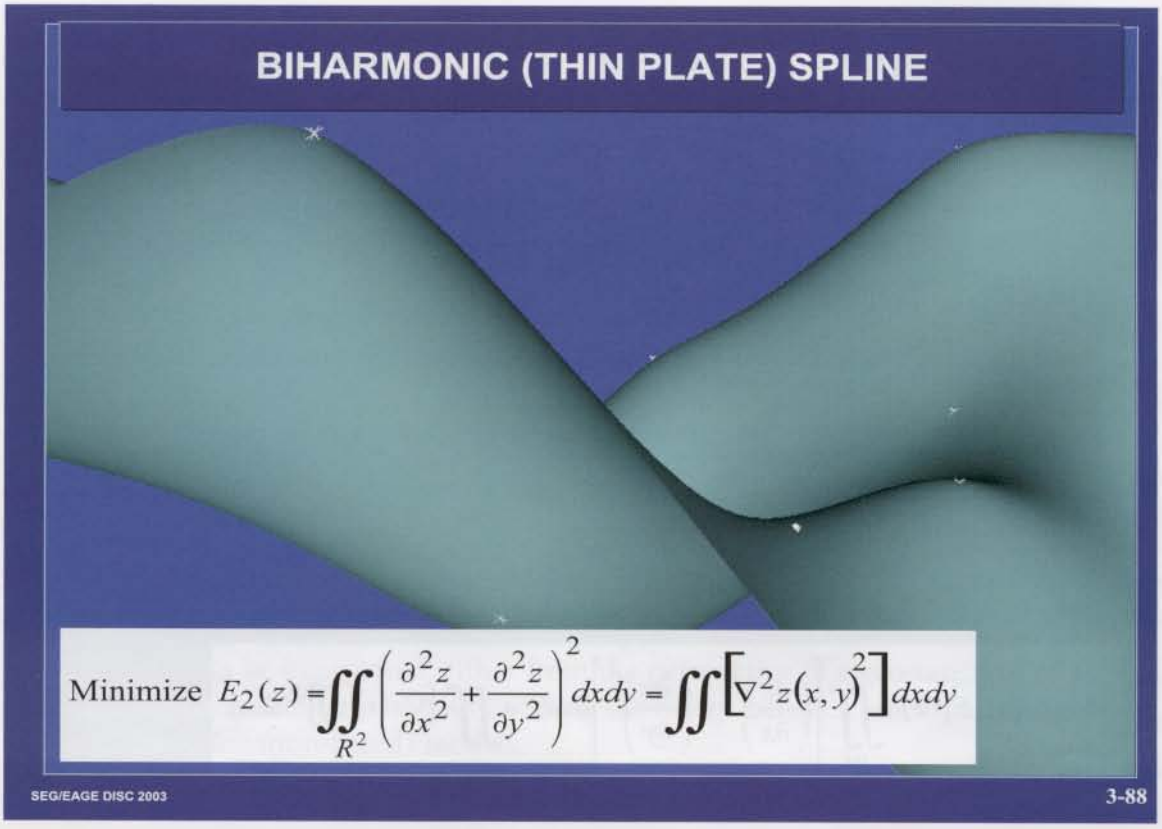
- Other spline functions

Mitas and Mitasova (1988), Mitasova and Mitas (1993), and Wessel and Bercovici

RELATIONSHIPS BETWEEN KRIGING INTERPOLATORS (COLEOU, 2002a)

- **Simple kriging is associated with zero-mean model.**
- **Universal kriging and kriging with an external drift are both equal to the trend calculated by kriging plus simple kriging of residual from this trend.**
- **When using coordinates instead of seismic data as an external drift, kriging with an external drift is equivalent to universal kriging.**
- **Bayesian kriging is very general and can replicate other methods: simple kriging (coefficients of mean fixed to zero), universal kriging or kriging with an external drift (no constraints on trend coefficients).**
- **Multi-collocated cokriging can replicate kriging with an external drift if the covariance models are all proportional to each other.**

SEG/EAGE DISC 2003 3-87



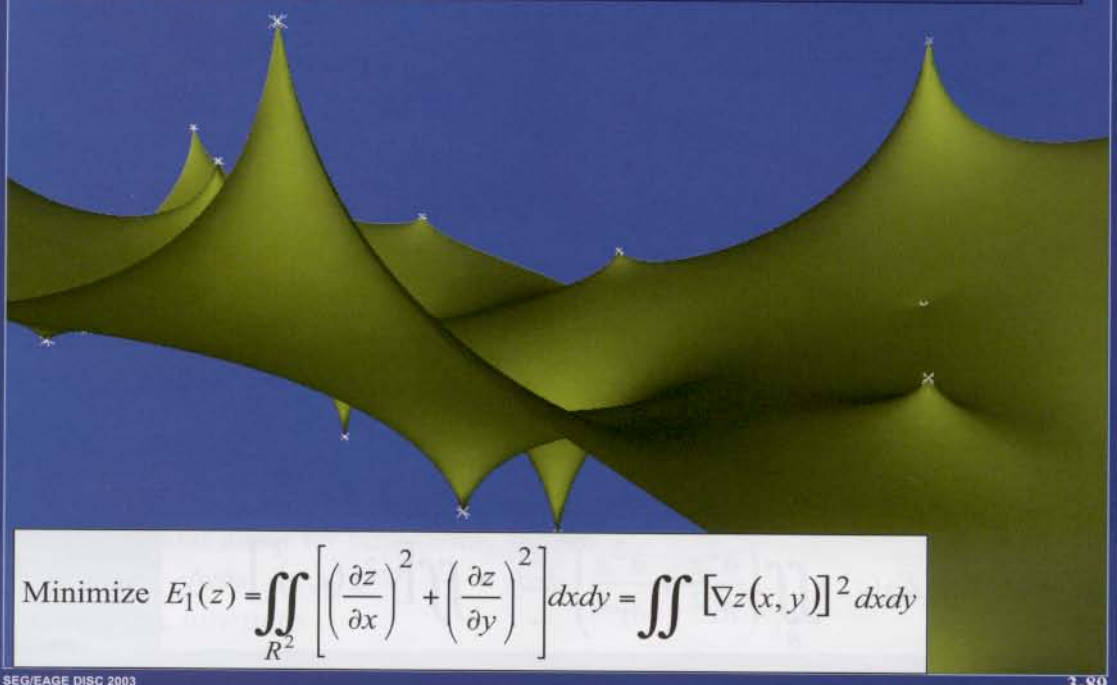
(1998) propose a more general definition of the energy functional than that of harmonic and biharmonic splines. They derive general (and complicated) analytical expressions for the interpolating function, minimizing this generalized functional with the constraint of honoring the data points. Because both the flexural rigidity of the plate and the tension are involved, this leads to interpolating functions that combine the advantages of harmonic and biharmonic splines. However, their complexity, and the number of parameters involved, are an issue.

- Smoothing splines

The interpolating spline method can be generalized to the smoothing of data affected by a measurement error. Instead of simply minimizing the energy functionals $E_1(z)$ and $E_2(z)$, as described in the previous sections, a sum of this functional plus a distance to the data is now minimized (see Fig. 3-91). This means that, contrary to the previous case, the interpolating function is not forced to exactly honor the data points. A certain tolerance is allowed, depending on the magnitude of the measurement error affecting the data, z_i.

A source of great discussion in the literature on interpolation is the choice of the coefficient θ, balancing the weight of the energy (or regularization) constraint versus the distance to the data points. The constant θ is usually calculated by cross-validation. Data points are dropped one after the other, and the selected value is that for which the estimation of the dropped data point, by use of the other points, is best.

HARMONIC (MEMBRANE) SPLINE



$$\text{Minimize } E_1(z) = \iint_{R^2} \left[\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 \right] dxdy = \iint \left[\nabla z(x,y) \right]^2 dxdy$$

SEG/EAGE DISC 2003

3-89

THE HARMONIC AND BIHARMONIC SPLINE MODELS

Harmonic splines

$$z_{hs}(x,y) = a_0 + \sum_{i=1,N} b^i f_1 \left[\sqrt{(x-x_i)^2 + (y-y_i)^2} \right] \text{ with } f_1(h) = Logh$$

(N+1) equations on the (N+1) coefficients:

$$\begin{cases} \sum_{i=1,N} b^i = 0 \\ z_{hs}(x_i, y_i) = z_i \end{cases}$$

Biharmonic splines

$$z_{bhs}(x,y) = a_0 + a_1 x + a_2 y + \sum_{i=1,N} b^i f_2 \left[\sqrt{(x-x_i)^2 + (y-y_i)^2} \right] \text{ with } f_2(h) = h^2 Logh$$

(N+3) equations on the (N+3) coefficients:

$$\begin{cases} \sum_{i=1,N} b^i = \sum_{i=1,N} b^i x_i = \sum_{i=1,N} b^i y_i = 0 \\ z_{hs}(x_i, y_i) = z_i \end{cases}$$

SEG/EAGE DISC 2003

3-90

3.8.3 *Comparison among kriging, splines, and radial-basis functions*

- Kriging and radial-basis functions

It is now obvious (simply compare Figs. 3-18 and 3-90) that splines and kriging functions share a similar expression, which is that of the general class of radial-basis functions. See, for instance, Franke and Nielson (1991) or Ahmed and Murthy (1997). Fig. 3-92 gives the generic expression of radial-basis-function interpolation and explains how each term is interpreted, depending on whether we are working from a spline or from a kriging perspective.

The multiquadric interpolation method of Hardy (1990), widely used in practice, is also a special case of radial-basis functions. It is a significant result, from a mathematical and computer point of view, that such a variety of interpolating functions can be reduced to a similar formalism. This will now allow us to throw a new light on a number of results.

- Special cases of radial-basis functions

In the following, we will look at a number of special cases of radial-basis functions and see how they can be interpreted, either from the perspective of a geostatistical or a spline model. This will help us start building the bridge between stochastic and energy-based (or regularization-based) models.

- *De Wijs variogram and harmonic functions*

Kriging with a constant trend and a De Wijs variogram is equivalent to 2D interpolation with harmonic splines. This variogram (Log$h$) has a slope of infinity at the origin, which corresponds to an extreme case of spatial variability for a model with zero nugget effect. As a result, kriging honors the data but rapidly returns to the mean. This explains the somewhat peaked behavior of the interpolating function around the data points (Fig. 3-89).

From the spline perspective, the interpolation surface (Fig. 3-89) is peaked around the data points, since it represents the shape of a membrane — a kind of tent — going through these data points. Away from the data, the membrane tends toward a flat surface. This interpolating function is a solution of the Poisson equation, which is used to describe many physical processes. It describes the shape of a soap film (Isenberg, 1992), of a thin rubber sheet (a membrane) forced to pass through fixed points, or the value of an electrostatic potential, over a surface, due to a two-dimensional finite distribution of charges (Feynmann et al., 1964, Chapter 12). Given a set of control points, the interpolating function takes the shape that minimizes its surface area. This is the familiar property of minimal surfaces, which is represented in nature by the shape of soap films.

Thanks to its harmonic property, Log$h$ is also associated with the familiar power spectrum of fractals (Fig. 3-93). This is no surprise, because we already saw (read, for instance, Section 2.4) that there was a close relationship between De Wijs's variogram and fractals.

- *Spline covariance and biharmonic equation*

Kriging with a linear trend and a generalized covariance equal to h^2Log$h$ is equiv-

THE SMOOTHING SPLINE MODEL (2D BIHARMONIC CASE)

Smoothing splines minimize:

$$E_2(z) + \theta \sum_{i=1,N} \frac{[z(x_i, y_i) - z_i]^2}{Var(\varepsilon_i)}$$

Their expression is the same as interpolating splines:

$$z_{ss}(x, y) = a_0 + a_1 x + a_2 y + \sum_{i=1,N} b^i f_2\left[\sqrt{(x - x_i)^2 + (y - y_i)^2}\right] \text{ with } f_2(h) = h^2 Logh$$

With (N+3) equations on the (N+3) coefficients:

$$\sum_{i=1,N} b^i = \sum_{i=1,N} b^i x_i = \sum_{i=1,N} b^i y_i = 0$$

$$z_{ss}(x_i, y_i) + b^i \frac{Var(\varepsilon_i)}{\theta} = z_i$$

SPLINES AND KRIGING TWO EXAMPLES OF RADIAL BASIS FUNCTIONS (RBF)

Interpreted as kernel of minimized operator by splines

Interpreted as Green function by splines

$$z_{rbf}(x, y) = Polynomial(x, y) + \sum_{i=1,N} b^i C\left[\sqrt{(x - x_i)^2 + (y - y_i)^2}\right]$$

Interpreted as trend by kriging

Interpreted as covariance by kriging

alent to 2D interpolation with biharmonic splines. This generalized covariance of order 1 is associated with a smoother random function than the De Wijs or the linear variogram model. As a result, the interpolation function is much smoother (Fig. 3-88). Away from the data, the map tends toward the value of a linear trend fitted through the data points.

From the spline perspective, the interpolation surface is a solution of the biharmonic equation, corresponding to the shape of an elastic plate forced to honor the data points at their location. Because it is an elastic plate rather than a membrane, the shape is very smooth close to the data points. Away from all data constraints, the thin plate spline comes back to an unconstrained flat (but not necessarily horizontal) shape.

Because of its biharmonic property, $h^2 \mathrm{Log} h$ is also associated with the familiar power spectrum of fractals (Fig. 3-93). The fractal nature of the covariance associated with the biharmonic spline has been recognized by Szeliski and Terzopoulos (1989).

- *Smoothing splines and error cokriging*

Now, let us examine the situation where the data points are affected by random measurement errors. This will lead the spline specialist to use smoothing splines, as presented in Fig. 3-91, and the geostatistician to use error cokriging as presented in section 3.3.1. If we express error cokriging as a function, it is easy to verify that there is equivalence between smoothing splines and error cokriging, as long as we use the trend and covariance models just discussed.

But, in the error cokriging framework, what happens to the important smoothing spline coefficient θ, which determines the relative weight of smoothing versus distance to the data? It has been shown by Matheron (1981a) that error cokriging using the covariance functions given in Fig. 3-94 is equivalent to calculating the smoothing spline associated with the same parameter θ. Thus, in geostatistical formalism, the problem of weighting the smoothness term versus the distance to the data is not raised. By fitting a covariance model and its coefficient to the experimental data, error cokriging automatically calibrates the variance of the measurement errors versus that of the covariance function.

- Recent developments

When we discussed kriging in a global neighborhood (Section 3.2.2), we stressed the fact that, due to computational limitations, this technique could only be used with fewer than, say, 1000 data points. So far, this limitation has also applied to radial-basis functions, which could not be used with more and a few thousand data points. Two recent papers in GEOPHYSICS challenge this view. Billings et al. (2002a, b) review the relationships between radial-basis functions, splines, and kriging, both for interpolation and smoothing purposes. They define continuous global surfaces as those surfaces satisfying the equations of Fig. 3-92. They use this new terminology because they consider that the class of radial-basis functions only encompasses those functions for which basis functions are radially symmetric. Billings et al. (2002a) convincingly promote and demonstrate the use of iterative techniques and efficient preconditioners for solving global neighborhood problems involving millions of data points. Billings et al. (2002b)

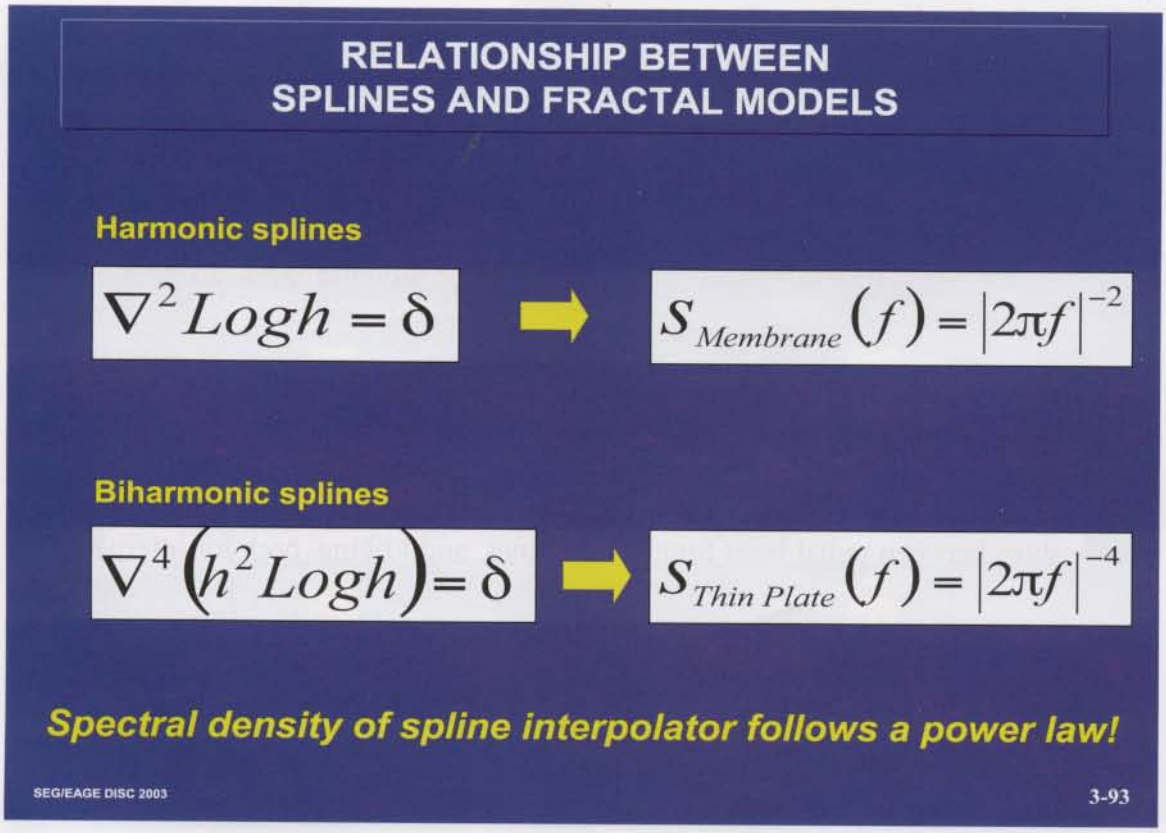
prefer to use smoothing splines rather than kriging, and recommend the use of generalized cross-validation for estimating the smoothing spline coefficient θ. If the approach proposed by Billings et al. (2002a) proves general enough, kriging in a global neighborhood may generalize to very large data sets, with a significant impact on many geostatistical algorithms, such as those used for geostatistical conditional simulation (see below).

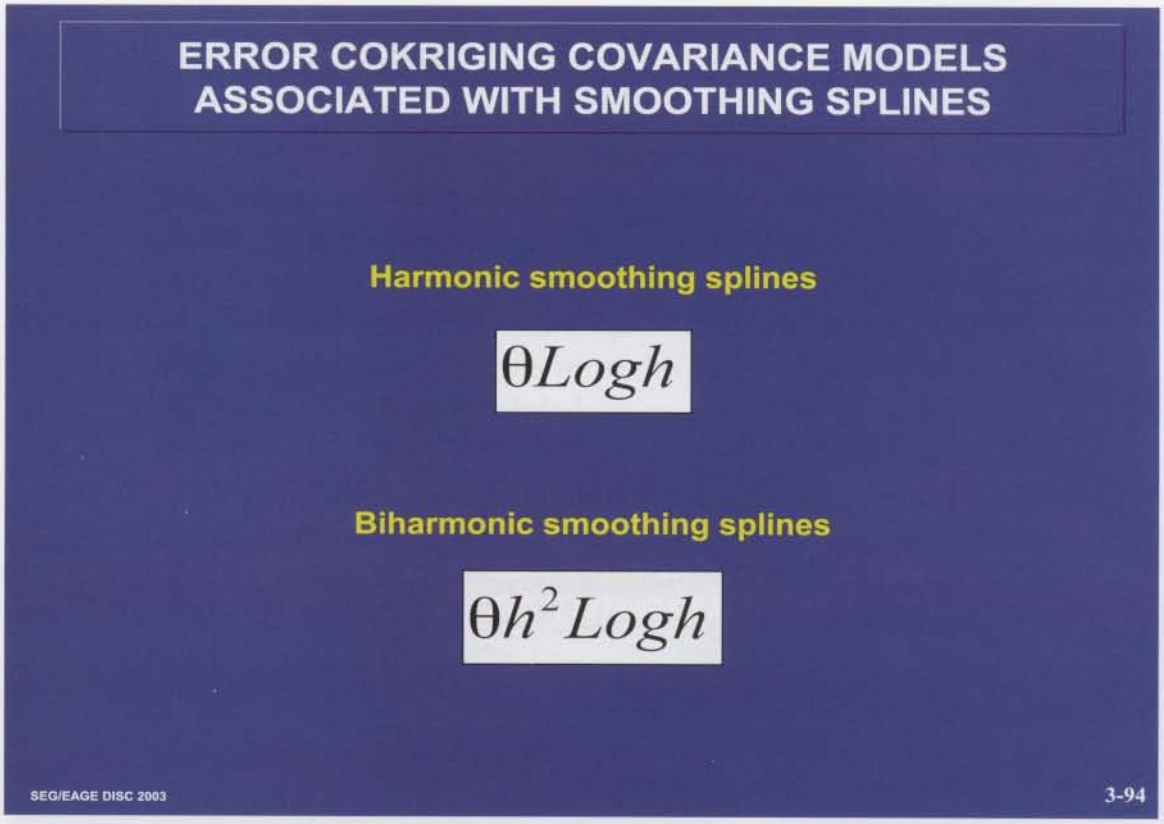
- Conclusion

Interpolating splines, in 2D, consist of the calculation of an interpolating function that minimizes an energy functional related to the stretching or a bending energy of a plate. The choice of such an energy functional is equivalent to fixing the degree of the trend function and the covariance model for kriging. In other words, fixing the energy — or regularization — term of splines is equivalent to fixing the a priori model for kriging.

Fig. 3-93 shows that there is a fundamental, inverse relationship between the spline functional and the covariance. Kimeldorf and Wahba (1970) demonstrated that this fundamental relationship also applies in the frame of discrete Bayesian statistics. The consequence of this relationship on the spectral density is straightforward. The spectral densities associated with the harmonic and biharmonic splines are power laws, which means that they represent fractal models. Szeliski and Terzopoulos (1989) convincingly demonstrate this relationship between splines and fractals.

Smoothing splines, in 2D, consist of the calculation of a function that minimizes





the sum of an energy functional — or regularization term — plus a distance to the data. This is equivalent to calculating the error cokriging interpolation associated with the a priori model corresponding to the regularization term. The coefficient of the covariance model is equal to the coefficient weighting the distance to the data against the regularization term. However, in the error cokriging formalism, this term is not arbitrary but instead results from the fit of the covariance model to the data.

Why this relationship between splines and kriging?

Kriging and splines originated from different worlds, that of approximation theory (splines) and that of probabilities (kriging). We have demonstrated the formal equivalence between the two methods. So far, this equivalence is purely formal and does not provide a fundamental clue of why this bridge exists between the two formalisms. The question is actually to relate an energy- or regularization-based formalism, to a probability-based one.

The idea of converting an energy function into a probability distribution comes from statistical mechanics, because the probability of a particular configuration is inversely related to its energy. Suppose we calculate the function $z(\mathbf{x})$ minimizing an energy functional $E(z)$. Using the results of Geman and Geman (1984), Szeliski and Terzopoulos (1989) associate a probability to this energy through the Boltzmann (or Gibbs) distribution, $p(z)$, defined as

$$p(z) = \frac{1}{Z} \exp\left[-\frac{E(z)}{T}\right]$$

(where Z and T are positive constants).

If we now have a number N of data z_i available, and if we make, as above, the assumption that these data are affected by independent and Gaussian measurement errors ε_i, of variance σ_i^2,

$$p(z_i/z) \propto \exp\left[-\frac{(z_i - z)^2}{2\sigma_i^2}\right]$$

This is the likelihood function, which we can combine, using Bayes's theorem, with the above a priori probability, to obtain the posterior:

$$p(z/z_i) \propto p(z_i/z)\, p(z)$$

In the Gaussian formalism, we wish to maximize the posterior to obtain the best estimate of z. This is equivalent to maximizing the logarithm, and we obtain the smoothing splines functional:

$$E(z) + \theta \sum_{i=1}^{N} \frac{(z_i - z)}{\sigma_i^2}$$

Fig. 3-95 summarizes the conclusion of these developments.

FROM ENERGY-BASED WORLD TO PROBABILISTIC WORLD THANKS TO THE GIBBS DISTRIBUTION

Energy

$$p(z) = \frac{1}{Z} \exp\left(\frac{-E(z)}{T} \right)$$

Probability

Using a regularization-based smoothness constraint is equivalent to using a correlated Gaussian field as the Gaussian prior

SEG/EAGE DISC 2003

3-95

4 Conditional Simulation for Heterogeneity Modeling and Uncertainty Quantification

4.1 Introduction

In the previous chapter, kriging proved to be an interpolation technique that was flexible enough to filter correlated or uncorrelated noise from the data or to combine seismic and well information. Thanks to the flexibility in the choice of the trend and the covariance model, kriging is closely related with splines, multiquadrics, and trend-surface analysis. Kriging, when it is based on the trend and covariance models fitted to the data, also provides an estimate of the uncertainty at every location of the map. However, kriging remains a deterministic approach that provides a very smooth image of geological variables that are, in most cases, very erratic.

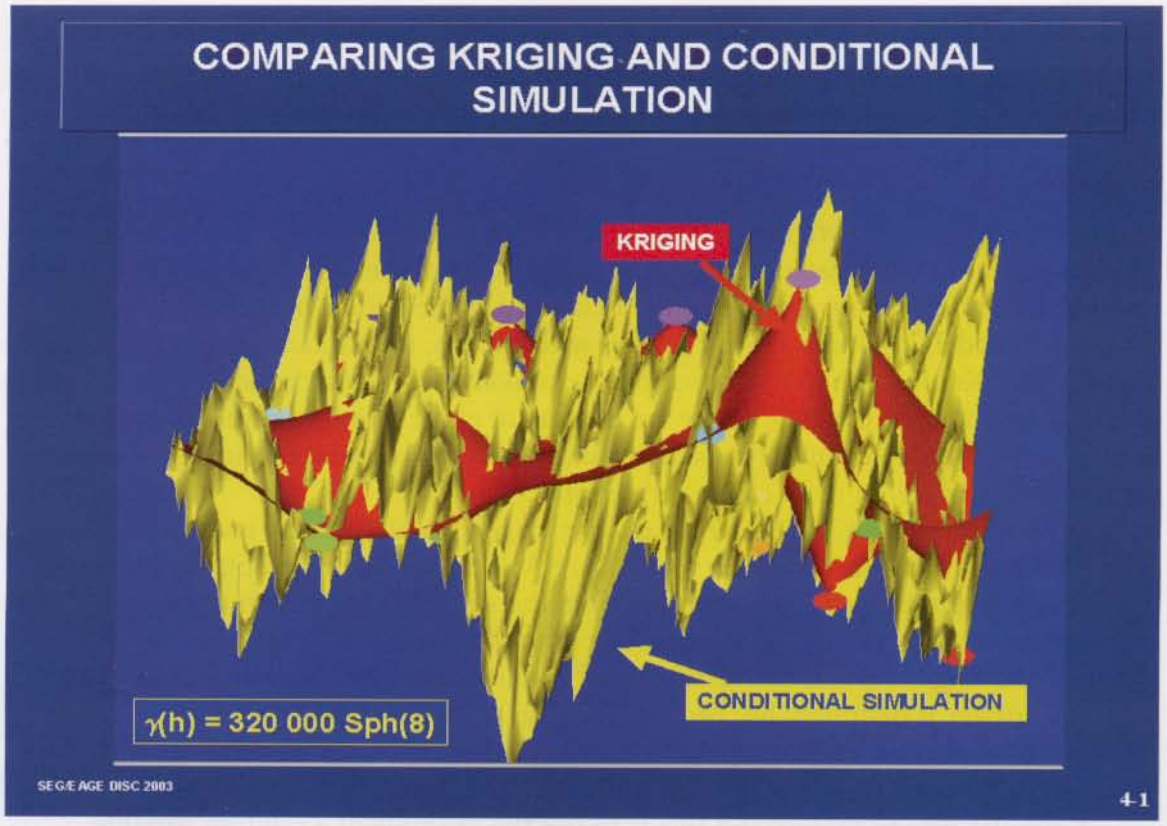
As an example, Fig. 4-1 shows in red the kriged surface already used in Fig. 3-24 (Hohn data set). This surface is very smooth. A variogram calculated on the points of this surface would be extremely different from the spherical variogram fitted on the data and used as input to kriging. Is there not a contradiction here? Should not the variogram of the kriged surface be the same as the spherical model used as input to kriging? The answer is definitely no, because the goal of kriging is not to generate a surface that mimics the actual variations of the interpolated variable, but to provide, at each location, an estimate that is as close as possible — on average — to the unknown value.

A simpler way to understand this is to use the example of a random variable, Z, taking the value –1 or +1, depending on the outcome of a game of heads or tails. A game will generate a sequence of values –1 and +1 (Fig. 4-2). The variogram is obviously a nugget effect, since two draws are uncorrelated with each other. At locations with no data points, kriging would be equal to the mean, which is zero (this also corresponds to the trend-surface-analysis estimate in the case of a constant trend; see Fig. 1-39). Thus, in the case of a pure nugget effect — the noisiest parameter one can think of — kriging is constant! This is hardly a representation of the actual heads-or-tails outcome! But still, as we saw earlier in Fig. 1-23, the mean is the parameter that is, on average, the closest to all possible values of the distribution.

The goal of geostatistical conditional simulation (GCS) is precisely to generate samples of surfaces or of 3D earth models that satisfy the input statistics (mean, variance, and variogram) instead of smoothing them, and that honor the data points. The yellow surface of Fig. 4-1 is a realization of GCS on the Hohn data set. This surface is far noisier than the red kriged surface, but its experimental variogram is equal to the spherical model fitted to the data. Before we discuss GCS in more detail, a few reminders about Monte-Carlo simulation may prove useful.

4.2 A Few Reminders on Monte-Carlo Simulation

Monte-Carlo simulation (MCS) is a powerful technique for calculating the pdf of the combination of several random variables. Take the example of Fig. 4-3. Standard original oil-in-place (STOOIP) is a function of gross-rock volume (GRV), porosity (Φ), net-



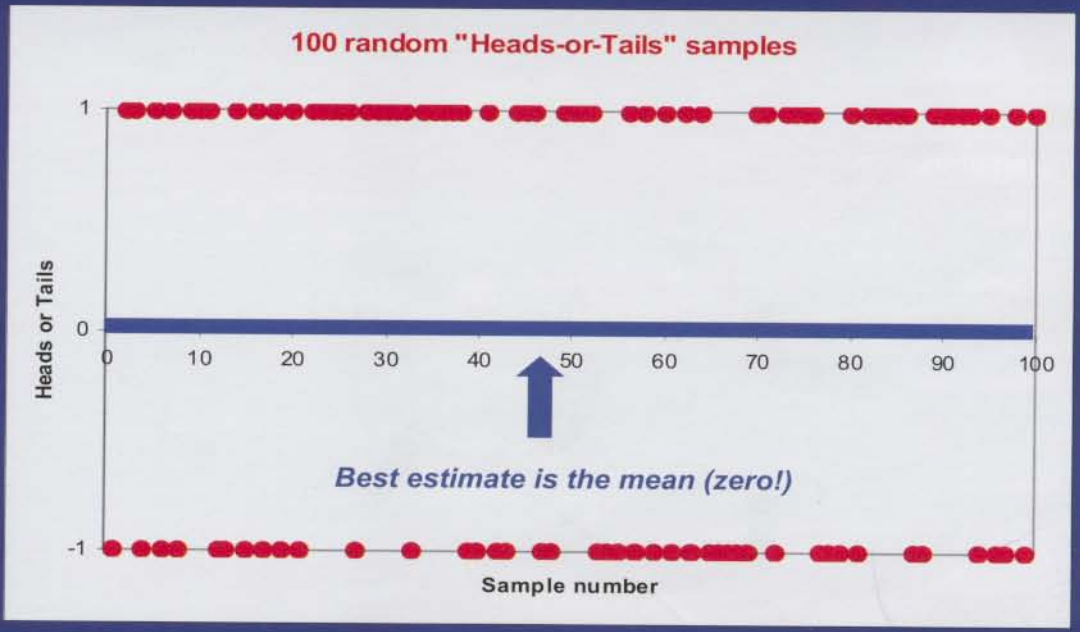
to-gross (N/G), water saturation (S_w), and formation volume factor B_o. If the pdf associated with each of these parameters is known, MCS consists simply of sampling each parameter a large number of times and calculating the STOOIP for each set of samples. Then the pdf of STOOIP can be derived from the histogram of the resulting STOOIP distribution.

Nowadays, spreadsheet add-ons are available on the market that provide a very easy user-interface to MCS. As a result, MCS is a generic approach now used in a large number of industries to quantify uncertainties (Fig. 4-4).

Figs. 4-5 and 4-6 explain some of the technicalities of MCS. We said earlier that MCS consists of sampling each parameter according to its input pdf. For continuous parameters, this sampling can be performed because of a very general result of probability theory. Fig. 4-5 explains how a sample from a pdf can be obtained by first sampling a uniform distribution between 0 and 1, then applying the inverse of the cdf to the outcome.

Often, we will also need to sample discrete random variables, such as lithology. For instance, we may know that the lithology at one location is either sandstone, shaly sandstone, silty shale, or shale. The four probabilities of occurrence add up to one and can be translated in four colors that each occupy a fraction of a segment of length 1 (Fig. 4-6). As in the previous case, a uniform distribution sampling can be used. The lithology outcome is then simply the one corresponding to the color in which the sample of the uniform distribution falls.

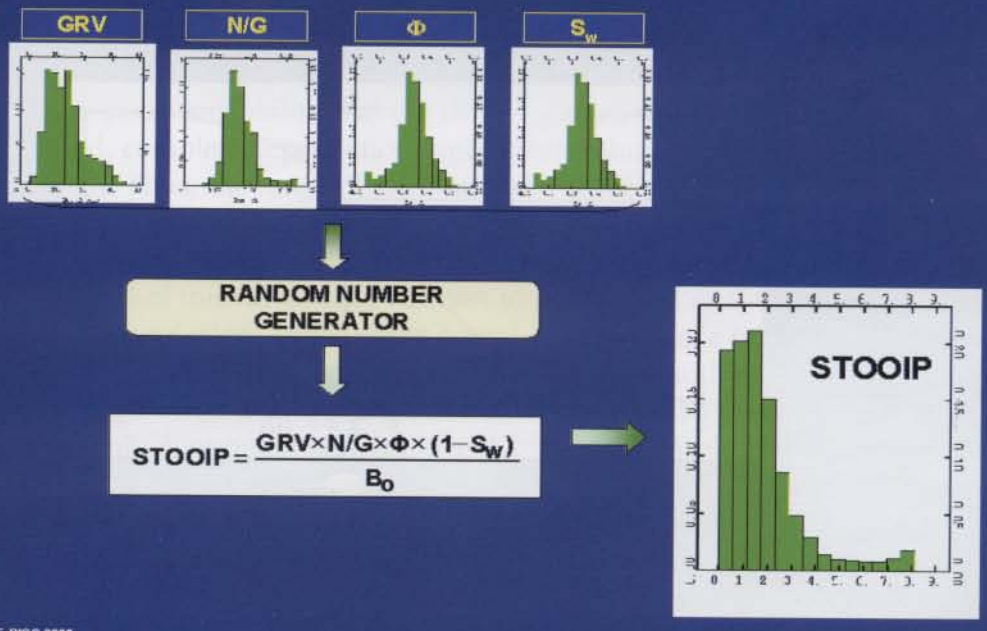
KRIGING IN THE FRAME OF A HEADS-OR-TAILS GAME



SEG/EAGE DISC 2003

4-2

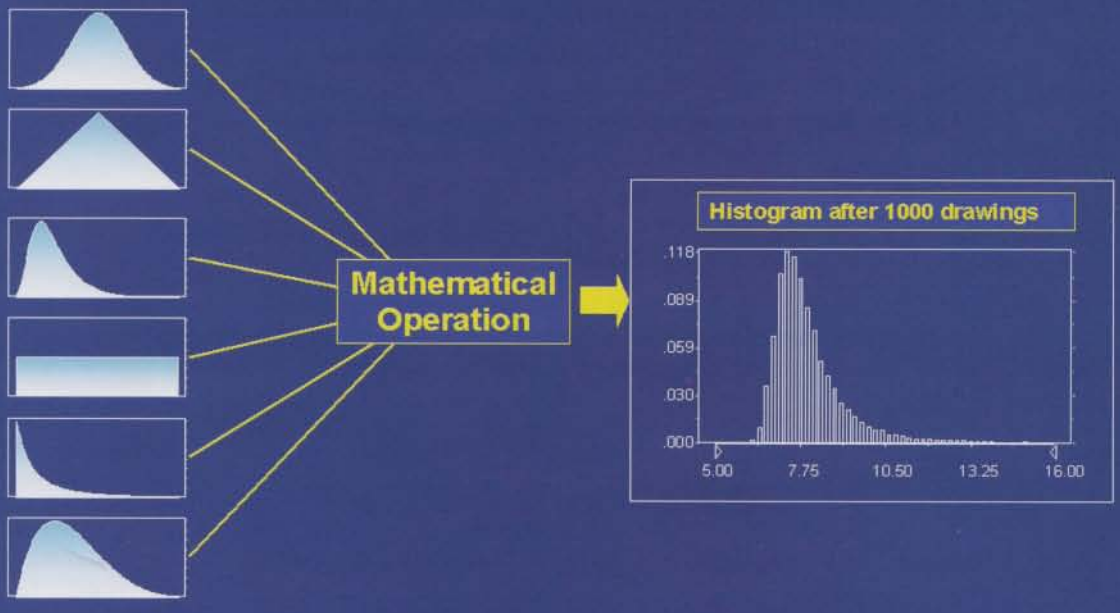
BASICS OF MONTE-CARLO SIMULATION (1)



SEG/EAGE DISC 2003

P. Delfiner 4-3

BASICS OF MONTE-CARLO SIMULATION (2)

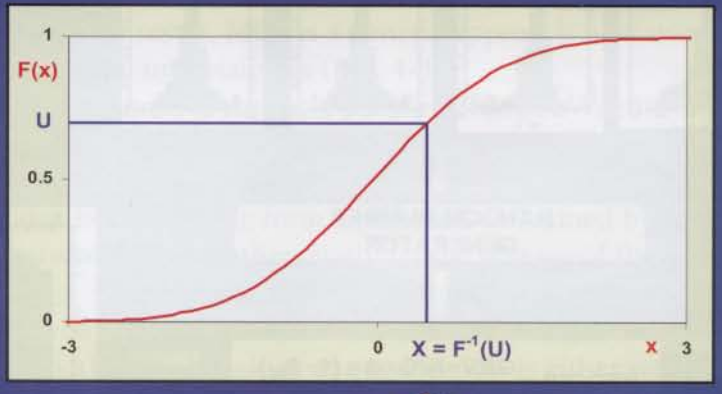


SEG/EAGE DISC 2003

4-4

SIMULATION OF A CONTINUOUS RANDOM VARIABLE WITH A CDF F(x)

Draw a uniform number U between zero and one



Find the value X associated with U using the formula $X=F^{-1}(U)$

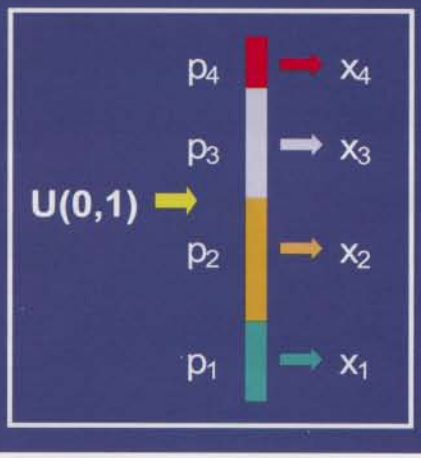
SEG/EAGE DISC 2003

P. Delfiner 4-5

SIMULATION OF A DISCRETE RANDOM VARIABLE

Assumption X takes n discrete values $x_1, x_2, \ldots, x_N$ with probabilities $p_1, p_2, \ldots, p_N$.

Method On the Y axis define segments of length $p_1, p_2, \ldots, p_N$. Simulate a uniform [0,1] and select the value corresponding to the segment in which the random point falls.

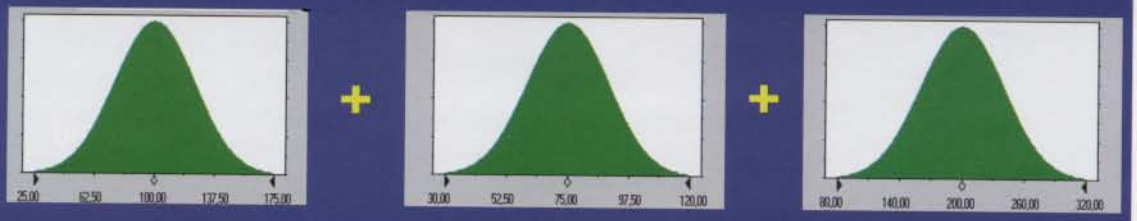


SEG/EAGE DISC 2003 P. Delfiner 4-6

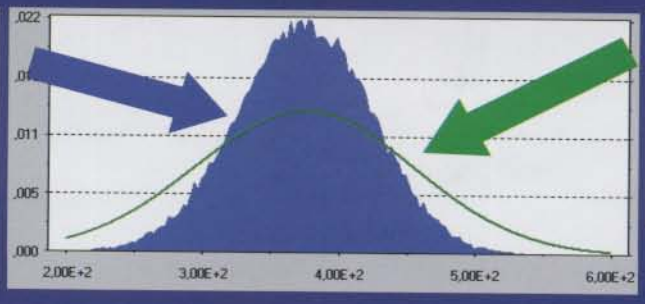
The most difficult issue in MCS is that of correlations between parameters. Take the example of three prospects located in the same play, each with a normal pdf for their STOOIP (Fig. 4-7). What will be the pdf of the complete volume, assuming that the three prospects are discoveries? In probability theory, one cannot avoid making correlation assumptions when combining several pdfs. These assumptions go from one extreme (pure independence between pdfs) to another (complete dependence). In the first instance, the three pdfs tend to compensate each other. Independence implies that there is a very small probability that the three volumes will be all low or all high. On the other hand, complete dependence implies that volumes are either all low, all medium, or all high. Thus, the range of uncertainty over the pdf of the sum is much larger in the complete dependence than in the complete independence situation. Note that the standard industry approach, which consists of adding the P90s and the P10s (see Fig. 1-25 for definition) of individual field reserves to get the P90 or the P10 of the company's total reserves, is equivalent to assuming complete dependence between all fields.

Fig. 4-8 is an interesting paradox. Assume we know nothing about two parameters. We only know that they are between 0 and 1, leading to the assumption that their value is uniformly distributed between 0 and 1. What about their sum? A casual run of a Monte-Carlo spreadsheet program leads to . . . a triangular distribution between 0 and 2, with a mode of 1! Why is the sum not distributed uniformly between 0 and 2? How is it possible that there is less uncertainty on the sum of two unknown parameters than on each of them individually? Again, the explanation lies in the parameter-correlation question, to which there cannot be an answer such as "I don't know." With Monte-

THE THREE PROSPECTS



Independence assumption: conclusion obtained by Monte-Carlo simulation (or by properly combining variances)

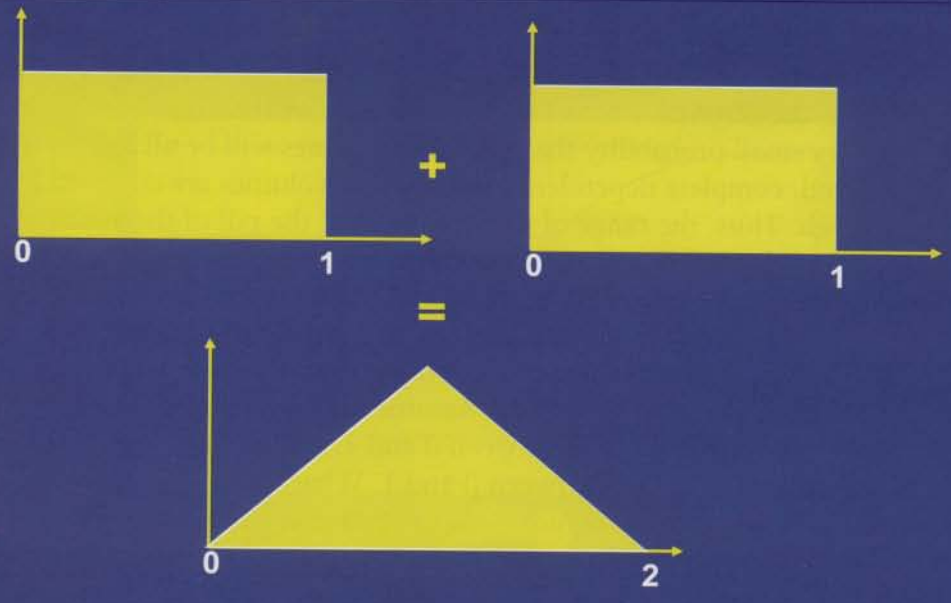


Full dependence assumption: conclusion obtained by simply adding min and max of prospects

SEG/EAGE DISC 2003

4-7

WHEN 1+1 IS NOT EQUAL TO 2



An independence hypothesis is a very strong hypothesis!!!

SEG/EAGE DISC 2003

4-8

Carlo software, not entering any correlation information in the input is usually interpreted as an independence assumption. As a result, there is a tendency to mutual compensation between the two parameters, leading to a sum that is closer to the middle than to the edge of the [0,2] interval.

There are different approaches for generating samples of parameters that are correlated with each other. The first approach consists of working sequentially. For instance, net/gross is sampled first, then porosity is sampled as a function of this net/gross, then saturation itself is controlled by porosity and net/gross. Another example is porosity and permeability. Porosity is sampled first using a Gaussian distribution, then the logarithm of permeability is derived from porosity using the regression equation associated with the permeability-porosity calibration. The reliability of the regression equation can be accounted for by adding a random term to the prediction of the logarithm of permeability from porosity.

The second approach consists of making an assumption — usually that it is normal — about the multivariate distribution of the correlated parameters to be sampled. For instance, it is assumed that the log of permeability and the log of porosity follow a binormal distribution with a correlation coefficient equal to ρ. There are many ways to generate samples from multivariate normal distributions; the approaches are simple for bivariate distributions (Johnson, 1987). When there are more than two variables, the general approach is discussed in Fig. 4-9. The trick consists of reducing the problem to that of generating uncorrelated univariate samples. The first approach consists of writing the sampled multivariate random variable as a linear combination of independent univariate samples, through Cholesky factorization. The other approach consists of working sequentially and using the property that, in the normal multivariate case, univariate conditional distributions are also Gaussian. These two approaches will also be used in the algorithms for the conditional simulation of continuous parameters. The use of multivariate normal distributions may not prove sufficient in situations where some of the variables are not normally distributed. This is why some MCS commercial programs prefer to use rank correlations, which are independent of the pdf of individual variables.

4.3 Conditional Simulations for Continuous Parameters

4.3.1 Example

MCS is a technique that takes the pdf of individual random variables as input and derives as output the pdf of any function of these variables (Fig. 4-4). Nowadays, parameters such as porosity, lithology, or net/gross are mapped in 2D or 3D. Because of the uncertainty affecting each of these models, we may wish to quantify the joint impact of these uncertainties on the STOOIP or on other results. We may also want to know what the heterogeneities actually look like away from the wells, in order to input a realistic model into flow simulation. GCS is the method that will allow us to generalize MCS to 2D and 3D models.

Fig. 4-10 is an example of acoustic-impedance kriging, as interpolated from six wells in a North Sea field. Away from the wells, kriging is very smooth, because there is no well information available, and the best estimate is the mean. However, this is not

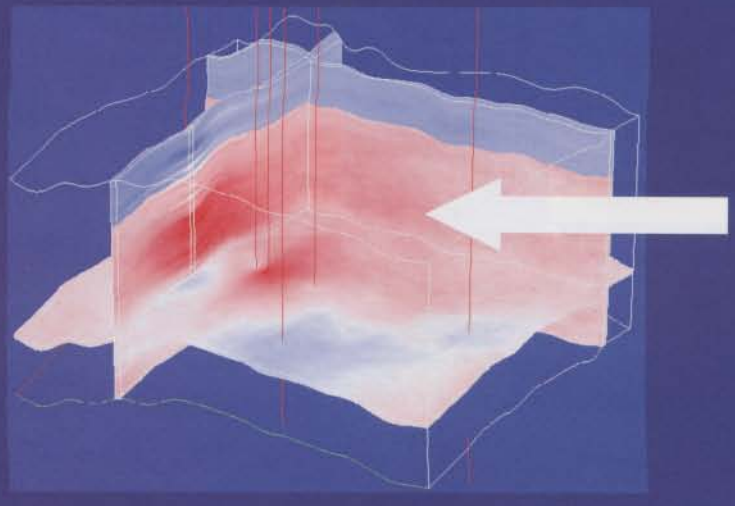
GENERATE A SAMPLE FROM A MULTIVARIATE NORMAL DISTRIBUTION (JOHNSON, 1987)

$$f(\mathbf{z}) = (2\pi)^{-N/2}|\Sigma|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{z}-\mathbf{m})'\Sigma^{-1}(\mathbf{z}-\mathbf{m})\right]$$

General approach: use a trick to only have to generate one sample of one univariate distribution at a time. Two techniques for this:

- Write $\mathbf{Z} = \mathbf{A}\cdot\mathbf{Y} + \mathbf{m}$ where **Y** is a vector of N independent normal variables.
- Generate Z_N, then Z_{N-1} / Z_N, then Z_{N-2} / Z_{N-1} ,Z_N ……
  Each conditional distribution is Gaussian with mean and variance equal to the mean and variance derived from the regression against previously simulated values

A KRIGING EXAMPLE IN 3D (LAMY ET AL., 1998b)

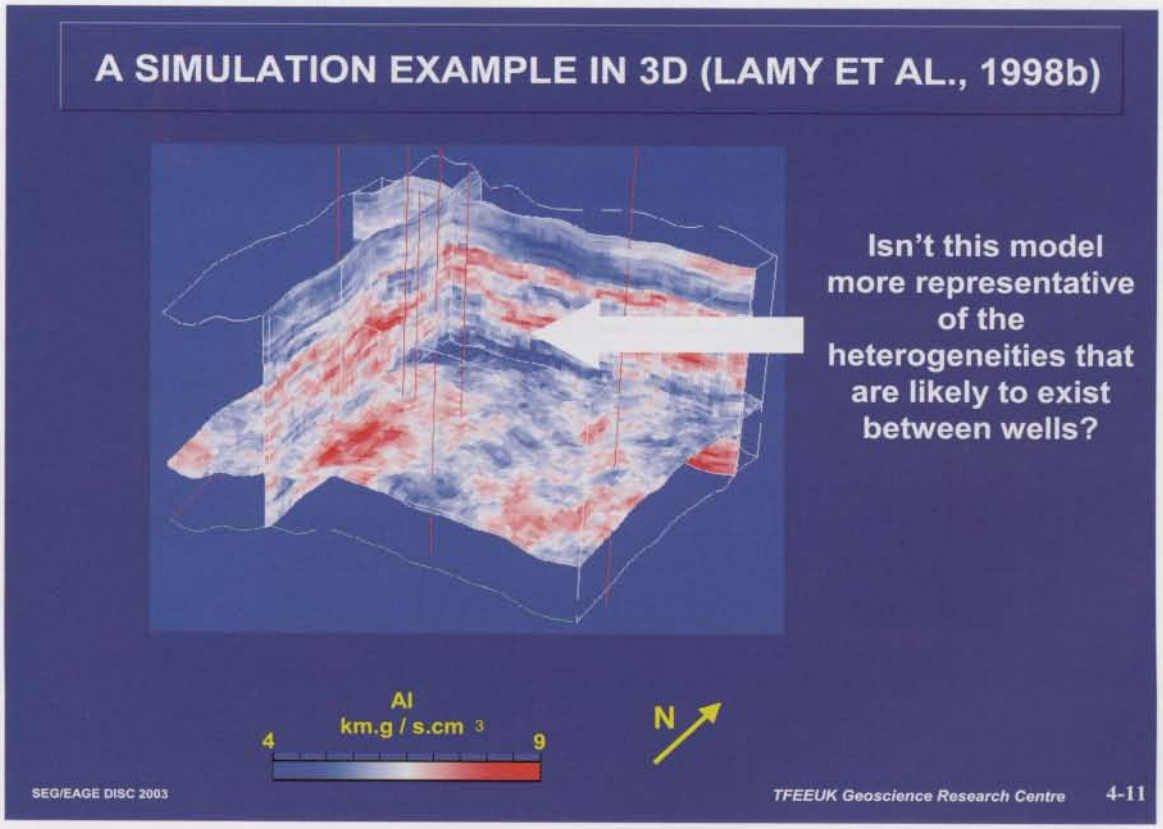


Why should the model be smooth precisely away from the data points?

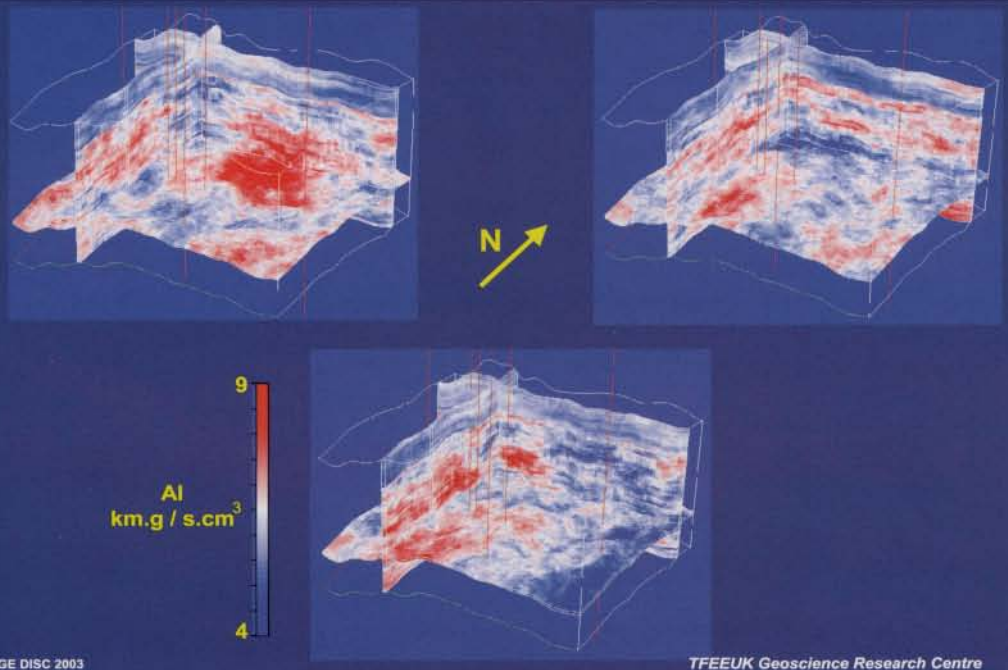
representative of the degree of variability of acoustic impedance away from the wells. Fig. 4-11 shows a realization of GCS, using exactly the same input parameters as the kriged model of Fig. 4-10. This realization is not optimal in the kriging sense, but it has the property of satisfying the variogram and also some a priori statistics on the mean and standard deviation, while still honoring the well data. Because these conditions are not sufficient to entirely constrain the model, a large number of 3D realizations can be generated that all satisfy these constraints (Fig. 4-12). Each GCS realization is representative of the heterogeneities that are likely to be encountered in the reservoir (Fig. 4-13).

A fundamental difference between kriging and a GCS realization is that, although both honor the wells, kriging becomes smoother away from the wells, whereas the GCS realization looks the same everywhere. If the well locations are not displayed on the model, it should be impossible to guess where there are on the basis of a GCS realization. On the other hand, their location is usually easy to spot on a kriged map. Fig. 4-14 illustrates this point on three 2D GCS realizations obtained from 70 wells. None of the three realizations show "bulls-eyes" at the well locations. The three intermediate nonconditional simulations will be discussed later.

What is the relationship between kriging and GCS realizations? As the number of realizations becomes very large — say, a few hundred — their mean becomes closer and closer to kriging. Why is this? Because all realizations honor the wells, they tend to be very similar to each other close to the wells. On the other hand, they are different away from the wells, and they become uncorrelated as soon as the distance from the wells is



GENERATION OF MULTIPLE REALIZATIONS (LAMY ET AL., 1998b)



SEG/EAGE DISC 2003 TFEEUK Geoscience Research Centre 4-12

TWO REASONABLE BUT INCOMPATIBLE OBJECTIVES

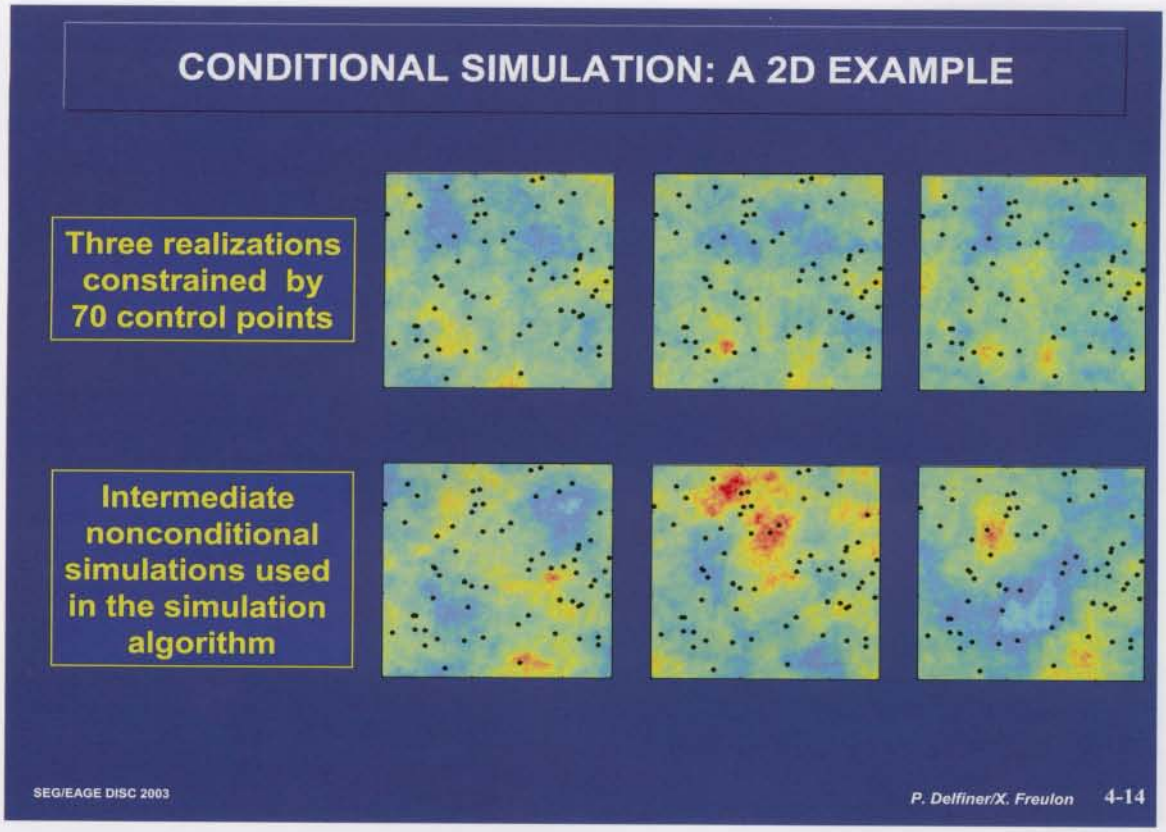
1. Minimize the interpolation error between wells

→ *KRIGING*

2. Give a representative picture of the heterogeneities between wells.

→ *CONDITIONAL SIMULATION*

SEG/EAGE DISC 2003 4-13

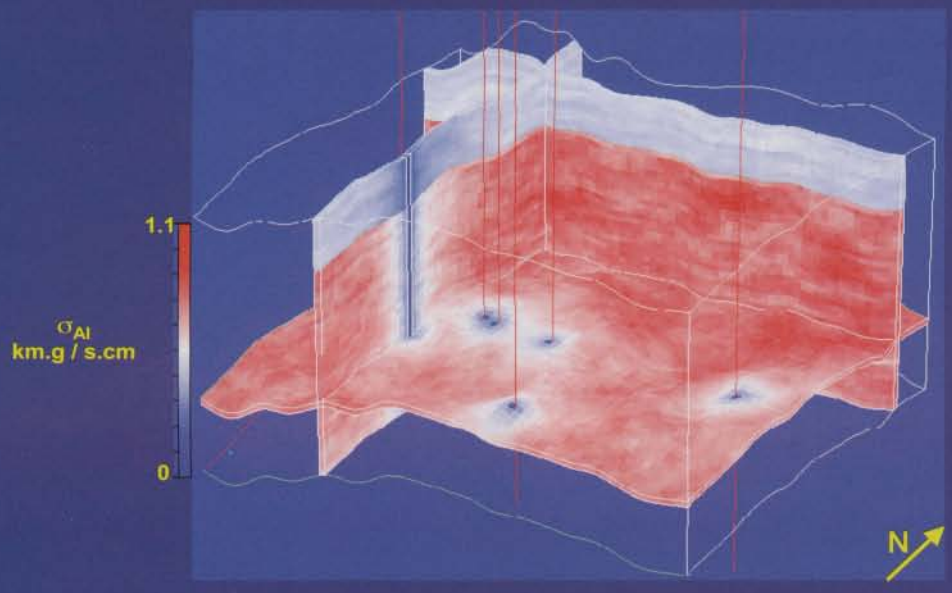


greater than the range of the variogram. This results in a significant compensation effect from one realization to another. Consequently, the mean of the realizations tends to be very smooth. If we now calculate at each location the standard deviation of the realizations, we find the kriging standard deviation, which is equal to zero at the wells because each realization honors these data (Fig. 4-15). Obviously, the GCS realizations also satisfy the spectral density associated with the input covariance or variogram, which was not true for kriging. Fig. 4-16 compares the properties of kriging and GCS.

Figs. 4-17 and 4-18 illustrate again the difference between GCS and kriging. We know that the larger the variogram range, the better the correlation between values away from each other. As a result, kriging remains controlled by the well data, as long as the distance from the wells is smaller than the variogram range (Fig. 4-17). GCS realizations look rather smooth, precisely because the variogram model has a large range. On the other hand, we know that kriging comes back rapidly toward the mean with short-range variograms (Fig. 4-18), whereas GCS realizations look quite noisy. In other words, the more randomness there is in the variable, the noisier the simulations . . . and the smoother kriging!

The assumption made by GCS is not that Nature is random, but that our knowledge of the reservoir is not sufficient to generate a single (deterministic) model of heterogeneities between wells. There are a number of possible scenarios given the data and the a priori variogram and trend model, and the probabilistic approach is very convenient for generating these scenarios (Fig. 4-19). Because of their great variability away from the wells, realizations are more realistic than most deterministic models, which are

STANDARD DEVIATION OF REALIZATIONS (LAMY ET AL., 1998b)



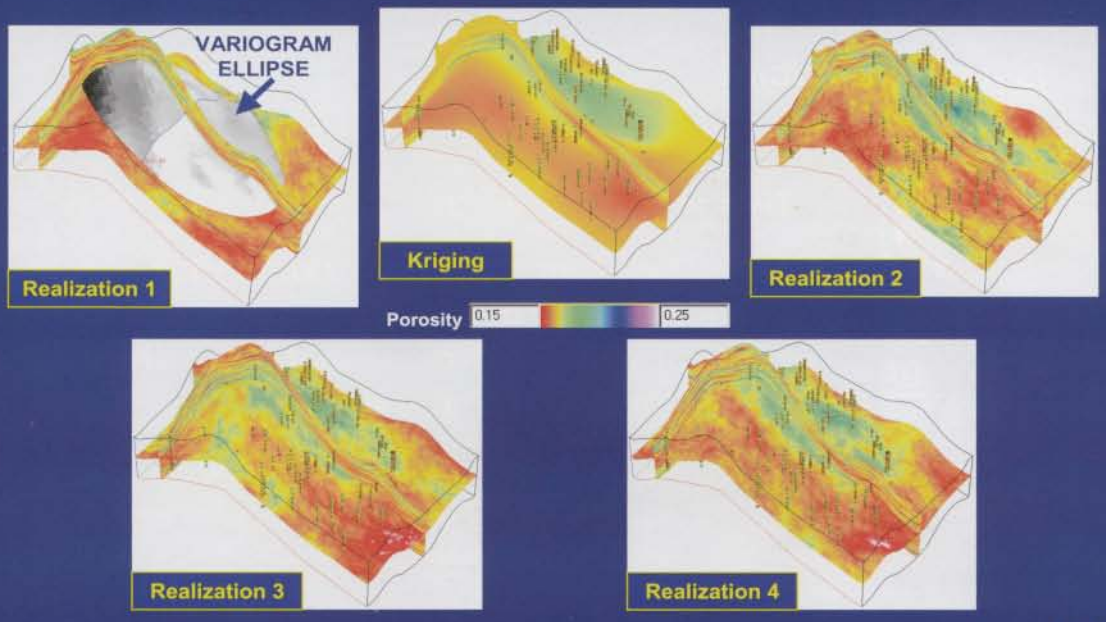
SEG/EAGE DISC 2003 TFEEUK Geoscience Research Centre 4-15

KRIGING OR CONDITIONAL SIMULATION?

	Conditional simulation	Kriging
Output	Multiple realizations.	One "deterministic" model.
Properties	Honors wells, honors histogram, variogram, spectral density.	Honors wells, minimizes error variance.
Image	Noisy, especially if variogram model is noisy.	Smooth, especially if variogram model is noisy.
Data points	Image has same variability everywhere. Data location cannot be guessed from image.	Tendency to come back to trend away from data. Data location can be spotted.

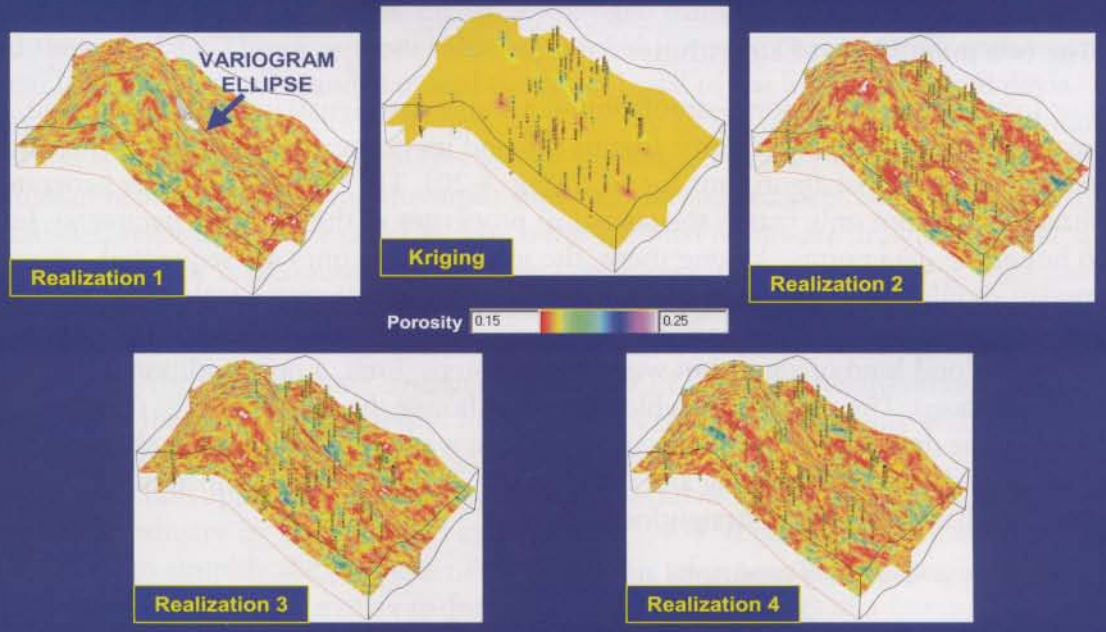
SEG/EAGE DISC 2003 4-16

COMPARING KRIGING AND CONDITIONAL SIMULATION
(VERTICAL & HORIZ. SPHERICAL VARIOGRAM RANGE =40% OF MODEL SIZE)

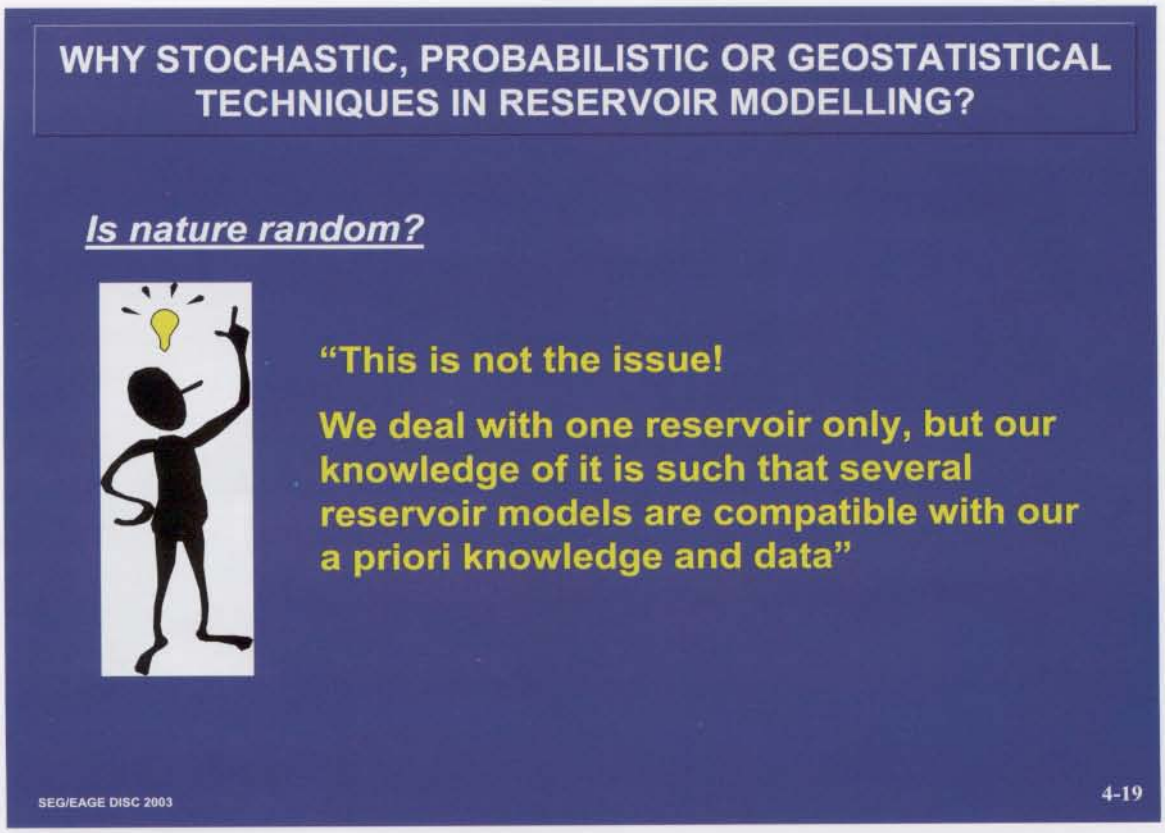


SEG/EAGE DISC 2003 Ph. Lamy 4-17

COMPARING KRIGING AND CONDITIONAL SIMULATION
(VERTICAL & HORIZ. SPHERICAL VARIOGRAM RANGE =5% OF MODEL SIZE)



SEG/EAGE DISC 2003 Ph. Lamy 4-18



usually far too smooth. We will see later how GCS is now routinely used to generate representative 3D heterogeneity models or to quantify the uncertainty associated with 3D models.

4.3.2 Algorithms

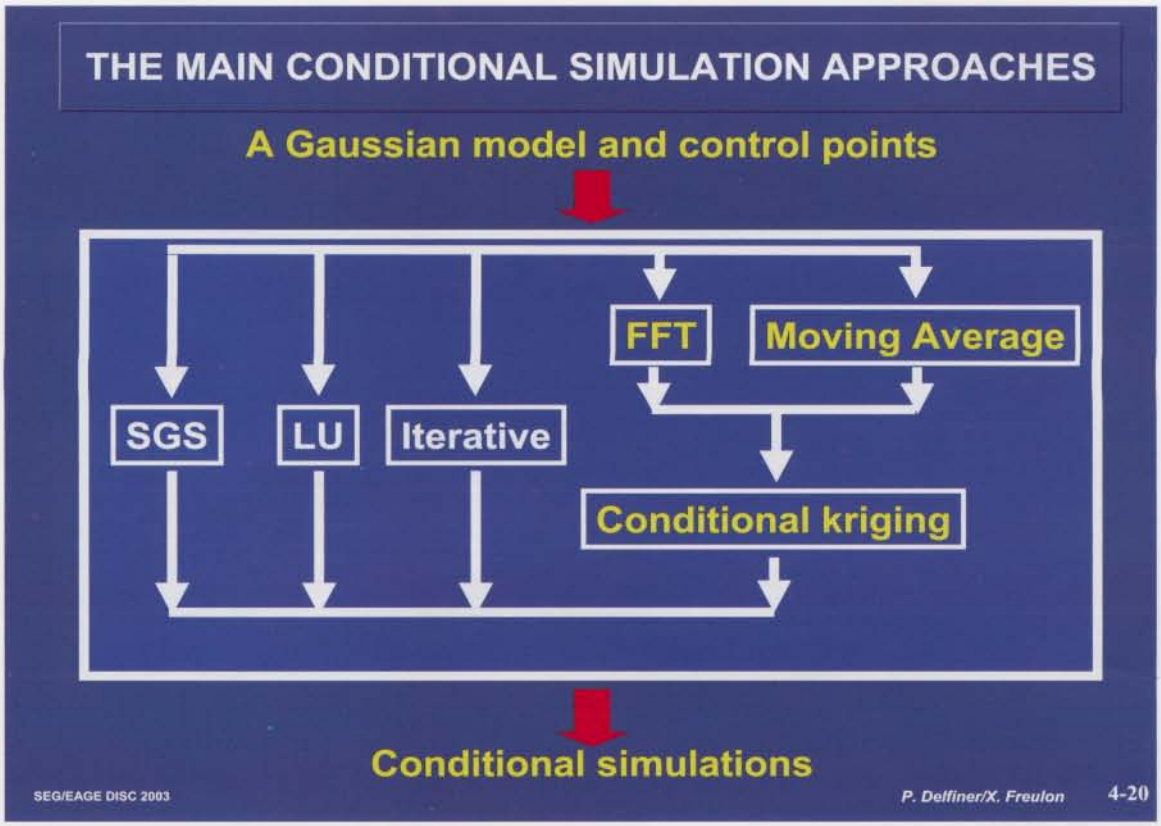
- The two main kinds of algorithms

Let us focus on the conditional simulation of a stationary Gaussian random function measured only at wells. We assume that its pdf and its variogram are known. There are two main categories of algorithms for GCS (Fig. 4-20). The first kind directly generates realizations that not only match the statistical properties of the modeled parameter, but also honor the data points. Among those, the most popular ones are sequential Gaussian simulation (SGS), LU matrix decomposition, and iterative techniques in a general sense.

The second kind of algorithm works in two steps. First, a nonconditional simulation is generated. The realization, which does not honor the data points, is then "conditioned" to the data points. In the discussion below, we have selected the techniques that appear to be the most interesting and widely used. They will be described in 2D, but their generalization to 3D is straightforward.

- *Direct generation of conditional simulations*

SGS (Deutsch and Journel, 1992) is probably the most popular and flexible tech-



nique today. Assume, for simplification, that a conditional simulation is generated on a 2D grid, where, say, three wells are available. The SGS approach is inspired directly from an already discussed result of the statistical literature (Fig. 4-9). The algorithm works sequentially, successively drawing and filling random locations of the plane (Figs. 4-21 and 4-22). At each new location, the value is kriged from the previously sampled values and the well data. Then a random value is sampled from the Gaussian pdf, with mean equal to the kriged value and standard deviation equal to the kriging standard deviation. Then the sampled value is merged with the rest of the data set, and a new random location is chosen. The data locations are left unchanged in the process. The result is a Gaussian realization satisfying the input statistics (mean, variance, and variogram). This is satisfied regardless of the order in which the sampled points are drawn. However, a random sequence is recommended to avoid spreading random artifacts. Fig. 4-23 provides examples of anisotropic SGS realizations on a North Sea field. The weakness of SGS is that it requires solving a kriging system at each location, which may prove time-consuming. Also, each kriging system is calculated in a moving neighborhood, which implies that the covariance function model is properly represented only within the distance of the moving neighborhood.

The second approach for direct conditional simulation was introduced to the petroleum industry by Davis (1987) and Alabert (1987). It is another way, as described in Fig. 4-9, to simplify the problem of GCS into that of generating independent random variables. The covariance matrix is decomposed into the product of a lower triangular matrix and its transpose. The conditioning by the data points is easily introduced as a

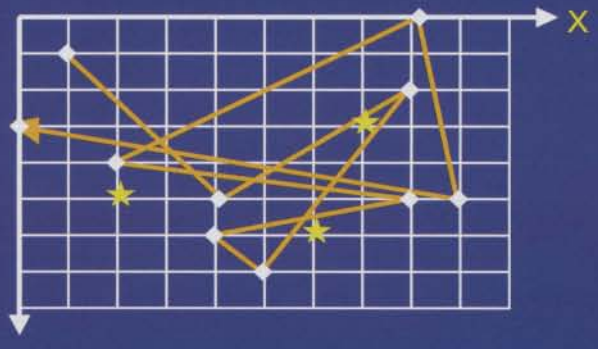
SEQUENTIAL GAUSSIAN SIMULATION

- Start with data at wells
- Random path through all locations

★ Wells

— Random path

◆ Location to be simulated

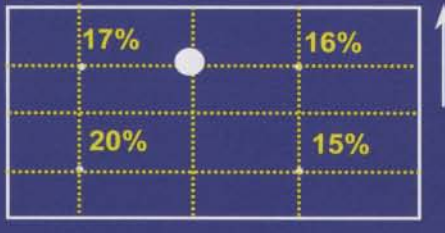


SEG/EAGE DISC 2003

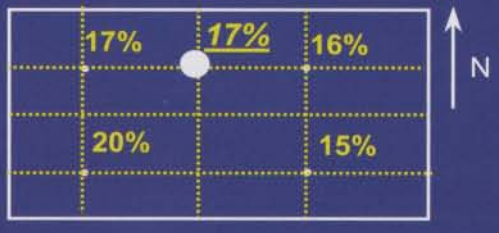
4-21

SEQUENTIAL GAUSSIAN SIMULATION (SGS) ALGORITHM

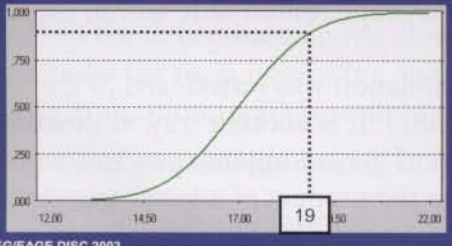
1. Draw random location between wells



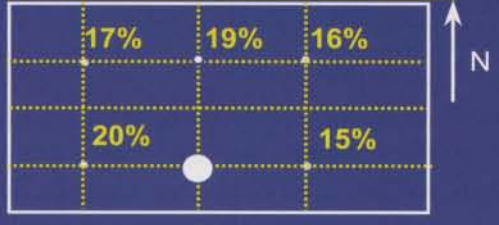
2. Estimate kriged value & std dev



3. Sample from Gaussian histogram

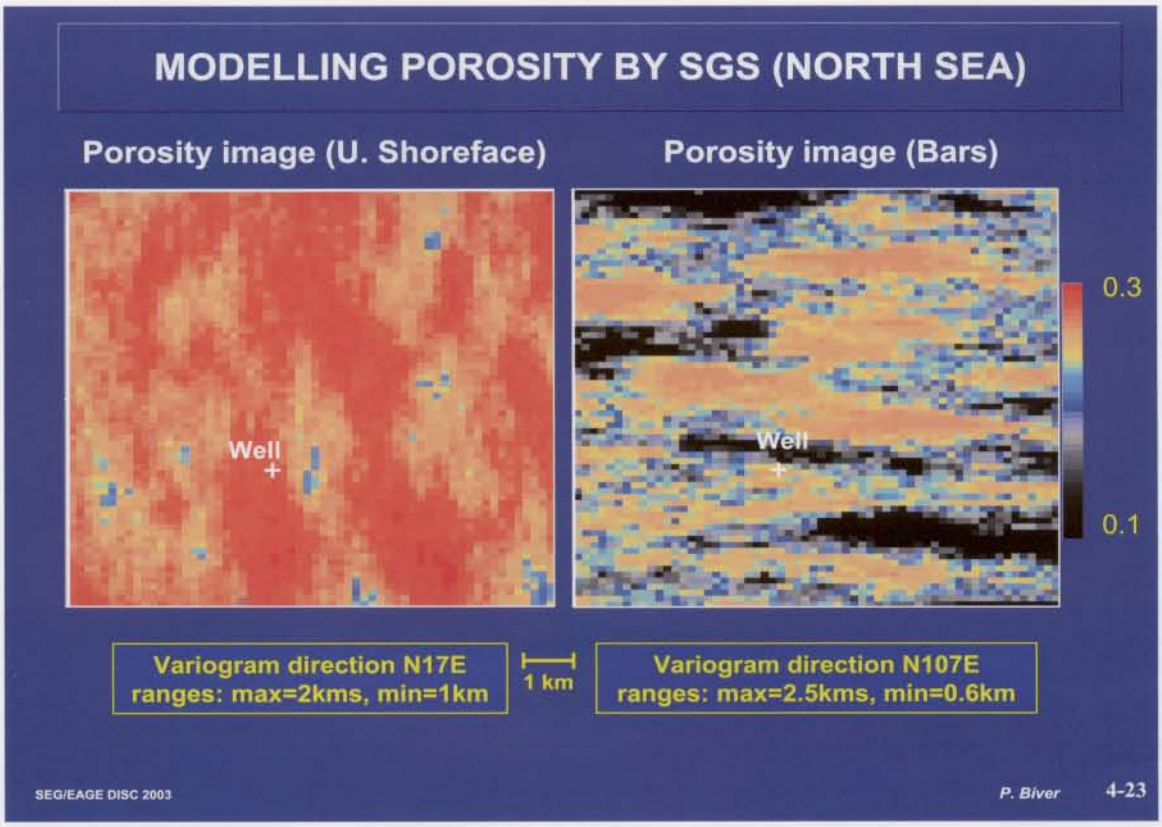


4. Merge sampled value and draw new location



SEG/EAGE DISC 2003

4-22



further constraint in the algorithm. The disadvantage of the method is that, if the simulation is calculated at 10,000 grid nodes, it requires the decomposition of the covariance matrix, whose dimension is $10{,}000^2 \times 10{,}000^2$.

The third approach is that of iterative techniques, which belong to the general class of Markov-Chain Monte-Carlo (MCMC) simulation (Gamerman, 1997). The Norwegian School (Hegstad et al., 1994; Omre and Tjelmeland, 1997) is very active and successful at developing and applying MCMC algorithms for sampling multivariate realizations constrained by wells, seismic, and production data. Their approach is Bayesian and consists of calculating the posterior multivariate pdf of the variables of interest and then sampling this pdf using the MCMC Metropolis-Hastings algorithm. Fig. 4-24 describes the Metropolis algorithm, which is a simplified version of Metropolis-Hastings, applying in situations where iterative perturbations are symmetrical. The algorithm starts from an initial random realization and converges toward the desired one thanks to a proper choice of iterative perturbations. The choice of the perturbations to apply is quite general and is not discussed here.

The Gibbs Sampler is the second prominent class of MCMC algorithms (Fig. 4-25). It was originally proposed by Geman and Geman (1984) and has found applications in the fractal modeling of natural terrain. Arakawa and Krotkov (1996) used the results of the work of Szeliski and Terzopoulos (1989) to generate fractal terrain models.

The simulated annealing algorithm used by Deutsch (1992) for generating geostatistical realizations constrained by well-test data also belongs to the class of MCMC techniques. Deutsch et al. (1996) generalize the approach to the incorporation of seis-

MARKOV-CHAIN MONTE CARLO (MCMC) (1) (GAMERMAN, 1997)

METROPOLIS ALGORITHM
Goal: sample from posterior $p(z_1, z_2, \ldots\ldots .., z_n)$

Start with initial $[z_1(0), z_2(0) \ldots\ldots .., z_n(0)]$

At each iteration t:

1. **apply perturbation to values $z_1(t-1), z_2(t-1), z_3(t-1), .. , z_n(t-1)$.**
2. **accept result of perturbation with probability:**

$$\min\left[1, \frac{p(\text{after perturbation})}{p(\text{before perturbation})}\right]$$

SEG/EAGE DISC 2003 4-24

mic data. Rather than formulate the problem in the Bayesian framework, where the prior distribution (geostatistical model) is clearly differentiated from the likelihood (fit to the seismic and production data), Deutsch et al. incorporate all the constraints (geostatistical constraints, fit to well data, well-test data, and seismic data) into a single global objective function.

MCMC algorithms are very time-consuming. However, they are rather straightforward to program, and they are able to address very general conditional-simulation problems that incorporate a great variety of data. They should be used only if noniterative approaches cannot address the problem. MCMC techniques give samples from the specified posterior distribution in the limit only. One method for deciding whether the algorithm is close to the limit is output analysis (Ripley, 1981). MCMC applications go far beyond geostatistical or petroleum applications. In the geophysical literature, the works of Mosegaard and Tarantola (1995) or Sen and Stoffa (1996) are important references.

- *Two-step approaches*

Two-step approaches start with generating realizations of nonconditional simulations. A nonconditional simulation matches the required statistics, but does not honor the data values.

A very popular method for generating nonconditional simulations is the Fourier integral method (Fig. 4-26). Yao (1998) provides numerous references to this very popular approach. She also proposes to develop this algorithm into a conditional one by adding an iterative identification of phases to honor data values at sample locations.

MARKOV-CHAIN MONTE CARLO (MCMC) (2) (GAMERMAN, 1997)

GIBBS SAMPLING
Goal: sample from posterior $p(z_1, z_2, ..., z_n)$

Consider also that the conditional distributions - called π - are available

Start with initial $[z_1(0), z_2(0) ....., z_n(0)]$

At each iteration t:

$z_1(t)$ drawn from $\pi[z_1/z_2(t-1), z_3(t-1) ....., z_n(t-1)]$

$z_2(t)$ drawn from $\pi[z_2/z_1(t), z_2(t-1), z_3(t-1) ....., z_n(t-1)]$

.....

$z_n(t)$ drawn from $\pi[z_n/z_1(t), z_2(t), z_3(t) ....., z_n(t-1)]$

THE FOURIER INTEGRAL METHOD FOR GENERATING NONCONDITIONAL SIMULATIONS (YAO,1998)

1. **Sample the covariance model to obtain the sequence of C(k), k=0, 1,..,K-1**
2. **Compute density spectrum by discrete Fourier transform of C(k):**

$$s(j) = \frac{1}{K} \sum_{k=0}^{K-1} C(k) e^{-2\pi kj/K}$$

3. **Derive amplitude spectrum from density spectrum by:**

$$|A(j)| = \sqrt{s(j)}, \; j = 0,..., K-1$$

4. **Generate random phase sequence $\phi(j)$ from uniform distribution in $[0,2\pi]$**
5. **Calculate Fourier coefficients:**

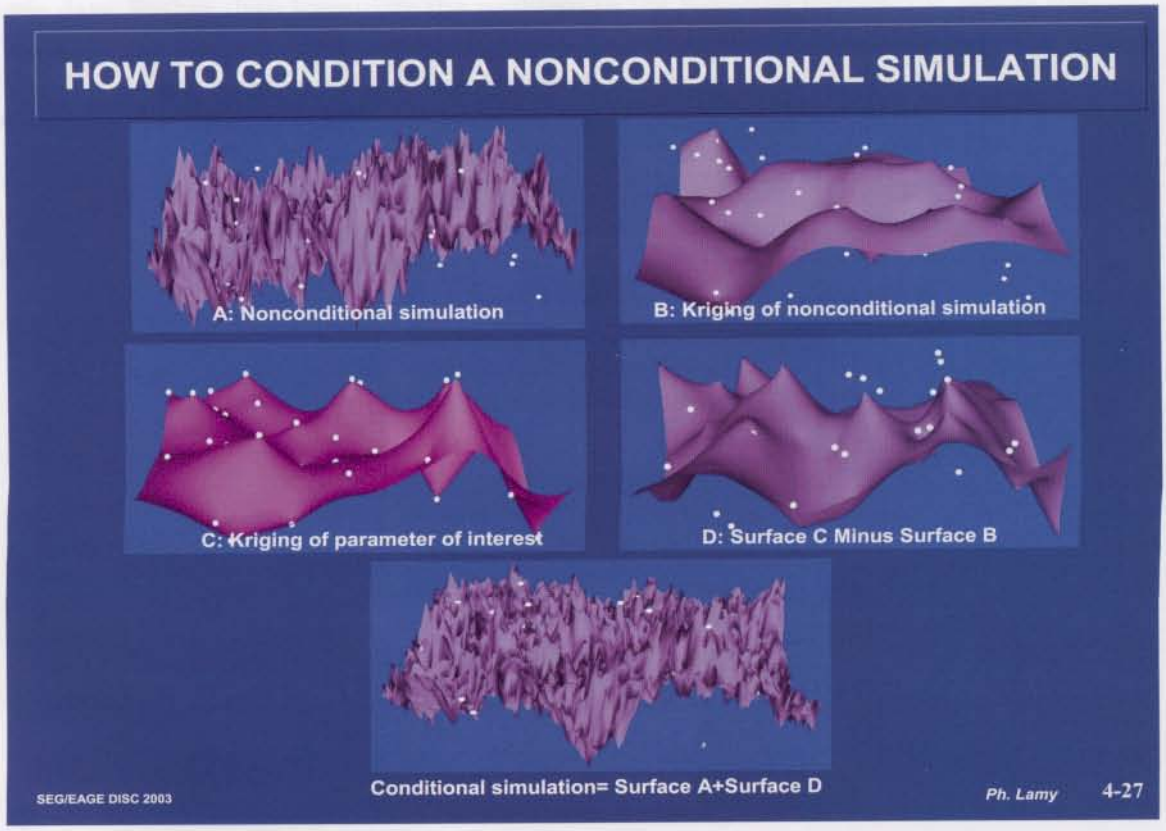
$$A(j) = |A(j)| e^{-i\phi(j)} \; j = 0,..., K-1$$

6. **Apply inverse discrete Fourier transform to complex coefficients A(j):**

$$Z(k) = \sum_{k=0}^{K-1} A(j) e^{2\pi kj/K}$$

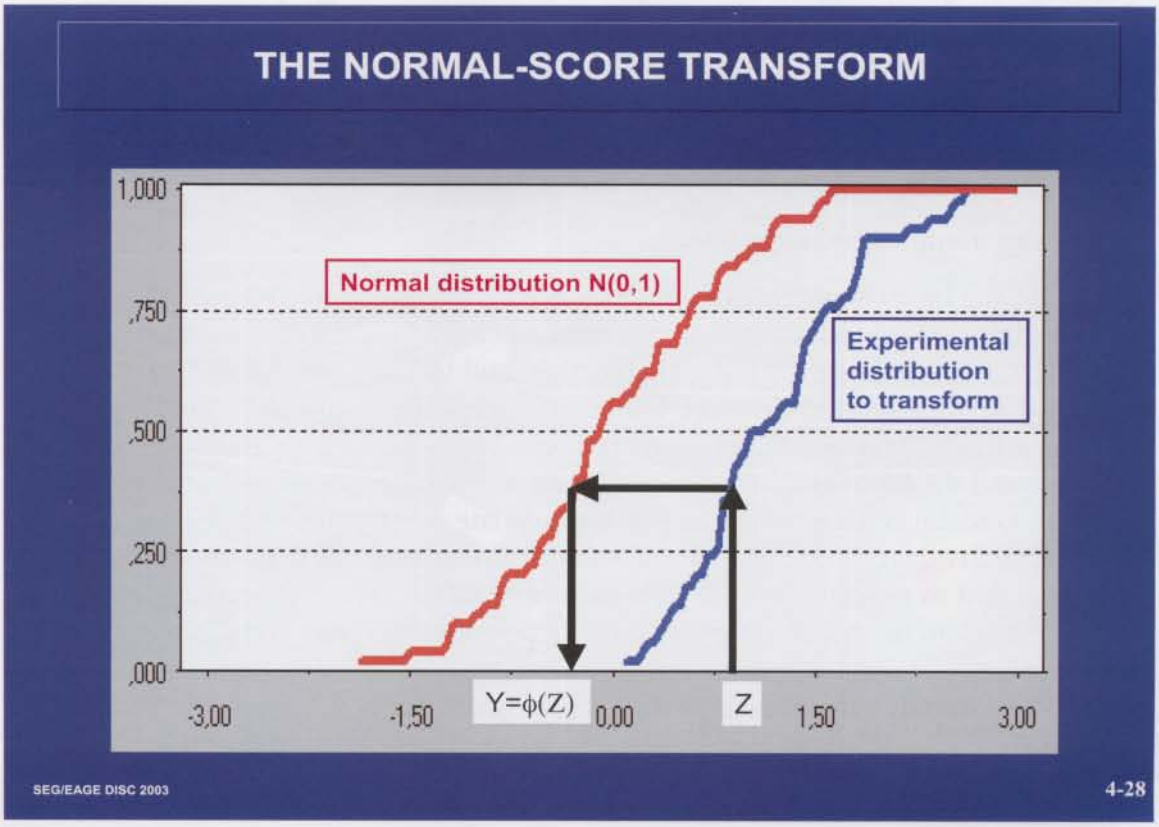
Another popular method, which Doyen (1988) used in his already-referenced paper, consists of convolving a white noise realization by a weighting function whose autocorrelation function is equal to the desired covariance. Exercise 3 illustrates how a simple averaging transforms the realization of an uncorrelated parameter into the realization of a correlated one. Oliver (1995) calls this the moving-average technique and generalizes it in 2D and 3D. For each covariance model to be simulated, Oliver calculates what he calls the "square root of the covariance operator." This is obtained by calculating the Fourier transform of the covariance function, then taking the square root of the result and deriving its inverse Fourier transform. In his paper, Oliver provides the analytical expression of the square root of commonly used covariance functions. The method is quite elegant but may prove time-consuming when covariances with large ranges are simulated.

How do we obtain a conditional simulation from a nonconditional one? The approach is widely used and explained by Chilès and Delfiner (1999). Fig. 4-27 summarizes its main steps, which can be interpreted in two different ways. According to one interpretation, a nonconditional simulation is generated first. Then a kriging-based smooth correction function is calculated, which corrects the values at the wells in order to make them honor the data. This is the approach described in Fig. 4-27. There is another way to look at it. First, the kriged surface interpolating the variable of interest is calculated. Then realizations of the kriging error are simulated and added to the kriging error. The result of each addition is a GCS realization. Fig. 4-14 shows examples of three nonconditional simulations and three conditional simulations from a case study.



- Nonstationary and non-Gaussian variables

By definition, SGS assumes that the simulated variable is Gaussian. The other methods — except for MCMC, which is very general — build the conditional simulations by multiple additions of elementary random variables. Because of the central limit theorem (Fig. 1-29), the resulting realizations will tend to be distributed normally. If we wish to simulate a non-Gaussian variable, two approaches are possible. The first one consists of initially transforming the data pdf into a normal distribution. This can be done using the normal-score transform. This simple transformation, described in Fig. 4-28, assigns to each value of the original variable another value such that the distribution of the transformed data is normal. After generating the simulation, a back-transformation must be applied (Deutsch and Journel, 1992). This approach, although more rigorous, also has a number of drawbacks. It is nonlinear, and it becomes tedious to use when a variety of data (seismic data, production data, etc.) are combined in the simulation process (Tran et al., 2001). Quite often the second approach is preferred, especially if the data histogram — as is usually the case with porosity data — is not too far from Gaussian. This approach consists of applying "Direct Simulation," that is, simulation on the untransformed data. In most situations, the variogram will be correctly reproduced. This is true, for instance, for SGS, as long as each simulated value is drawn from a local distribution whose mean and variance are obtained by simple kriging (Journel, 1993). The histogram will not be properly reproduced, but a number of techniques have been proposed to reproduce this histogram (Tran et al., 2001).



The generalization to nonstationary variables is simple, because the trend is a deterministic function. Only the residual from the trend changes from one realization to another. This also applies to GCS with an external drift. Only in the case of Bayesian kriging, where the coefficients of the trend are themselves random, does the trend model change with the realization (see, for instance, Abrahamsen et al., 1991).

4.4 Cosimulation

By "cosimulation," we mean all the situations in which the GCS of one parameter must incorporate information from another parameter, which provides extra information about the first one. In earth modeling applications, cosimulation can apply to a number of possible situations (Haas et al., 1998).

4.4.1 Collocated simulation with seismic data as a secondary variable

- Handling linear relationships

This is the situation that has already been encountered with collocated cokriging. A seismic attribute, known as the secondary variable, is statistically correlated with the parameter of interest, known as the primary variable. Information from the secondary variable must then be incorporated into the simulation of the primary variable, in order to keep the consistency between values of the primary and of the secondary variable. Usually, the collocated cokriging approach is applied jointly with SGS.

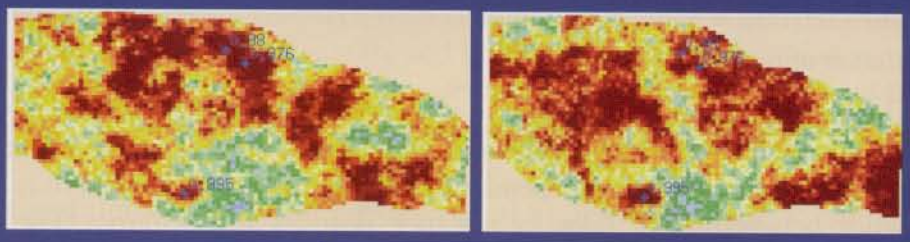
The SGS technique can be straightforwardly generalized to cosimulation (Gomez-Hernandez and Journel, 1993). The only difference with the single-parameter situation is that, instead of calculating the kriging estimate and standard deviation at each new sampled location (Fig. 4-22), this time the cokriging estimate and standard deviation are calculated. Fig. 4-29 shows an example of the simulation of net/gross constrained by a seismic attribute. Here, the seismic parameter is considered to be deterministic, or exact, because is does not change from one realization of the primary variable to another.

- Handling nonlinear relationships

Gastaldi et al. (1998) argue that linear relationships are too limited to predict reservoir thickness from seismic parameters under tuning conditions. In a case study in which 15 wells and a 3D seismic survey are available, they build a large number of "pseudo-wells" and derive a synthetic seismic from them. Modeling results show (Fig. 4-30) that reservoir thickness, the parameter of interest, can be predicted from the "reservoir isochron" and the "stratigraphic isochron" seismic attributes, using two nonlinear relationships. Gastaldi et al. assume that this multivariate relationship can be described by a joint probability distribution function. A multiparametric kernel density estimation method is used to estimate this pdf from the synthetic data set. Collocated cokriging, which is based on the correlation coefficient between the primary and the secondary variable, cannot be applied in this situation. However, the Bayesian formalism of Doyen et al. (1996), already presented in Section 3.6.4, can be applied, because the likelihood function is quite general and is not limited to linear relationships. Using SGS, Gastaldi et al. generate a large number of reservoir thickness realizations constrained by seismic

COLLOCATED COSIMULATION EXAMPLE

Seismic amplitude map



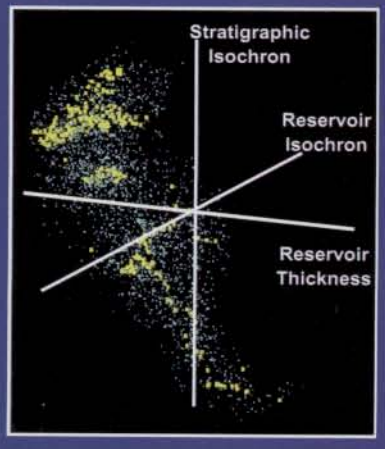
Net to Gross maps (2 realizations)

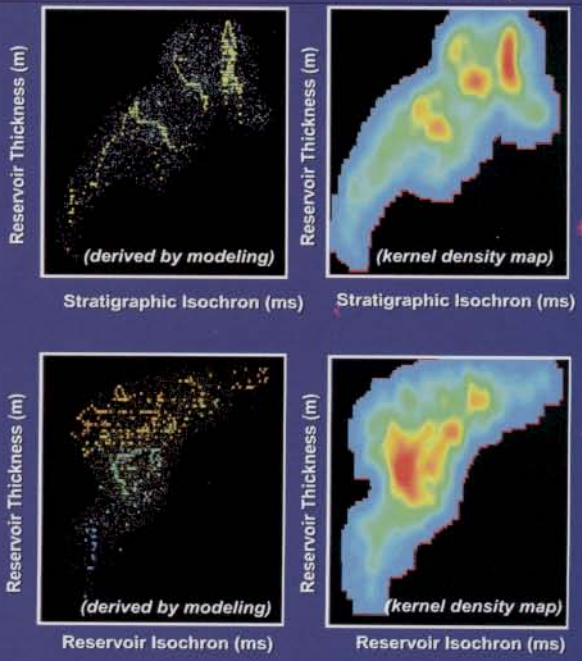
SEG/EAGE DISC 2003

P. Biver 4-29

FITTING A NONLINEAR RELATIONSHIP BETWEEN PRIMARY AND SEISMIC VARIABLE (GASTALDI ET AL., 1998)

KERNEL DENSITY ESTIMATION





SEG/EAGE DISC 2003

4-30

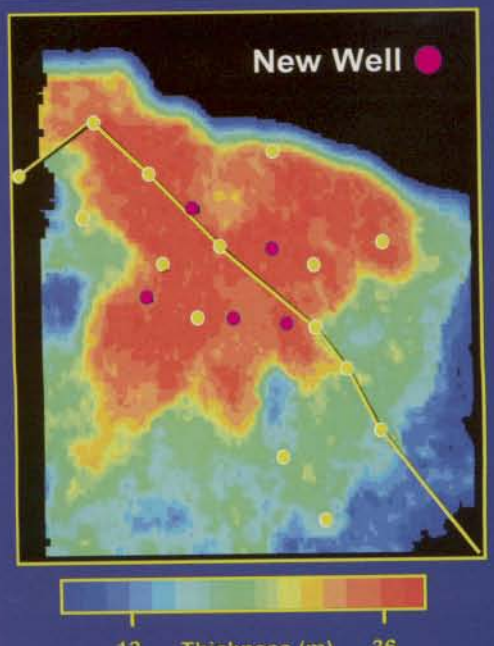
and derive the best geostatistical estimate of thickness — a kind of generalized nonlinear cokriging — by averaging the Bayesian simulations. In their example, estimation at five blind wells (Fig. 4-31) shows that the method gives better results than kriging or than standard "linear" collocated cokriging.

Before Gastaldi et al., Bashore et al. (1993) developed a somewhat similar technique, which they called the cloud transform, for handling nonlinear relationships. Based on the crossplot of porosity versus acoustic impedance, they build several porosity histograms associated with each class of impedance values. This is a similar philosophy to that of the kernel density estimation. Then, they generate the SGS conditional simulations by sampling a porosity value from the histogram associated with each seismic-impedance value. The approach they use to ensure lateral continuity between sampled porosity values is different from and somewhat less general than that of Gastaldi et al.

4.4.2 *Joint simulation of two parameters*

This is exactly the generalization of the problem already discussed with joint Monte-Carlo simulation. We need to generate conditional simulations of two parameters at the same time, but cannot generate them independently because we know that there is a degree of correlation between them. Take, for instance, porosity and net/gross (Fig. 4-32). Here, cosimulation appears to be the right approach. With each realization, two new maps of porosity and net/gross are generated. Both parameters are treated as stochastic. However, as explained by Gomez-Hernandez and Journel (1993), the sequential simulation algorithm can still be applied.

RESERVOIR THICKNESS PREDICTION AT 5 NEW WELLS, (GASTALDI ET AL., 1998)



	Absolute Error (m)	
	Average	Maximum
Kriging	3.7	7.1
Cokriging	2.5	3.8
Bayesian simulation	1.8	2.7

SEG/EAGE DISC 2003

4-31

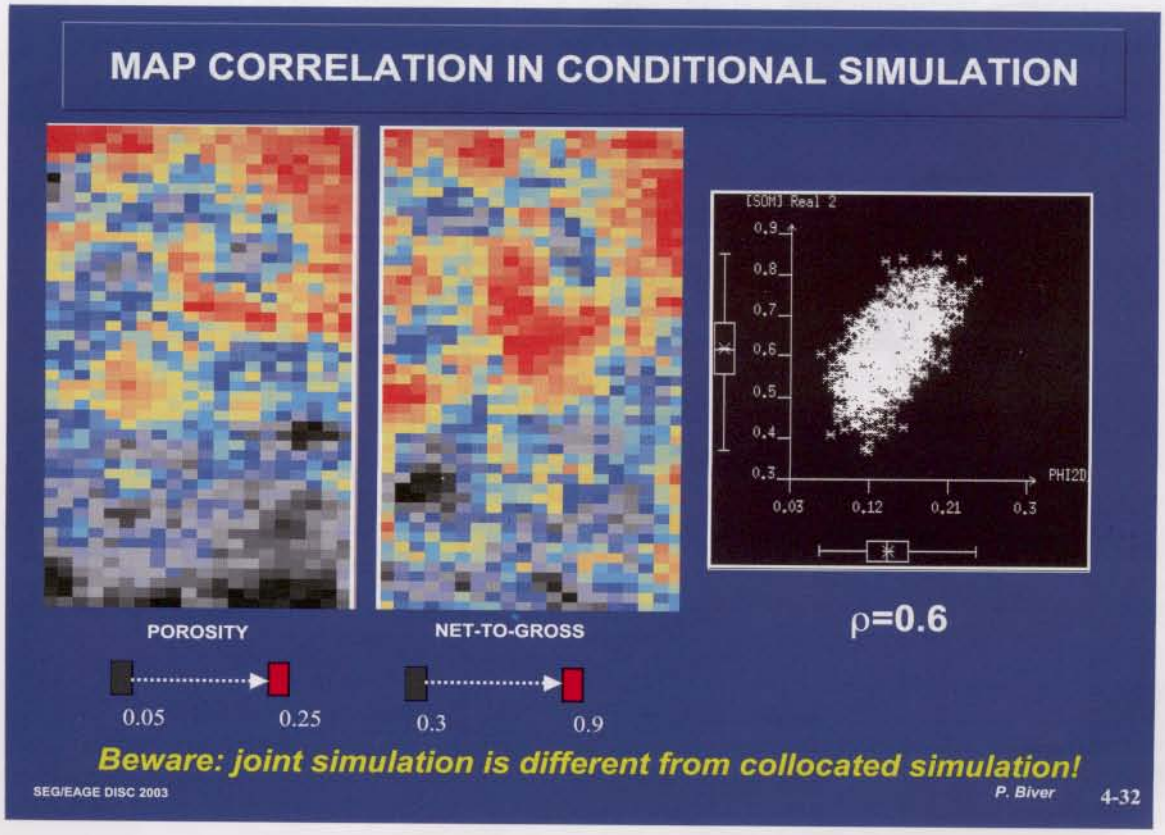
4.4.3 Cascade or parallel conditional simulation of several parameters?

Often, as Haas et al. (1988) discuss, different simulated variables are unlikely to play the same role. For instance, we may first consider a porosity simulation constrained by seismic attributes, and then other variables such as permeability or saturation can be generated from the porosity model. In this case, it is more logical to perform successive simulations, beginning with the most informed variable and using the previously simulated variables to constraint further steps. In other words, it may be better to perform cascade instead of parallel simulations (Fig. 4-33). This is an approach that has been clearly formalized by the Norwegian School, using diagrams such as that of Fig. 4-34. This diagram translates into a cascade simulation approach, where a reflection coefficient realization is first obtained from seismic data (this will be discussed later in the geostatistical inversion chapter), then acoustic impedance is derived from the reflection coefficient, then both permeability and porosity are derived from the acoustic impedance (Fig. 4-35). Another plausible cascade simulation approach would be to derive porosity from acoustic impedance, then permeability from porosity.

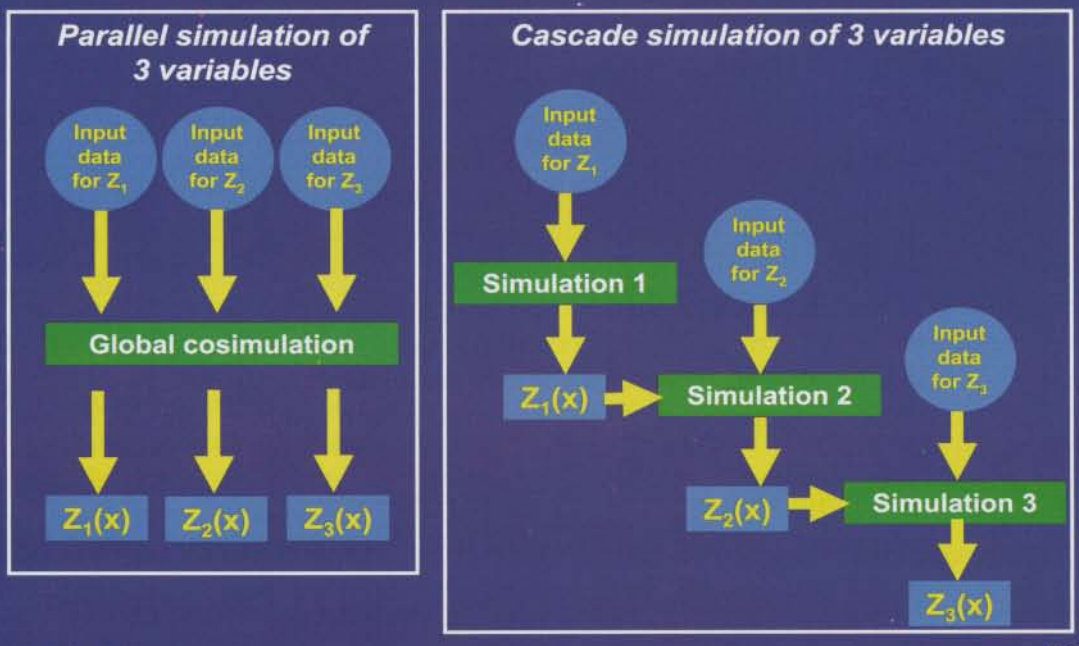
4.5 Conditional Simulation for Geological Facies Modeling

4.5.1 Introduction

After discussing the conditional simulation of continuous parameters such as porosity or water saturation, we are now going to address the simulation of discrete parameters,



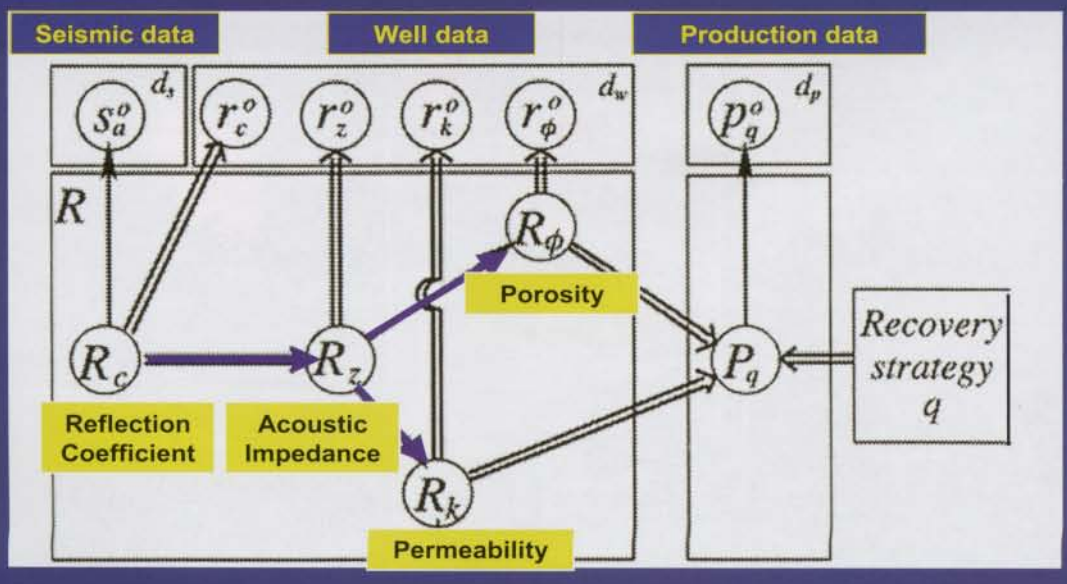
CASCADE OR PARALLEL SIMULATION? (HAAS ET AL, 1998)



SEG/EAGE DISC 2003

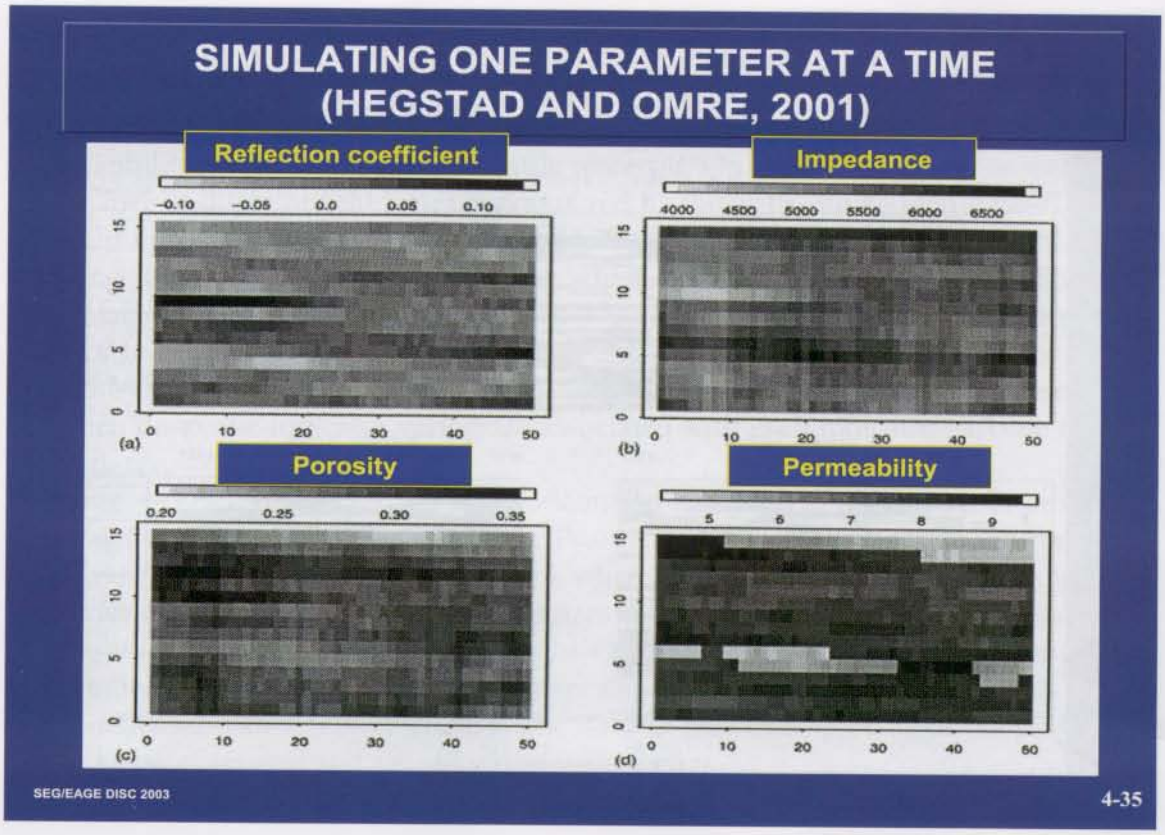
4-33

SIMULATING ONE PARAMETER AT A TIME (HEGSTAD AND OMRE, 2001)



SEG/EAGE DISC 2003

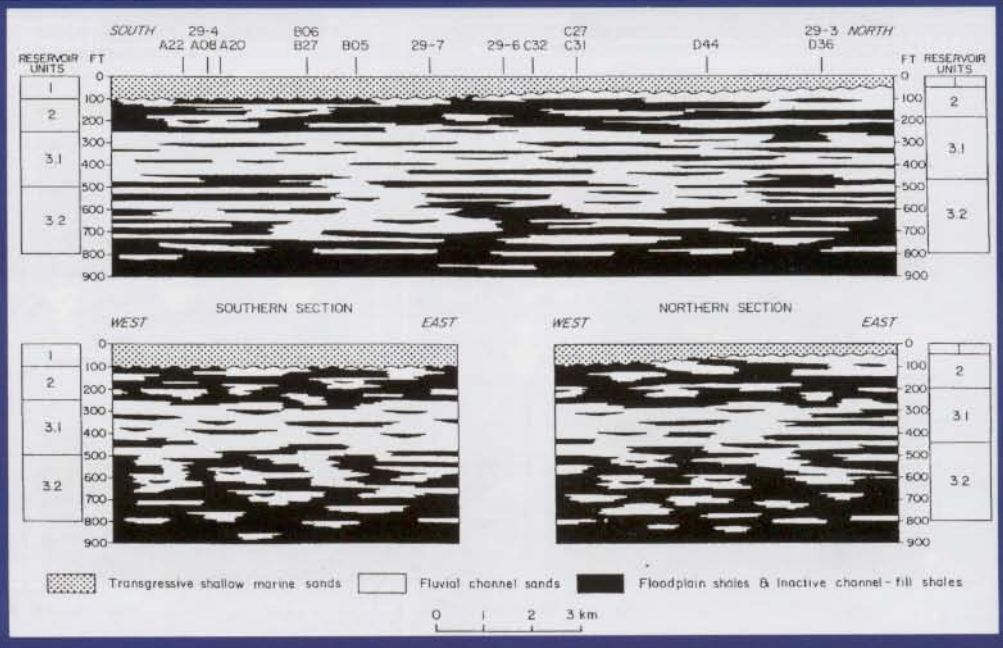
4-34



such as geological facies, lithology, or rock types. This is a very important topic in reservoir characterization and is described in more detail by Dubrule (1998). Here, we will focus on what needs to be known in order to incorporate seismic data into earth models.

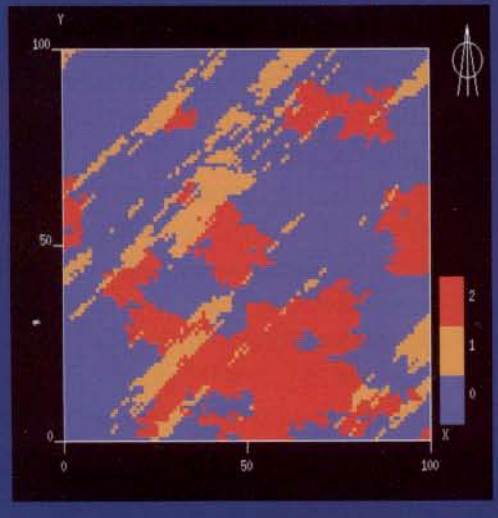
Geological quantification is a topic that has always raised much interest and debate among geologists. Fig. 4-36 shows three hand-drawn geological cross-sections of the Statfjord Formation in the Brent field (North Sea). Depending on the depositional environment he/she is dealing with, and using the well data as a constraint, the reservoir geologist can draw sketches of the distribution of sands and shales. Unfortunately, hand-drawn cross-sections are limited in that they do not lead to a 3D model, and they represent only one possible model among an infinity of scenarios matching the wells and compatible with the depositional environment. In the early 1980s, it became clear that geostatistical techniques could help generate such 3D geological scenarios. These scenarios will never be quite as realistic or "geologically loaded" as those produced by a geologist, yet they present the advantage of being multiple and three-dimensional. Today, there are two major classes of techniques available for generating 3D stochastic models: pixel-based and object-based models (Fig. 4-37). Object-based models (Clemetsen et al., 1990; Damsleth et al., 1990) assume that geological bodies such as channels or crevasse splays can be described using simple geometrical shapes. On the other hand, pixel-based models adopt a more modest approach by simply making assumptions about the statistical relationships between the facies types present at individual grid cells of the earth model.

HAND-DRAWN CROSS-SECTIONS FROM STATFJORD FORMATION, BRENT FIELD (JOHNSON & KROL, 1984)

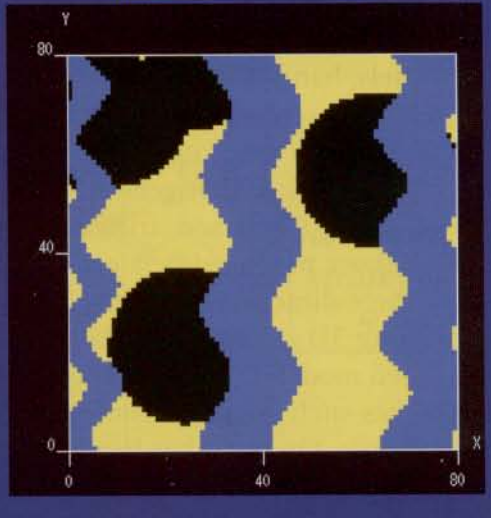


TWO MAIN KINDS OF STOCHASTIC TECHNIQUES FOR GENERATING 3D GEOLOGICAL MODELS

Pixel-based



Object-based



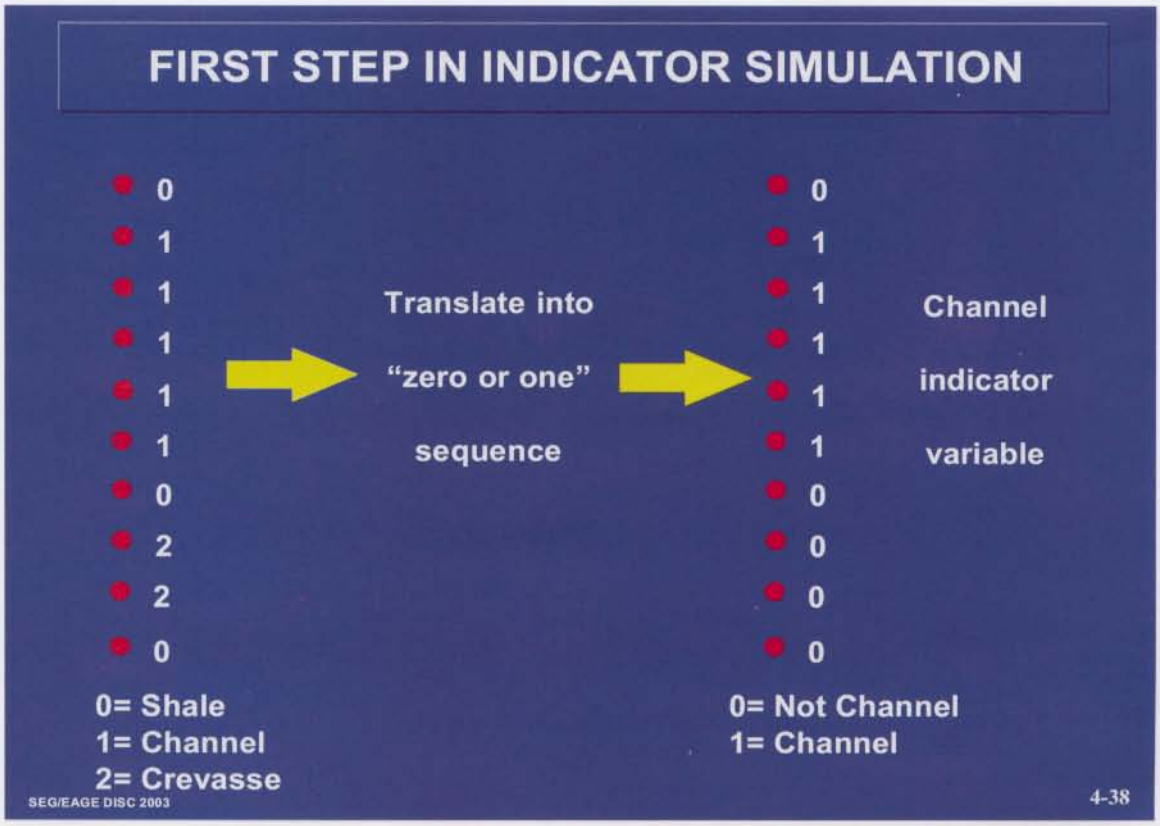
4.5.2 Pixel-based models

- Indicator simulation

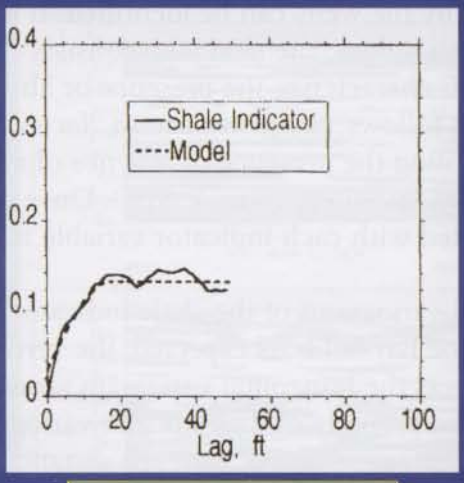
The most popular pixel-based approach is that based on the indicator variogram. Take the example of wells drilled in a fluviatile reservoir. On the basis of cores or well logs, a sequence of different lithologies encountered by the wells can be identified. If the different lithologies are coded by different integer values, the next step consists of defining "indicator variables" (Fig. 4-38), which characterize the presence or absence of the facies of interest along the wells (in what follows, the generic term "facies" will be used to designate a discrete variable representing the presence or absence of a geological feature, such as depositional environment, lithology, or rock type). Once this has been done, the indicator variogram associated with each indicator variable is calculated.

Fig. 4-39 shows the vertical and horizontal variograms of the shale indicator variable calculated within a small area of the Prudhoe Bay field. As expected, the vertical variogram range is on the order of meters, whereas the horizontal variogram range is on the order of kilometers. Fig. 4-40 summarizes the property of the indicator variogram and its associated covariance. The indicator variogram measures, for each distance, the probability that two facies found this distance apart are different. Of course, this probability increases with distance. There is also a direct relationship between the proportion of the facies of interest and the sill of the variogram.

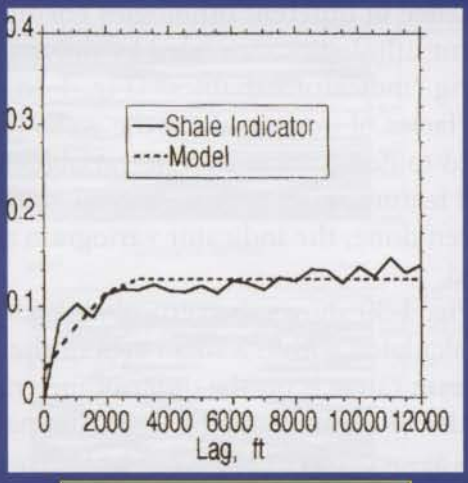
The most common models that are used for indicator variograms are the spherical



INDICATOR VARIOGRAMS FOR SHALE INDICATOR VARIABLE, PRUDHOE BAY, USA (PEREZ ET AL., 1997)



Vertical variogram

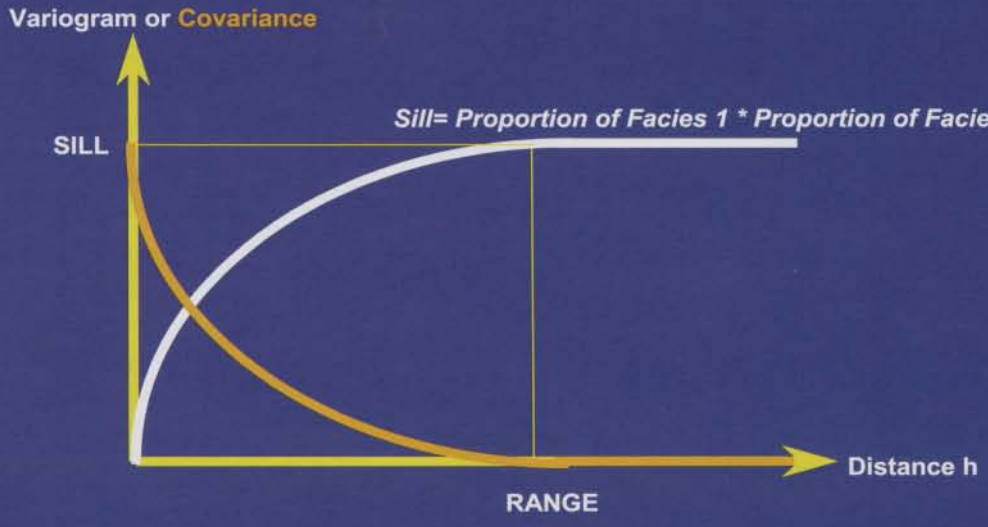


Horizontal variogram

SEG/EAGE DISC 2003

4-39

INDICATOR VARIOGRAM AND COVARIANCE



$\gamma(h)$ = 1/2 Average $[I(x)-I(x+h)]^2$ = 1/2 Proba (Facies found at distance h away are different)

COVARIANCE = SILL - VARIOGRAM

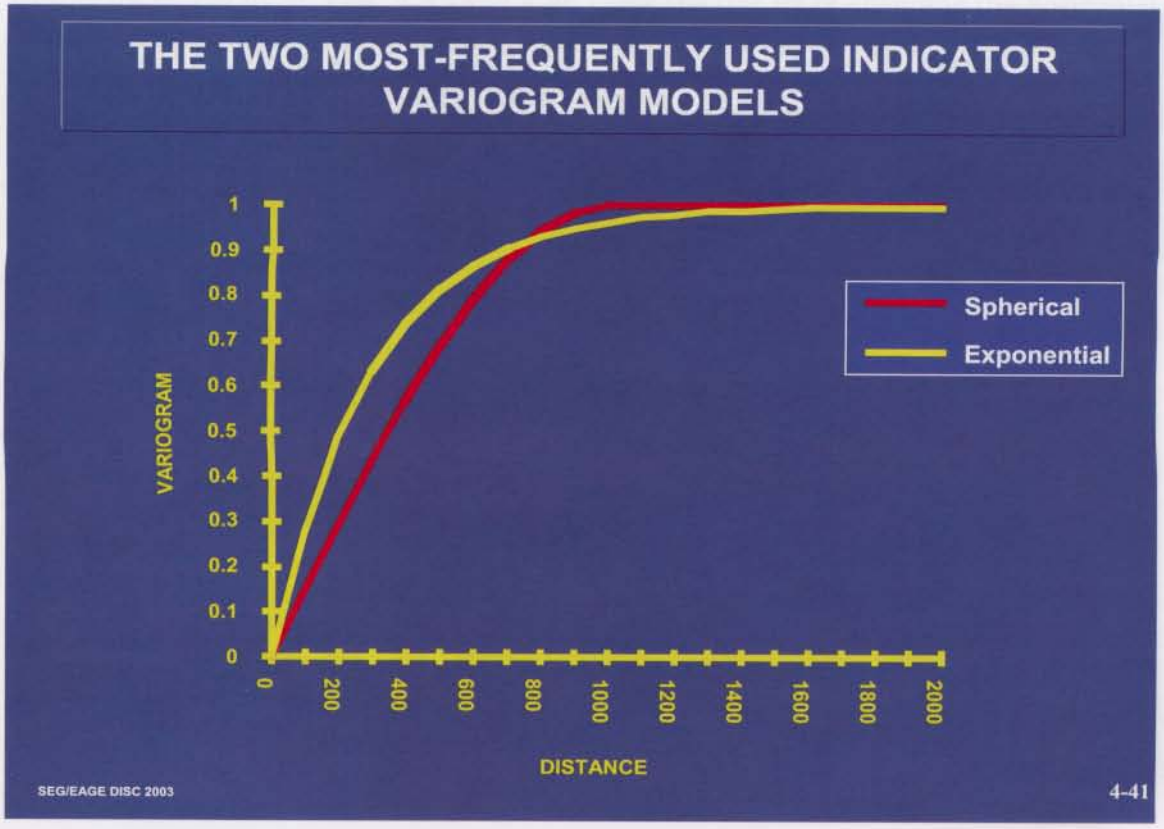
SEG/EAGE DISC 2003

4-40

and exponential models (Fig. 4-41). Understandably, the Gaussian model is too smooth to represent the kind of discontinuities associated with the changes of an indicator variable from 0 to 1.

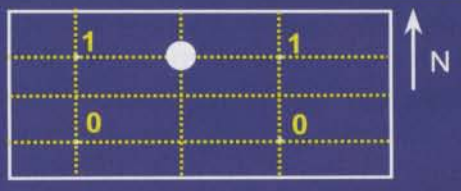
Once a model has been fitted to the experimental variogram, indicator simulation is usually performed using the sequential indicator simulation (SIS) algorithm originally developed by the Stanford School (Journel and Gomez-Hernandez, 1989). SIS (Fig. 4-42) is the generalization of SGS to discrete variables. The only slight difference lies in step 2 (compare Figs. 4-22 and 4-42). With SGS, step 2 consists of kriging the mean and standard deviation of the variable at the current location. With SIS, step 2 consists of kriging, from the surrounding indicator values, the probability of having a value of the indicator function equal to one. This probability is equal to the mean of all possible values of the indicator variable at that location.

To better understand the geological implications of using one indicator variogram model rather than another, we have computed nonconditional realizations for different kinds of models. The same random visitation path was used for the different models, in order to stress the impact of the choice of the variogram model, independently from that of the random path. Fig. 4-43 shows that there is very little difference between a realization generated by a spherical and that generated by an exponential variogram. We just expect the exponential model realizations to be a little bit noisier, because of the steeper slope of this variogram model at the origin. In the following examples, we will always use the spherical variogram model. Fig. 4-44 gives an idea, through the display of four realizations, of the change from one realization to another while keeping the

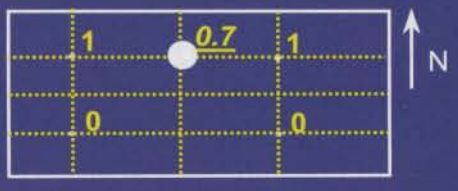


SEQUENTIAL INDICATOR SIMULATION (SIS) ALGORITHM

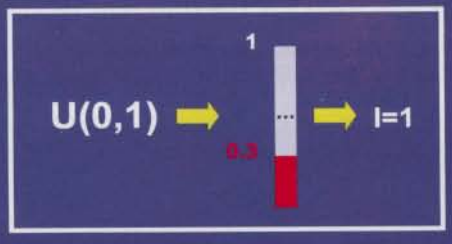
1. Draw random location between wells



2. Krige P(I=1) at random location



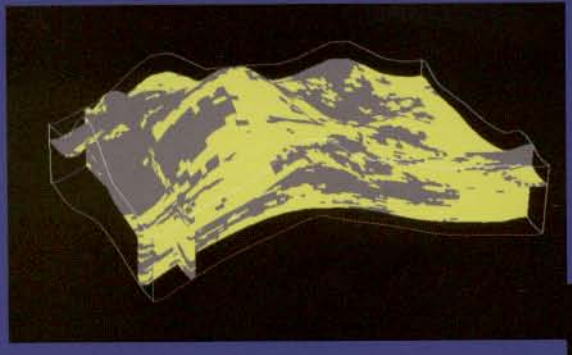
3. Sample indicator with probability 0.7

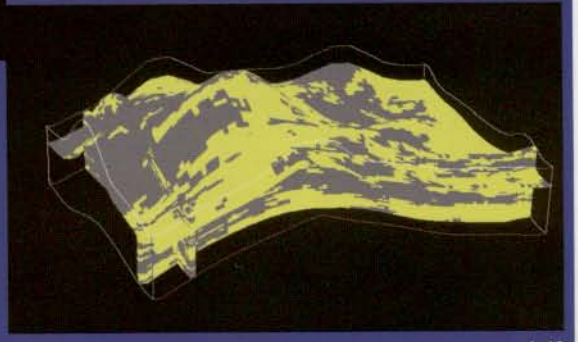


4. Merge sampled value & draw new location

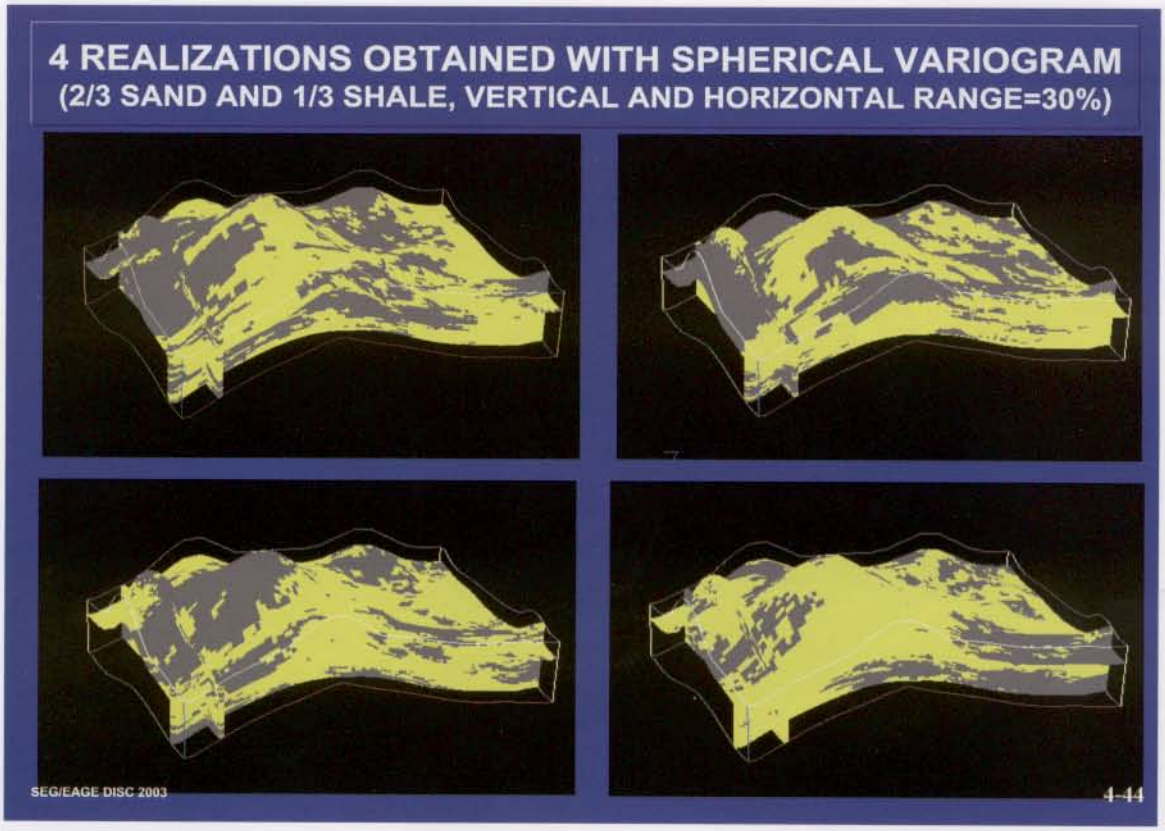
IMPACT OF CHANGING INDICATOR VARIOGRAM MODEL (2/3 SAND AND 1/3 SHALE, VERTICAL AND HORIZONTAL RANGE=30%)

Spherical





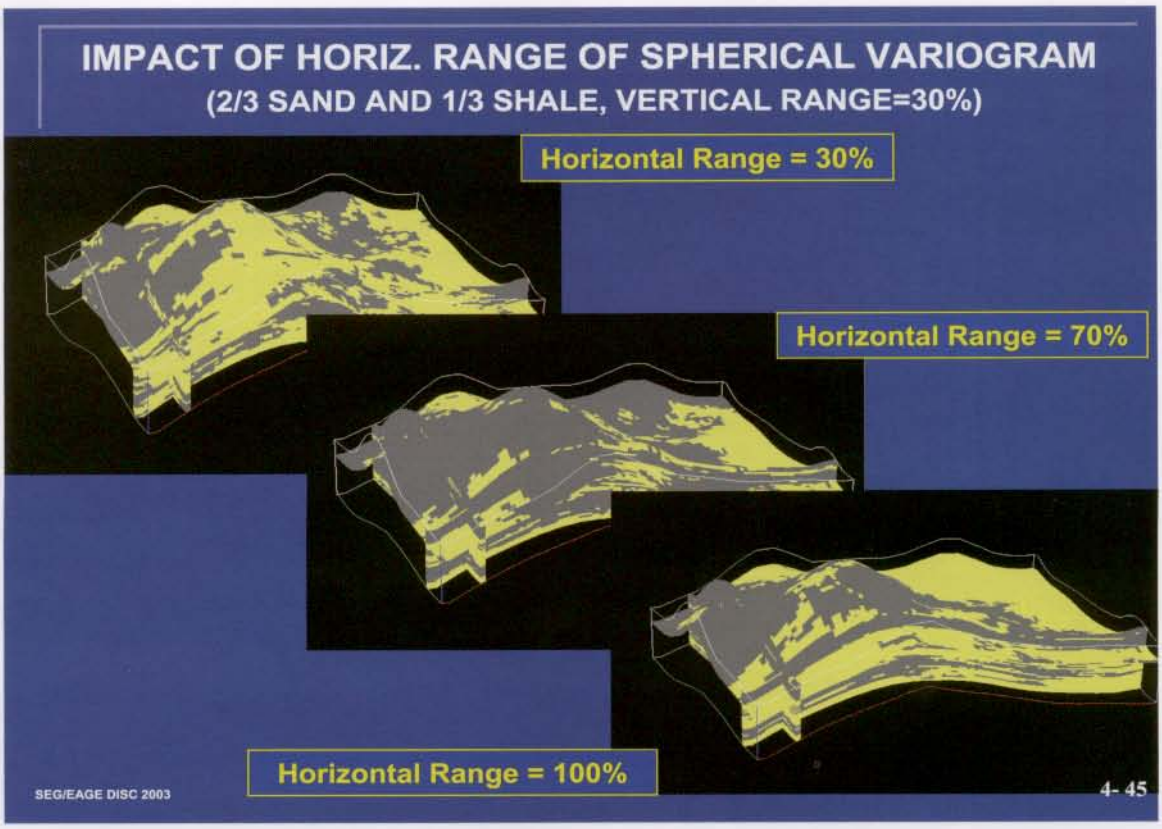
Exponential

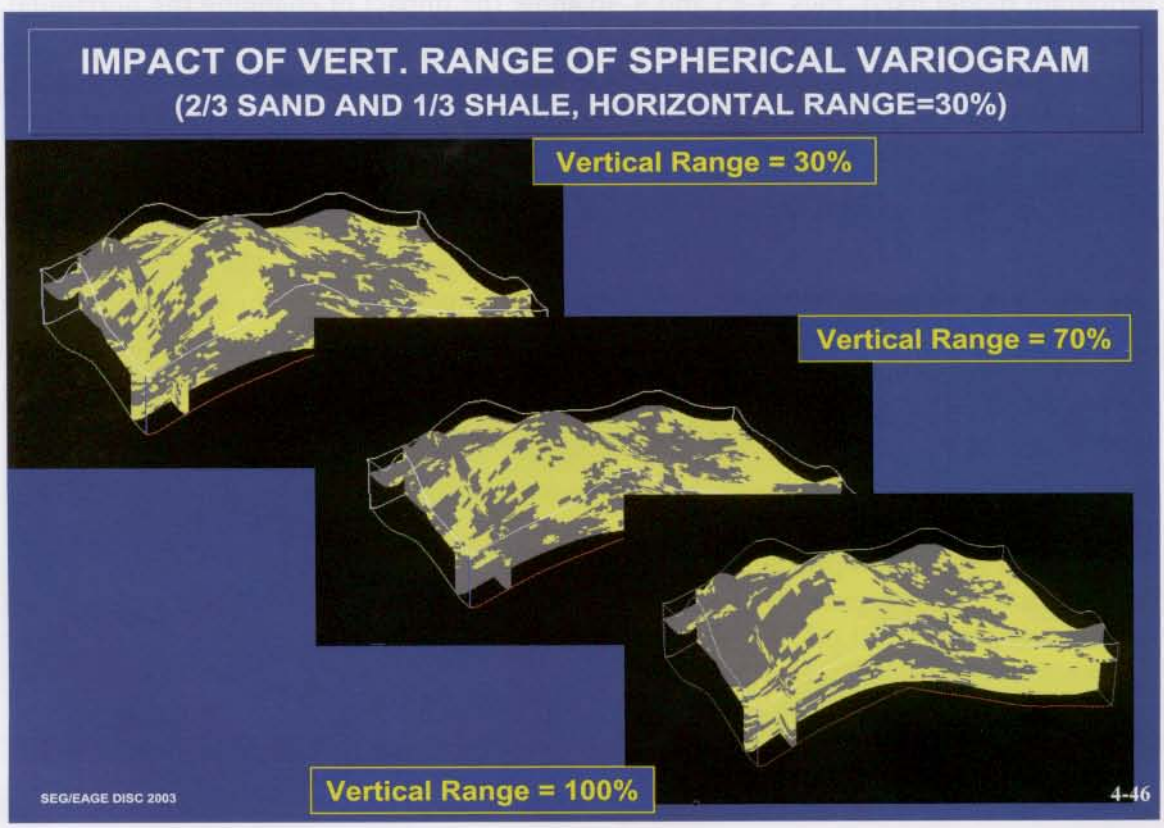


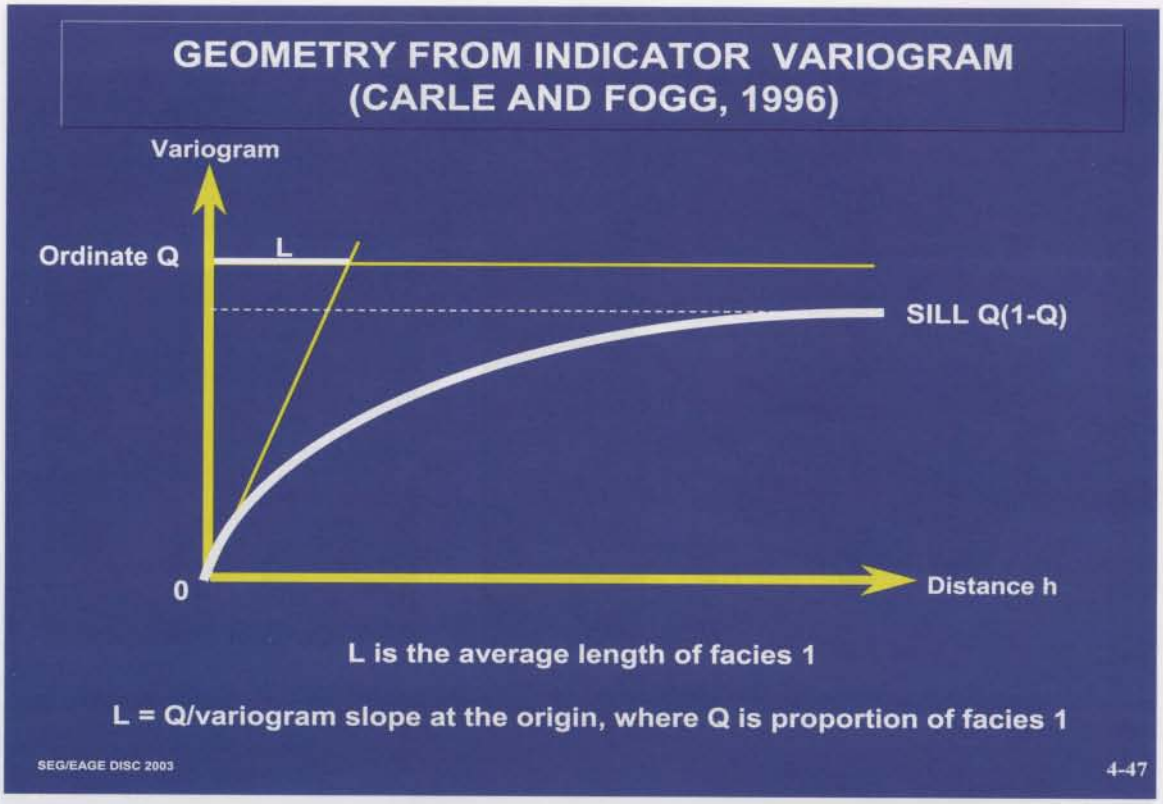
variogram model fixed. Figs. 4-45 and 4-46 show the impact of changing the horizontal or the vertical range of the variogram model.

Too often, the range of an indicator variogram is considered to be representative of the size of the individual facies bodies. This is wrong. To be convinced of this, just think that the indicator variogram of the "shale" occurrence is the same as the indicator variogram of the "non-shale" occurrence. How could the range be representative of both the size of shale and non-shale bodies? Fig. 4-47 clarifies the relationship between the size of a geological object and the range of the indicator variogram. The spatial facies proportion Q plays a significant role in the calculation. The greater the proportion, the more the regions of occurrence of the facies of interest will be connected with each other, thus increasing their average lateral extent. Carle and Fogg (1996) derived the relationship between the parameters of the indicator variogram and the transition probabilities used in Markov-Chain simulation (which should not be confused with the already discussed MCMC). A recent paper by Ritzi (2000) further develops the results of Carle and Fogg.

The previous examples show that the outcome of SIS is not very strongly constrained in terms of geometry. This is because the indicator variogram itself does not carry much geometrical information. This may prove to be a blessing in situations where the geometry of various geological bodies is poorly known. In carbonate reservoirs for instance, SIS is often used, and we will see later a very good example of it (Figs. 4-63 to 4-66). However, in fluvial or fluvio-deltaic environments, for instance, it may be necessary to control the geometry of the simulated bodies. This is the goal of


IMPACT OF HORIZ. RANGE OF SPHERICAL VARIOGRAM
(2/3 SAND AND 1/3 SHALE, VERTICAL RANGE=30%)
Horizontal Range = 30%
Horizontal Range = 70%
Horizontal Range = 100%
SEG/EAGE DISC 2003
4- 45


IMPACT OF VERT. RANGE OF SPHERICAL VARIOGRAM
(2/3 SAND AND 1/3 SHALE, HORIZONTAL RANGE=30%)
Vertical Range = 30%
Vertical Range = 70%
Vertical Range = 100%
SEG/EAGE DISC 2003
4-46



object-based models. But before presenting these models, let us discuss other pixel-based techniques used in the industry.

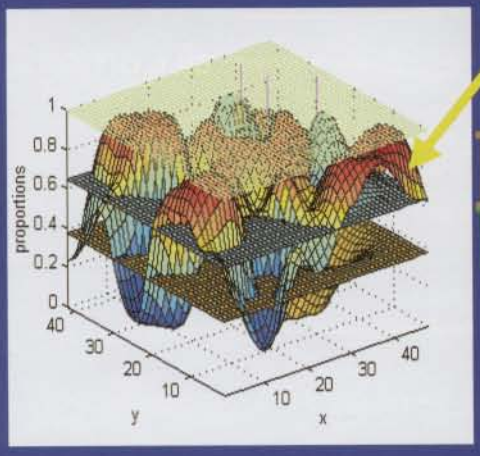
- Other pixel-based models

The truncated Gaussian simulation method was developed by the Heresim Group (Rudkiewicz et al., 1990). Fig. 4-48 is an illustration of the technique and an example of the kind of 2D realization obtained with a three-facies model. In this 2D example, the method consists of generating the pixel-based model in two steps. First, a continuous Gaussian simulation is generated using a standard simulation technique, then two cut-offs are applied to it, defining three different "facies." Because of the way the facies map is constructed, spatial relationships between different facies are automatically introduced. For instance, in Fig. 4-48, direct transition from facies 1 to facies 3 is not possible without going through facies 2. These implicit relationships may prove useful when they correspond to the modeled geology. In other situations, they are a limitation of the method. Truncated Gaussian simulation can also be used for generating realizations composed of two facies only (Fig. 4-49).

The generalization of Markov-Chain simulation (not to be confused with MCMC) from one to three dimensions has always proved very difficult. In a recent paper, Parks et al. (2000) propose to use the flexibility of simulated annealing algorithms to generate grids with Markov statistical structures honoring the wells. This is a promising area of research, because the quantification of geological patterns by Markov chains has always been attractive to geologists. However, convincing 3D applications are yet to be produced.

A 2D EXAMPLE OF TRUNCATED GAUSSIAN SIMULATION

Define 2 truncation thresholds



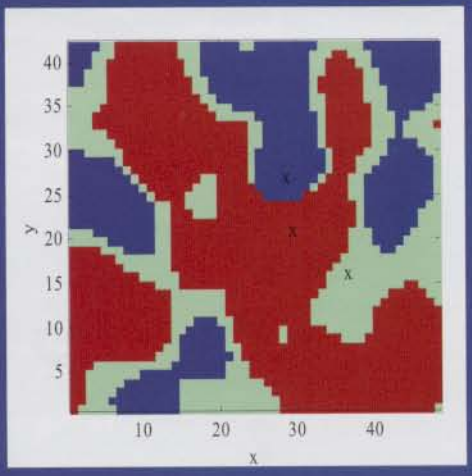
Surface to truncate

Facies 3

Facies 2

Facies 1

Generate facies map

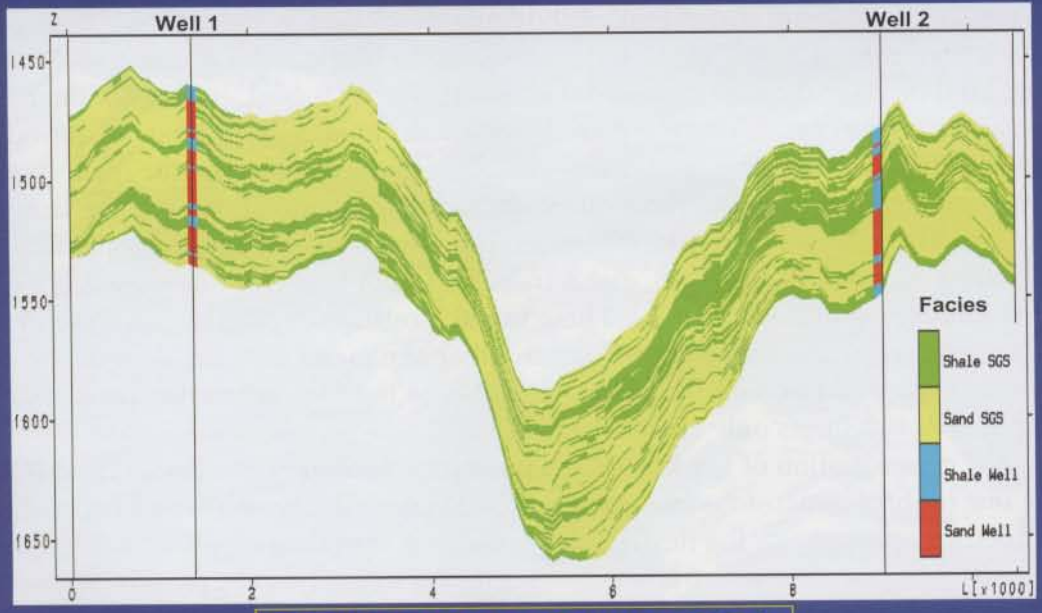


Facies 1 2 3

SEG/EAGE DISC 2003

P. Biver 4-48

AN EXAMPLE OF TRUNCATED GAUSSIAN SIMULATION



Simulated facies (Range = 2 Km)

SEG/EAGE DISC 2003

X. Freulon 4-49

Strebelle and Payrazyan (2002) argue that the variogram, which is a two-point measure of spatial variability, cannot describe realistic geological features. They develop multiple-point geostatistics using a training image instead of a variogram to quantify geological heterogeneity. The training images are nonconditional representations of the geology of interest that can be hand-drawn by a geologist or produced by geological modeling software (process models).Then the method "learns" what statistical patterns are contained in the training images and reproduces these patterns using a conditional simulation approach. Existing seismic data can also be used to further constrain the probability to find a given lithology or facies at a location of the model. Multiple-point geostatistics addresses a real issue, that of injecting more geological information in geostatistical models. However, at this stage, it remains very difficult to use, and a number of issues have to be addressed to make the workflow applicable in routine petroleum applications.

The use of Markov random fields is not to be confused with that of Markov chains or MCMC. Markov random fields are very popular in image analysis (Besag, 1974) for quantifying the statistical relationship between one pixel and the surrounding pixels. (I. Eidsvik et al., personal communication, 2003) express the probability of having sand at one grid cell as an exponential function of the number of sand grid cells in the neighborhood. Farmer (1988, 1992) introduced these models to the oil industry, but they have proved difficult to use for quantifying geology. This is because, as with the method of Strebelle and Payrazyan, they require the use of a training image for calibrating the multipoint statistics quantified by the Markov random field model.

4.5.3 Object-based models

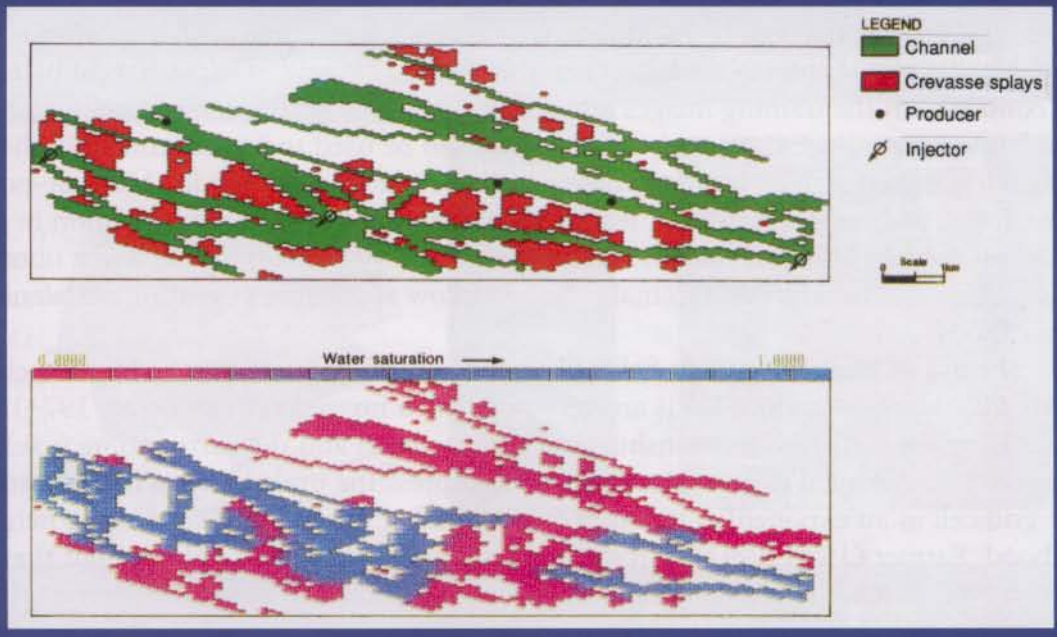
The Norwegians pioneered the development of object-based models, which produced satisfactory representations of the distribution of channels in fluvio-deltaic formations within some of their giant fields (Fig. 4-50). The difference from SIS realizations is striking. Instead of dealing with "salt-and-pepper" models, poorly constrained geometrically, we now have distributions of objects with well-defined shapes and varying size.

The model of Fig. 4-51 was produced within a fluviatile formation in the North Sea. In this formation, we were dealing with meandering channels, the architecture of which could be better constrained using object-based models rather than using SIS. In Fig. 4-52, we are also dealing with fluvial channels, but they are assumed to be straighter than in the previous case.

Object-based modeling simply assumes that the various facies are associated with well-defined geometries and that their size (width, thickness, and length) is random and can be statistically quantified by the reservoir geologist. Object-based simulation can also incorporate constraints about the relative positions of different geological objects (for instance, crevasse splays, which result from sediment spills on the edges of channels, must always be located close to a channel, see Fig. 4-52). Trends can also be incorporated for controlling the proportion of various objects vertically or spatially.

Object-based simulation algorithms work easily in situations where the well spacing is much greater than the lateral dimension of the modeled objects. However, conditioning is much more difficult when sand-body dimensions are large compared with well spacing, because well-to-well geological correlations are difficult to handle

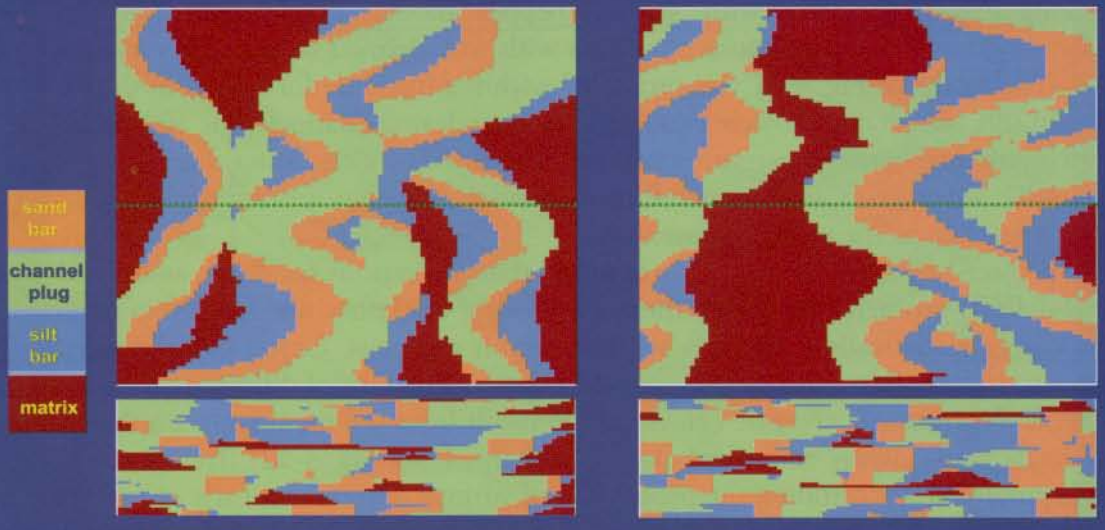
AN EARLY EXAMPLE OF OBJECT-BASED MODEL, NESS FORMATION, NORTH SEA (GUNDESO AND EGELAND, 1990)



SEG/EAGE DISC 2003

4-50

OBJECT-BASED MODELS FOR MEANDERING CHANNELS

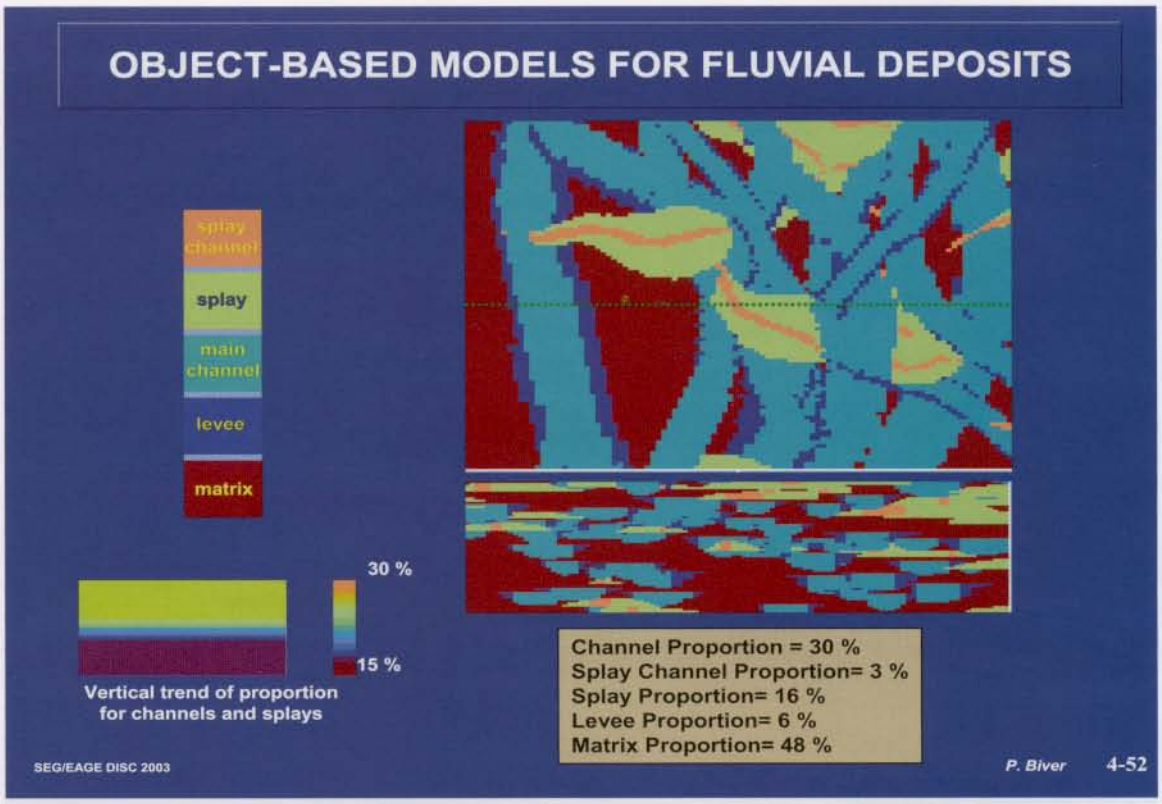


Sand Bar Proportion = 22 %
Silt Bar Proportion = 19 %
Plugs and Matrix Proportion = 59 %

Sand Bar Proportion= 23 %
Silt Bar Proportion= 23 %
Plugs and Matrix Proportion = 54 %

SEG/EAGE DISC 2003

P. Biver 4-51

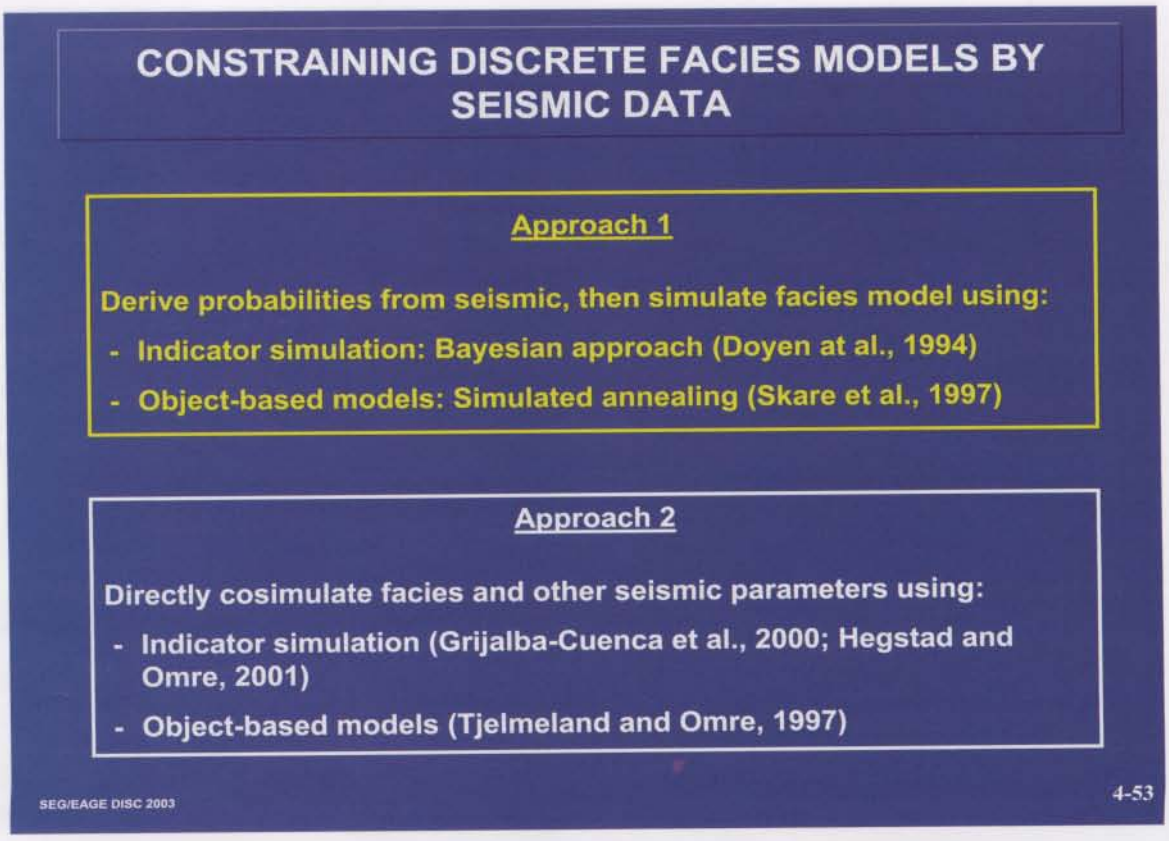


(Dubrule, 1998). Object-based models are not "better" or "worse" that indicator simulation. They are just better suited to certain kinds of geological environments. So far, their main application has been in fluvio-deltaic environments.

4.5.4 Facies models constrained by seismic information

We have already seen how information from a seismic attribute could be incorporated to reduce the uncertainty on the kriging or conditional simulation of the primary variable, when this variable was continuous. There may also be situations where the seismic data provide direct information about the distribution of various geological facies in the reservoir. Everybody has seen these impressive seismic-attribute maps from deep offshore reservoirs, showing clear meandering channel patterns. The challenge in constraining facies models by seismic data is that a discrete parameter (presence or absence of a given facies) must be constrained by a continuous (seismic-attribute) parameter.

We may wish to constrain an indicator simulation or an object model using seismic data. In both situations, two approaches are possible (Fig. 4-53). The first one is to calculate the probability that a facies is present at a given location, then use this probability as a constraint in the simulation. The other one is to directly produce a joint simulation of facies and acoustic impedance, making sure that they are consistent with seismic data. In this section, we will discuss techniques whereby facies distributions are constrained by 3D probability models derived from seismic. Later on, in the geostatistical inversion chapter, we will see how joint realizations of facies and acoustic impedance distributions can be directly constrained by seismic data.



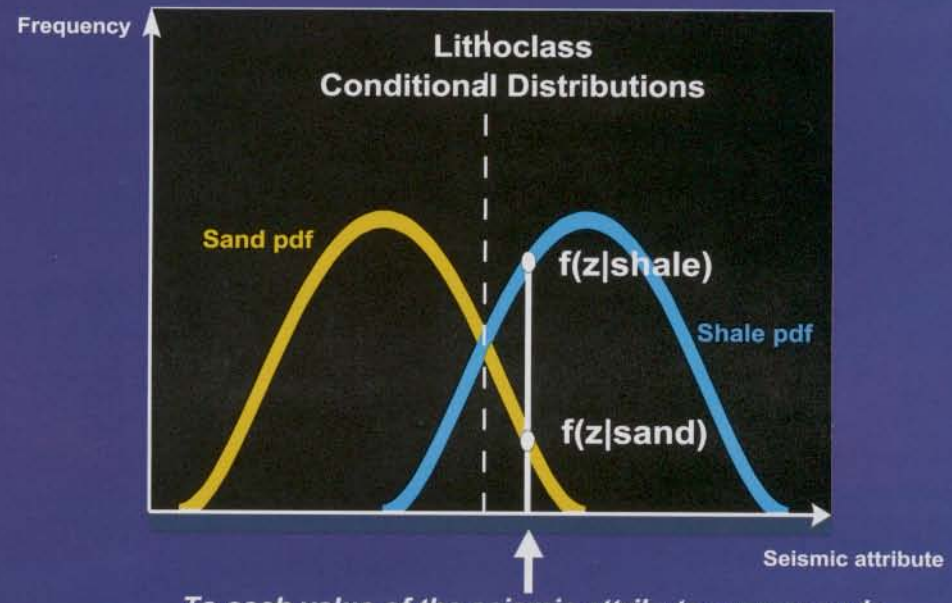
- *Constraining pixel-based models by probabilities derived from seismic data*

Doyen et al. (1994) propose a very elegant approach that is a generalization to discrete models of the methodology developed for continuous variables (Figs. 3-73 to 3-75). Commonly, when relating facies information with seismic data, a plot such as that of Fig. 4-54 is derived from the seismic-to-well calibration. This plot gives the seismic-attribute histogram associated with each facies (here, the two facies are shale and sand, and are called lithoclasses). The more difference there is between histograms, the more hope there is that seismic will help discriminate between lithoclasses. Now, consider a grid cell of the earth model where the seismic attribute has been mapped. Based on the value of this attribute, a probability of the presence of shale or sand can be derived (Fig. 4-54). Doyen et al. interpret this value as a likelihood function, similar to the one derived from the crossplot in the case of continuous variables (Fig. 3-73).

Now, assume we are running SIS (Fig. 4-55). We know (Fig. 4-42) that SIS provides, at each sampled location, the probability of finding sand or shale, on the basis of the facies encountered at the wells and at the previously simulated locations. This can be interpreted as an a priori probability, based on wells and on the a priori geological model quantified by the variogram.

Now we have two pieces of information about the probability of finding shale or sand: the prior probability given by SIS, and the likelihood given by the seismic attribute. Doyen et al. simply combine these two pieces of information using Bayes's theorem (Fig. 4-56). The actual simulation runs are performed using SIS, with the only

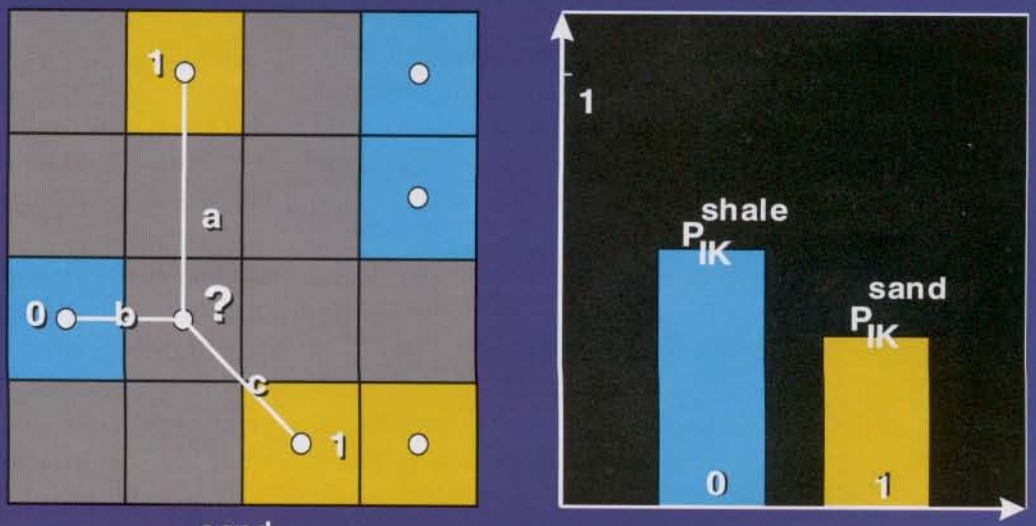
CONSTRAINING FACIES SIMULATIONS BY PROBABILITIES. QUANTIFYING THE LIKELIHOOD (DOYEN ET AL., 1994)



To each value of the seismic attribute corresponds a probability of having sand or shale

4-54

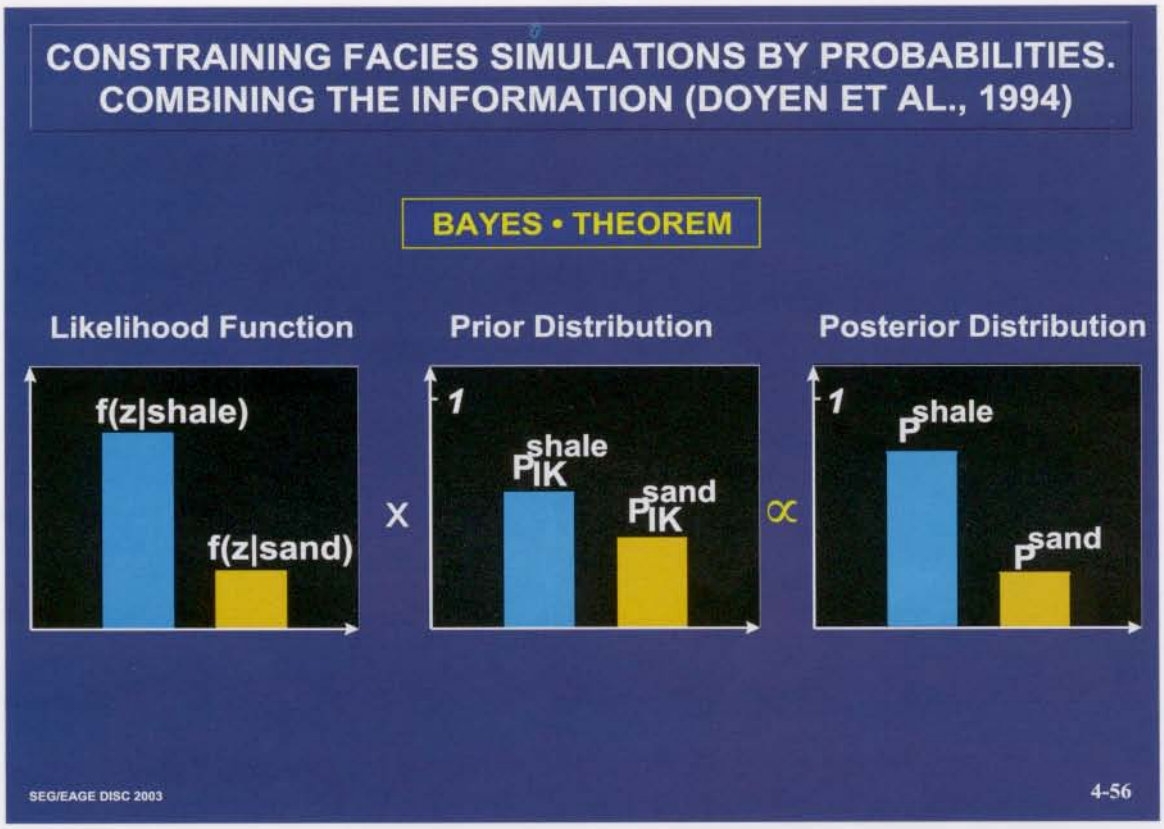
CONSTRAINING FACIES SIMULATIONS BY PROBABILITIES. QUANTIFYING PRIOR PROBABILITIES (DOYEN ET AL., 1994)



$$P_{IK}^{sand} = a \cdot 1 + b \cdot 0 + c \cdot 1 + d = 0.4$$

$$P_{IK}^{shale} = 1 - P_{IK}^{sand} = 0.6$$

4-55



modification that, at each new sampled location, the probability of finding each facies is calculated using the posterior rather than the prior probability.

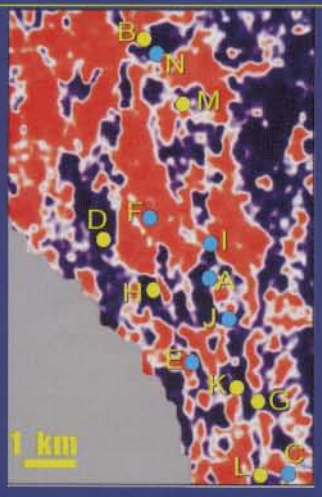
Figs. 4-57 and 4-58 are an application of Doyen et al.'s methodology in the Ness Formation of the Oseberg field (North Sea). Seismic amplitude is first calibrated against occurrences of channels at the 14 wells, showing that there is a contrast in reflection strength between channel and non-channel deposits. However, there is also some overlap between the two histograms, making the discrimination difficult on the basis of seismic data. In addition to the approach just described, Doyen et al. constrained the SIS realizations using locally identified channel directions, as picked on seismic data by the interpreter. As a result, the realizations of Fig. 4-58 show a significantly curved shape.

Insalaco et al. (2001) present an application of this approach to the detailed geological modeling of a West Africa turbidite deposit. Based on the pdfs associated with each individual facies, they produce a global histogram of acoustic impedance per facies (Fig. 4-59). Then, after translating this histogram into a likelihood function, they produce a model of the probability of encountering each facies at every grid cell of the earth model (Fig. 4-60 shows the probability of encountering facies 8). This information is validated and, if needed, modified by the sedimentologist. Using this as input, a realization of the 3D facies model is produced (Fig. 4-61).

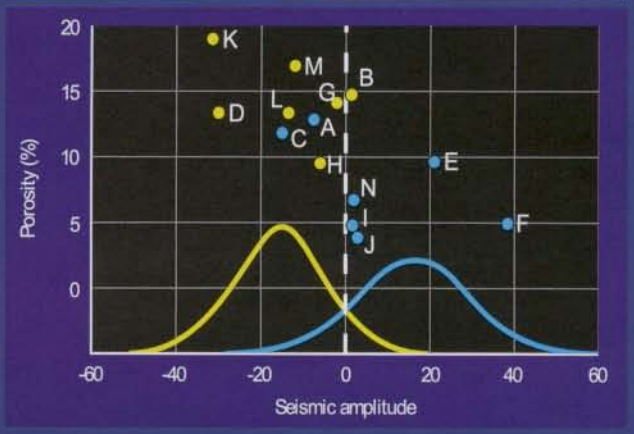
Lo and Bashore (1999) propose a similar approach to those of Doyen et al. (1994) and Insalaco et al. (2001). They obtain a 3D density model by inversion, then translate it into probabilities of various facies being present. Facies realizations are constrained by

CONSTRAINING FACIES SIMULATIONS BY PROBABILITIES. CASE STUDY (DOYEN ET AL., 1994)

Seismic amplitude map at Base Upper Ness



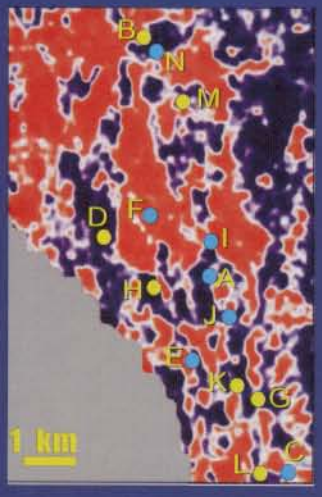
Likelihood function derived from well data



● channel well ● non-channel well

CHANNEL SAND SIMULATIONS CONSTRAINED BY SEISMIC AND WELLS (DOYEN ET AL., 1994)

Seismic amplitude map at Base Upper Ness

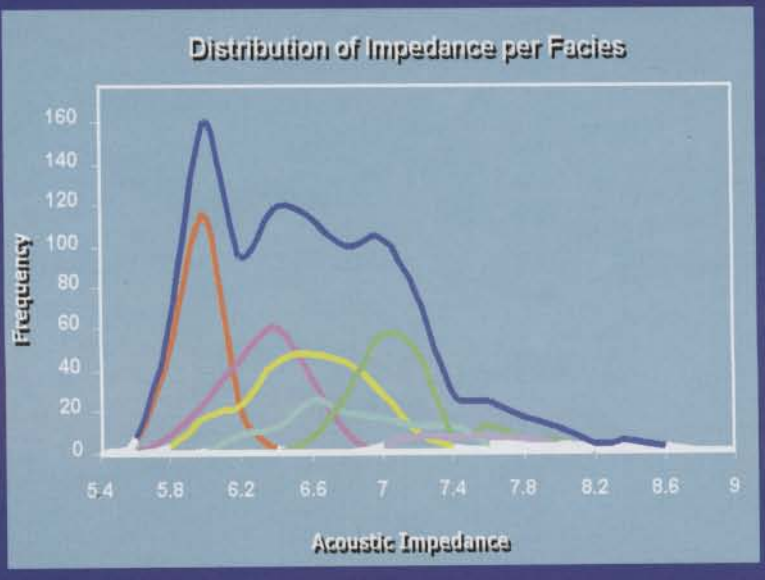


Realization 1

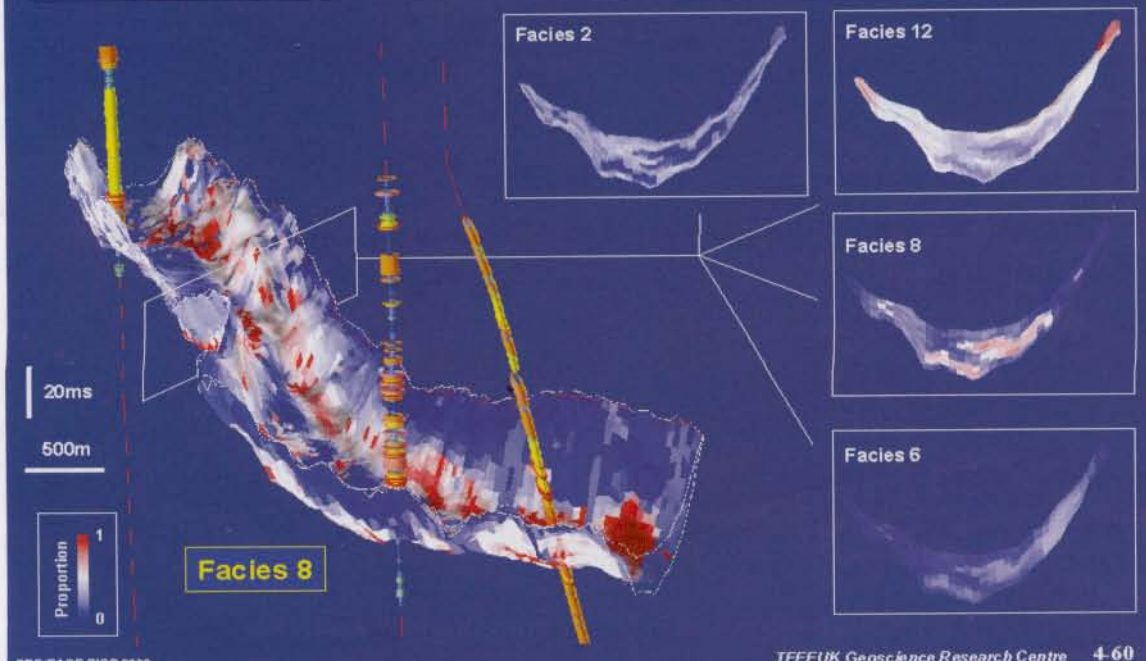
Realization 2

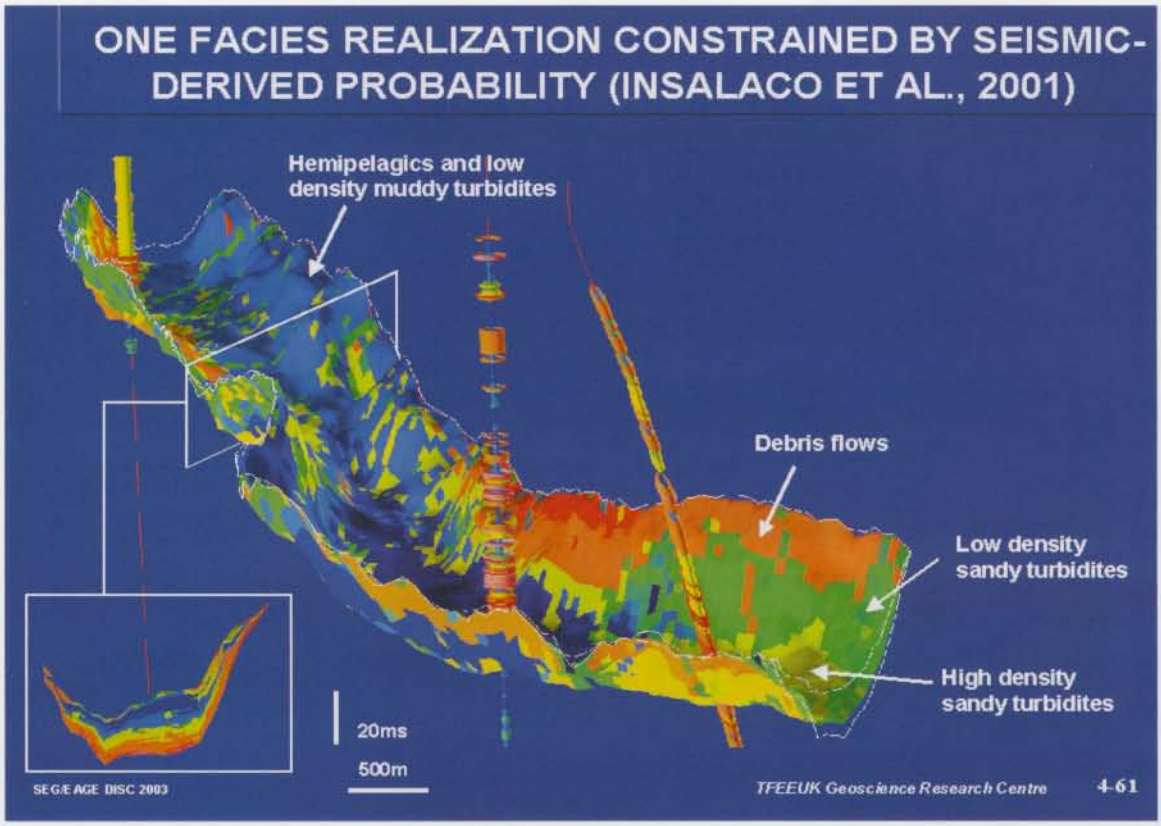
● channel well ● non-channel well

CONSTRAINING FACIES SIMULATIONS BY PROBABILITIES (INSALACO ET AL., 2001)



FACIES PROBABILITIES DERIVED FROM ACOUSTIC IMPEDANCE (INSALACO ET AL., 2001)



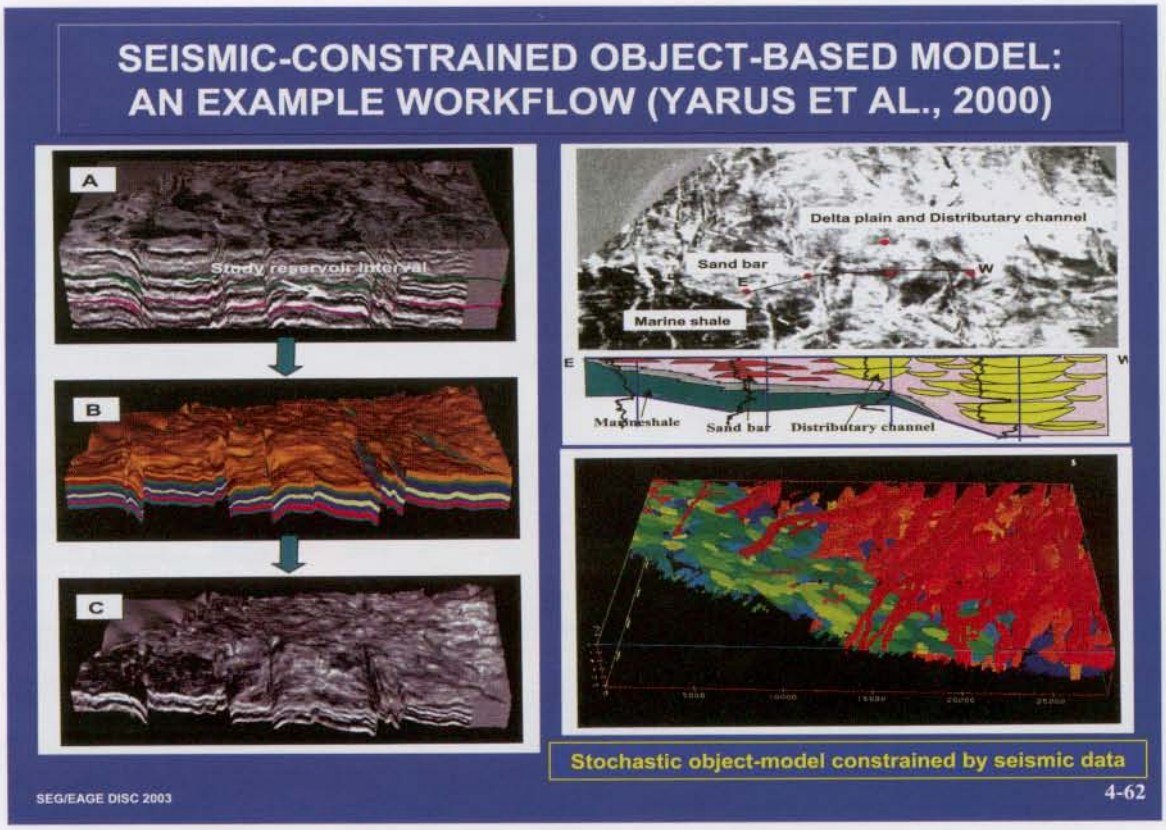


these probabilities and used as a template to predict porosity from acoustic impedance, using the crossplot associated with the facies present at each location.

- Constraining object-based models by seismic data

It is fair to say that there is not much experience in the industry about constraining object-based models by seismic data. The Norwegian School (Macdonald et al., 1995; Skare et al., 1997; and Holden et al., 1998) has been a leader in developing this approach, which assumes that — as the previously described method does — thanks to seismic data, it is possible to define a likelihood function at each grid cell of a 3D earth model. Then a simulated annealing algorithm is used to generate realizations by iteratively adding and subtracting channel bodies from the simulated volume. At each iteration, a new objective function is calculated. The channel is accepted if the objective function is decreased and is removed otherwise. One of the terms of the objective function evaluates the likelihood function of the new realization. The higher this likelihood, the smaller the objective function.

Yarus et al. (2000) describe an application of such a methodology for constraining an object-based model by seismic data in a Tertiary reservoir in the Gulf of Thailand (Fig. 4-62). The environment is tide-dominated, associated with shallow-marine deposits. A 3D probability cube was derived from the seismic data for each facies in each zone. Then an object-based model was used to stochastically model the distribution of bar and channel facies bodies within a shaly background.



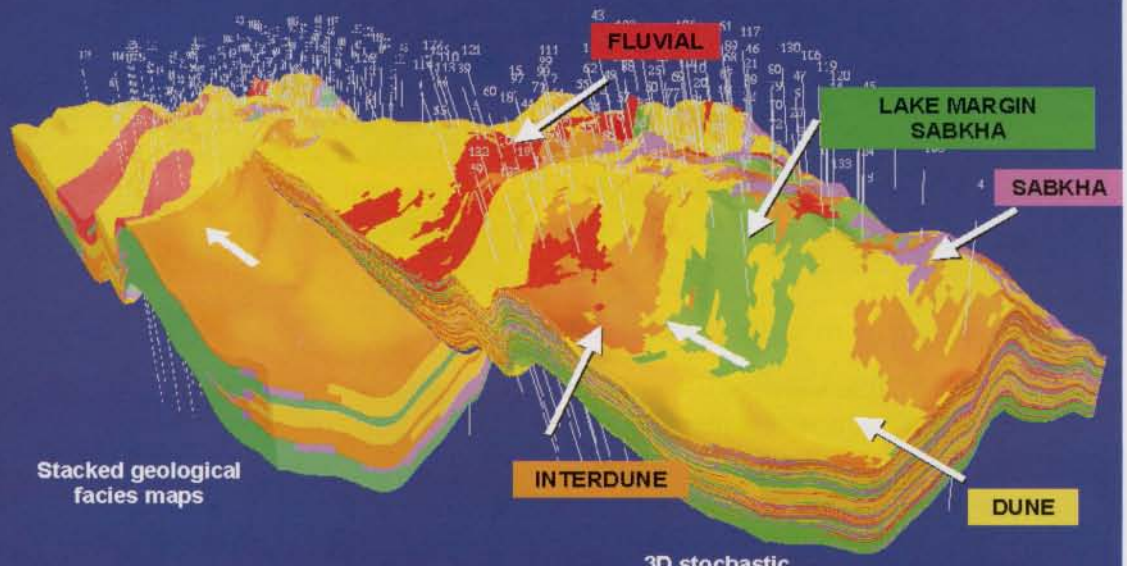
4.5.5 *Hierarchical modeling of geology and petrophysical parameters*

- An example

We have seen numerous examples of techniques for the conditional simulation of discrete and of continuous variables. So far, the techniques have been presented somewhat independently from each other. However, the strength of these techniques lies in the ability to combine them together for modeling different reservoir features at different scales. Alabert and Massonnat (1990) and Damsleth et al. (1990) give very good examples of early multiscale heterogeneity modeling studies, and Haldorsen and Damsleth (1990) provide a good overview of the topic.

Rather than formalize this, let us take an example from Al Qassab et al. (2000). This covers the construction of a 3D earth model of the Unayzah reservoir in the Hawtah field (Saudi Arabia). A large number of well data is available, and the reservoir is subdivided in 13 layers. First, a depositional facies model is constructed using SIS, under the control of sedimentary maps produced by the geologist (Fig. 4-63). The second step consists of producing a rock-type model accounting for the fact that, within a depositional facies, rock types can change. Three rock types are identified: reservoir, intermediate, and nonreservoir. Fig. 4-64 compares the rock-type model we would obtain by using (left) or not using (right) the control provided by the depositional-environment-facies model. Obviously, the model controlled by the depositional environment is better. Thus, at each simulation step, the generation of stochastic simulations is performed independently within each region defined by the previous steps.

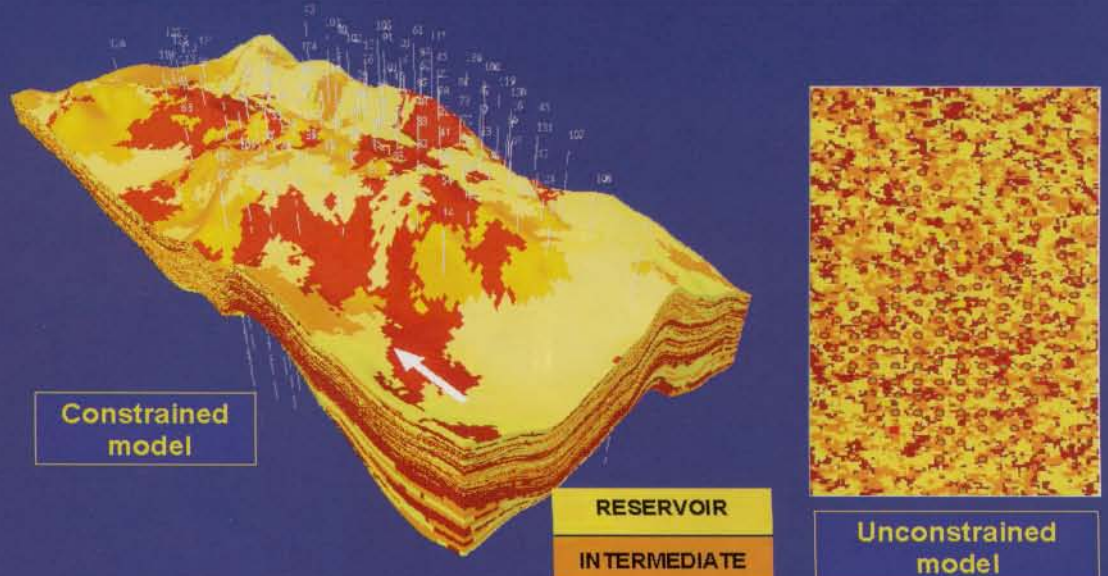
3D MODEL OF DEPOSITIONAL ENVIRONMENT FACIES, UNAYZAH RESERVOIR, HAWTAH FIELD, SAUDI ARABIA (1) (AL QASSAB ET AL., 2000)



SEG/EAGE DISC 2003

4-63

3D PETROPHYSICAL ROCK MODEL, UNAYZAH RESERVOIR, HAWTAH FIELD, SAUDI ARABIA (2) (AL QASSAB ET AL., 2000)



SEG/EAGE DISC 2003

4-64

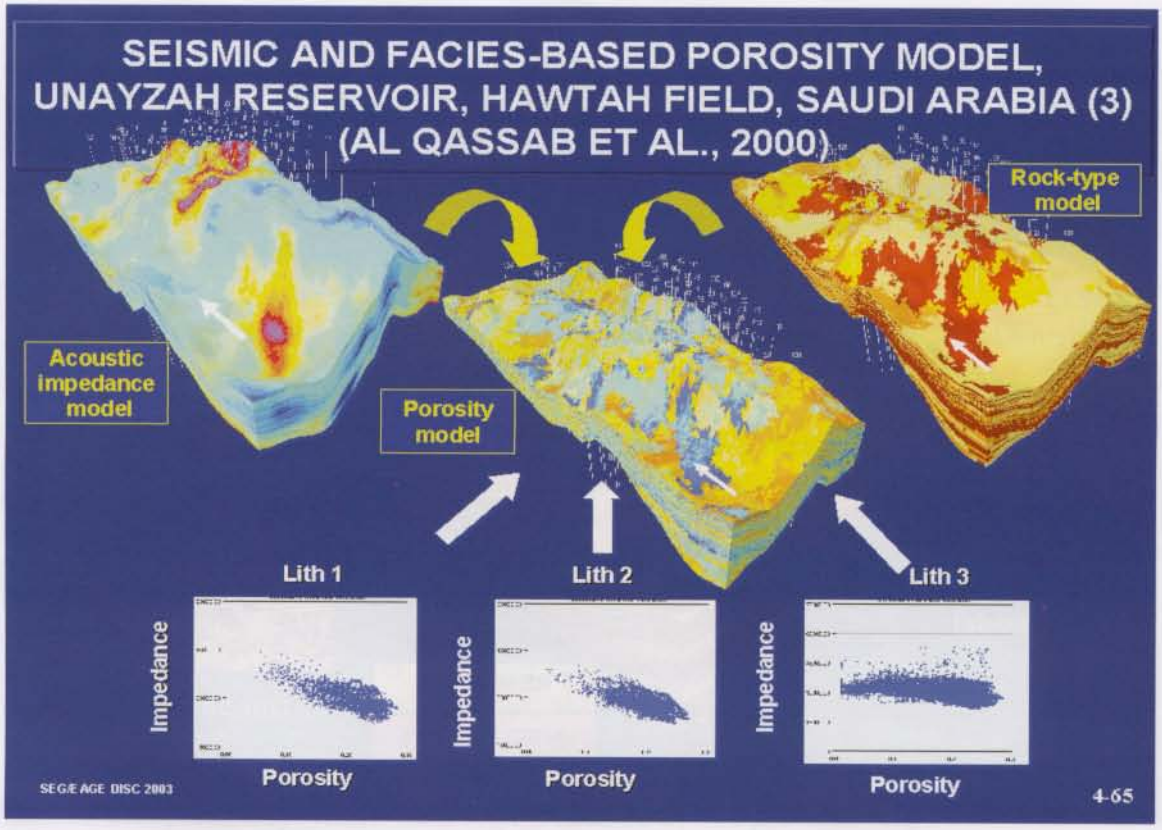
Seismic acoustic-impedance information was also available. It resulted from a stochastic inversion exercise (we will discuss stochastic inversion later on). Crossplots (Fig. 4-65) showed that the better the rock type in terms of reservoir quality, the better the linear relationship between porosity and acoustic impedance. Within each rock type, as determined by the 3D rock-type realization, it was thus possible to model porosity by collocated cosimulation using the acoustic-impedance value at each sampled location and the correlation coefficient corresponding to the rock type. The last stage consisted of simulating permeability (Fig. 4-66) under the control of the simulated porosity values within each facies.

The geostatistical realizations obtained by Al Qassab et al. were used as input to flow simulation, and provided a much better history match than would flow simulations run on conventional (smoothly interpolated) models of porosity and permeability.

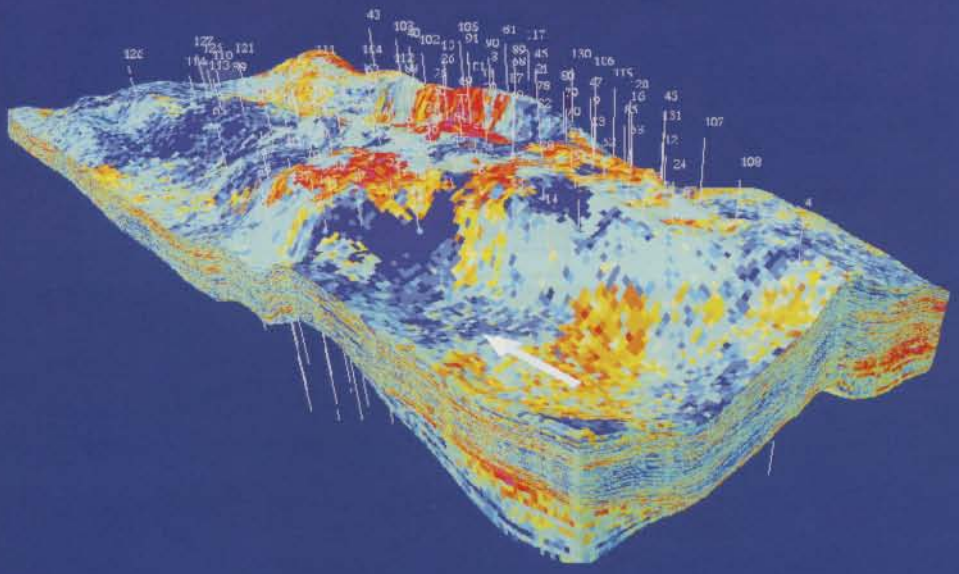
- The heterogeneity modeling toolkit

The previous example shows what must be available from a heterogeneity-modeling toolkit (Fig. 4-67). Fig. 4-68 takes the classification of reservoir architectures proposed by Weber and van Geuns (1990) and assigns to this classification various geostatistical simulation techniques that seem most appropriate for each situation.

The old "garbage in, garbage out" expression applies to everything we have seen about 3D heterogeneity modeling. All these geostatistical simulation techniques need some quantified geological information as input. In the last 20 years, the industry has



POROSITY-DERIVED PERMEABILITY MODEL, UNAYZAH RESERVOIR, HAWTAH FIELD, SAUDI ARABIA (4) (AL QASSAB ET AL., 2000)



SEG/E AGE DISC 2003

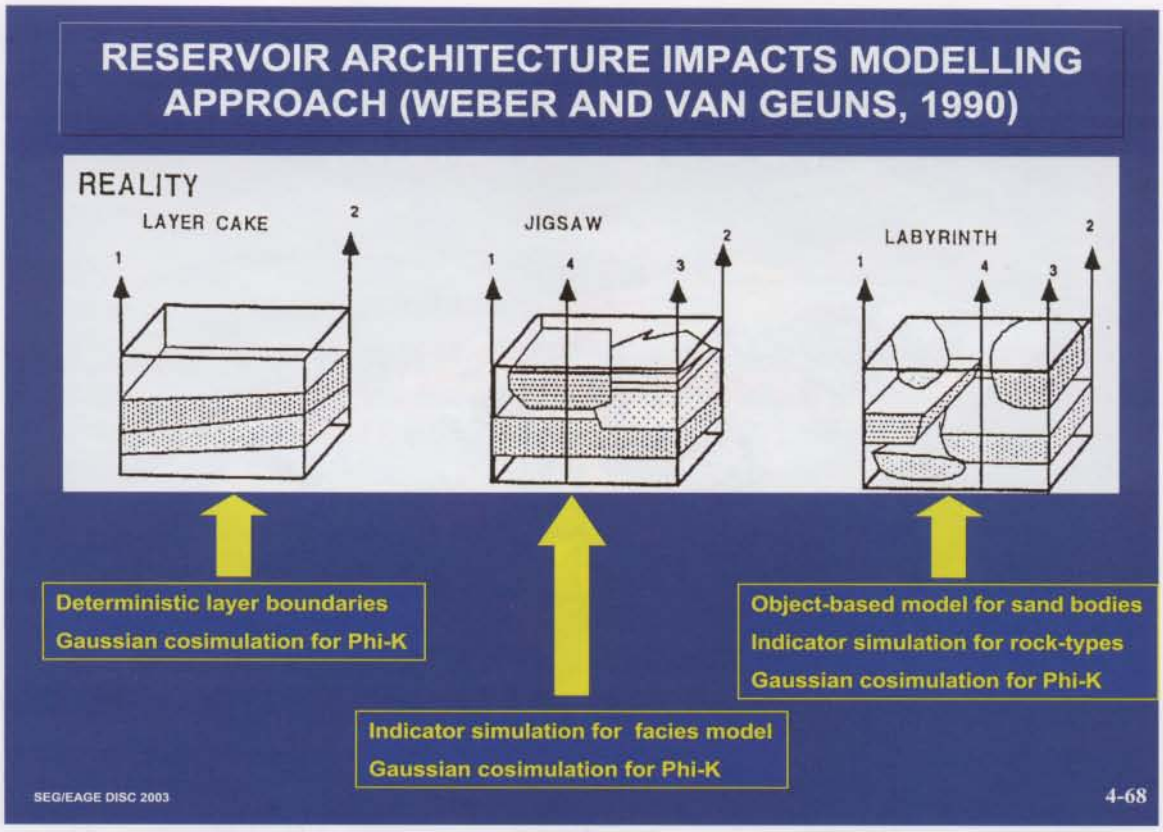
4-66

TOOLKIT NEEDED FOR GEOSTATISTICAL MODELLING

- Depositional environment modelling
- Rock-type or lithology modelling
- Joint petrophysical property modelling

SEG/EAGE DISC 2003

4-67

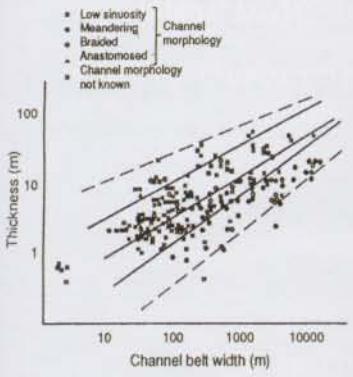


made a significant effort to collect quantitative information at all scales. Examples and references of such collection are given in Fig. 4-69.

A good criterion for deciding whether to use geostatistical heterogeneity modeling is to ask the geologist whether he or she would be able to hand-draw a representation of the geological model. If this is not the case, there is probably not enough information to build a 3D geological model, whether hand-drawn or generated by geostatistics. This is a very important point, stressing that geostatistics is not a substitute for geological knowledge, but rather a tool for quantifying this knowledge.

EXAMPLE OF QUANTITATIVE GEOLOGICAL INFORMATION AT DIFFERENT SCALES

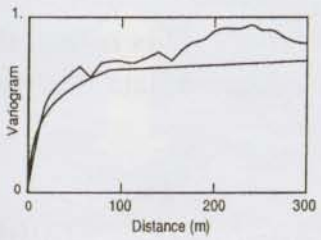
SIZE OF GENETIC UNITS



WIDTH/THICKNESS RELATIONSHIPS
VARIOUS OUTCROPS
(Bryant and Flint, 1993)

KILOMETRIC TO HECTOMETRIC SCALE

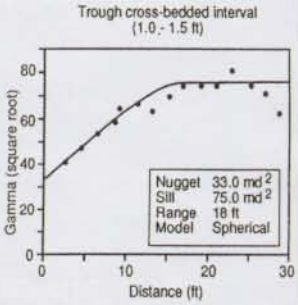
LITHOLOGY INDICATOR VARIOGRAM



SANDSTONE LITHOLOGY VARIOGRAM
MIDDLE JURASSIC. YORKSHIRE
(Ravenne at al., 1988)

HECTOMETRIC TO DECAMETRIC SCALE

PERMEABILITY VARIOGRAM



PERMEABILITY VARIOGRAM
THE FERRON SANDSTONE
(Tyler et al. 1991)

DECAMETRIC TO METRIC SCALE

SEG/EAGE DISC 2003

4-69

Geostatistics, Inverse of the Covariance, and Filtering

Let us discuss what Claerbout and Brown (1999) define as the geoestimation problem. Using their terminology, we deal with data, d, and a roughening filter, F.

Clarbout defines the Prediction Error Filter (PEF) as the operator F, such that, if it is applied to the data, the white noise is obtained ("PEF output is white"). This means that the autocorrelation of the filtered data is a white noise:

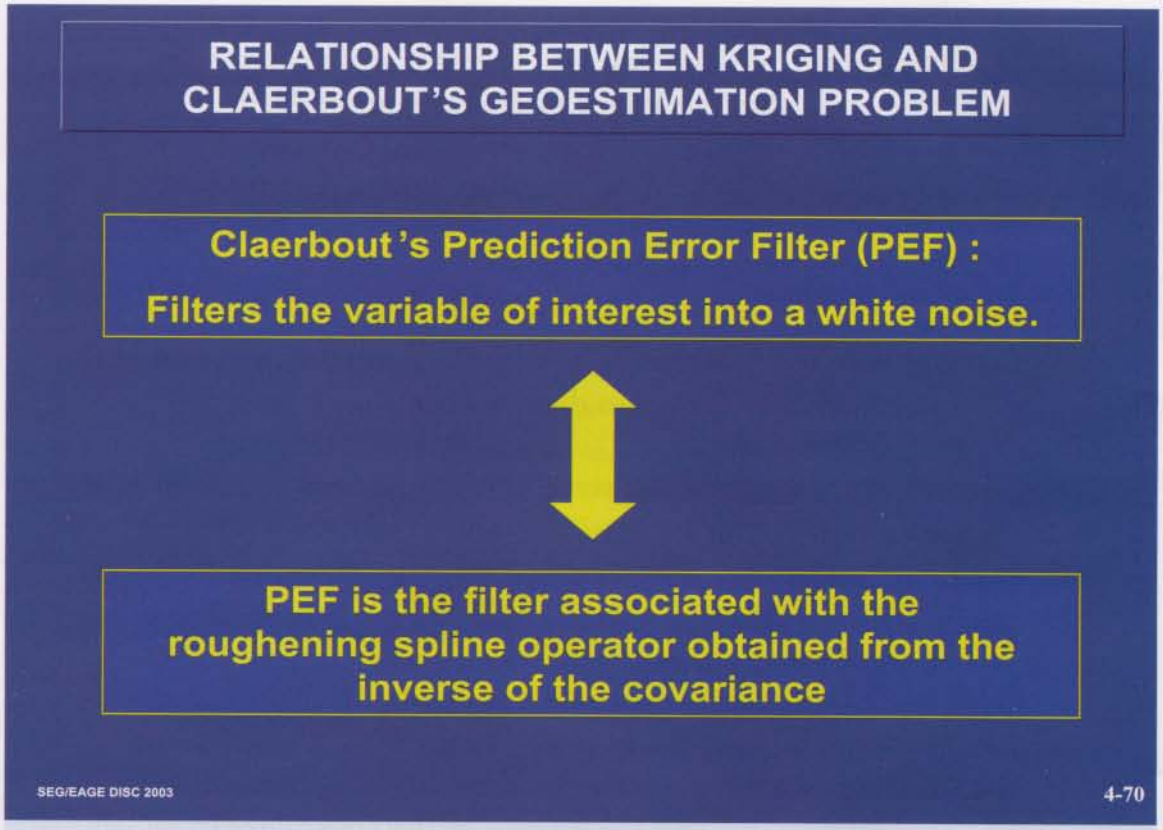
$$Fd * (Fd)' = \delta$$

This implies that the spectrum of F is the inverse of the data spectrum, or that F'F is the inverse of the covariance.

Thus, using our terminology and the results of Section 3.8.3 and Fig. 3-93, we can anticipate that our equivalent to Claerbout's PEF will be the spline operator associated with the covariance of the data (Fig. 4-70)! Claerbout (2002) mentions that "the PEF plays the role of the so-called inverse-covariance matrix in statistical estimation theory," which agrees with the relationship of Fig. 3-93.

Even more interestingly, Claerbout (2002) explains that "of all the assumptions we could make to fill empty bins, one that people usually find easiest to agree with is that the spectrum should be the same in the empty-bin regions as where bins are filled." This is another way of justifying the use of geostatistical simulation for generating representative samples of a spatial parameter. Claerbout shows that representative samples of such a model can be obtained by applying 1/F to a white noise, U, which is easy to prove considering that the spectrum of F is the inverse of the data spectrum.

Claerbout and Brown (1999) show that a variety of synthetic images can be produced with this approach, which is very similar to the moving-average method of Oliver (1995). Oliver applies the square root of the covariance function to a white noise. In the previous development, 1/F plays this role.



The work of Kane et al. (2001) discusses the analogy between deconvolution and kriging. Their results agree with the previous ones because, as mentioned by Claerbout, "PEF is also called deconvolution."

The previous discussion may be a bit confusing, because our geostatistical developments have been based on the assumption that $z(\mathbf{x})$ was a function of $\mathbf{x}$. In practice, we deal with regular grids of interpolated or simulated values, and geostatistics can also be expressed using a discrete formalism similar to that of filtering theory (Matheron, 1981b).

Assume we calculate kriging on a regular grid, and that all values of this grid are called z_u. The data points z_i are assumed to be among these grid points. In the zero-mean case, if C_{uv} is

the covariance between two grid points u and v, Matheron showed that the kriged grid values minimize:

$$\sum_{u,v} z_u C_{uv}^{-1} z_v$$

under the constraint that, at grid points where a data point i is present, the value z_i is honored. We recognize here the familar expression of the multivariate Gaussian distribution (Fig. 1-38).

Thus, thanks to the use of the inverse covariance, kriging also appears to minimize an energy-like expression. From the filtering point of view, the interpretation is that, after specified filtering (as expressed by the inverse of the covariance function), the data have minimum energy. When the discrete version of splines is used (Briggs, 1974) the quadratic form that is minimized is similar to that minimized by kriging. Similar comments can be made about the properties of PEF. Again and again, we see this duality between the spline operator and the inverse of the covariance.

In the case where data values y_i are affected by measurement errors, error cokriging minimizes this time:

$$\sum_{u,v} z_u C_{uv}^{-1} z_v + \theta \sum_i \frac{(z_i - y_i)^2}{\sigma_i^2}$$

The first term can also be interpreted as the regularization term of the energy function traditionally used in inverse problems. In the regularization context, the expression of the quadratic form is driven by smoothing considerations. Kimeldorf and Wahba (1970) argue that the choice of the regularization operator is merely driven by computational convenience. On the other hand, geostatisticians implicitly derive this operator from the a priori geological knowledge, as quantified by the covariance function.

5 Geostatistical Inversion

5.1 Introduction

In the previous chapter, we saw how geostatistics can be used to generate 3D heterogeneity models satisfying a number of input statistical parameters, such as the variogram, or equivalently, the spectral density. Relationships between simulated parameters and seismic-derived information were statistical, usually coded as a linear correlation between the seismic parameter and the simulated variable. This was the "primary- versus secondary-variable" approach. We also saw earlier that the geostatistical paradigm, based on the concept of trend and covariance, can be considered to be an approach for coding the a priori model that is often used in Bayesian terminology.

In parallel with geostatistical developments, inversion technology made significant progress in the last three decades (Fig. 5-1). Thanks to the Bayesian formalism promoted by authors such as Tarantola (1987) and Duijndam (1988), the standard optimization-based deterministic approach (minimization of an objective function with a regularization term) was improved, and it became possible to produce an estimate of uncertainty together with the inverted acoustic-impedance block.

Logically, the idea emerged in the early 1990s to apply the conditional simulation approach to acoustic-impedance inversion to produce multiple 3D realizations, all constrained by seismic data. This resulted in geostatistical-inversion methodology. In the following, we will consider that geostatistical and stochastic inversion are exactly the same thing.

5.2 Basics of Geostatistical Inversion

The method discussed here is that presented in the papers of Bortoli et al. (1992) and Haas and Dubrule (1994). Geostatistical inversion (GI) consists of generating 3D acoustic-impedance realizations, all constrained by seismic data. The input to a geostatistical-inversion study is similar to that of any conventional simulation study, with the important addition of a 3D seismic block (Fig. 5-2). Bortoli et al. asked the question: How, using SGS, can we make sure that not only the well data and the variogram are honored, but also the 3D seismic block? They proposed the solution described in Fig. 5-3, which specifically addresses a well-known inversion issue: how to make sure that every single seismic trace is inverted, while preserving the lateral continuity between inverted traces (that is, while avoiding the limitations of a single trace algorithm).

The algorithm is just an extension of SGS. At each sampled location, a large number of — say, 100 — local acoustic-impedance trace realizations are produced. These traces are all convolved with the seismic wavelet, and the acoustic-impedance trace providing the best match with the seismic trace at that location is selected. Then the algorithm moves to another location. When the whole space is filled with traces, we have obtained one global realization. Of course, more than one global realization (say, 100) is produced.

The algorithm is illustrated using a North Sea example. First, an acoustic-impedance global realization was generated using only one local realization at each sampled

EVOLUTION IN DATA INTEGRATION (BOTH WITH SEISMIC AND PRODUCTION DATA)

Classical inversion
deterministic optimization with regularization term

One model

Calculation of a posteriori model under multi-Gaussian assumptions

One model and associated uncertainty model

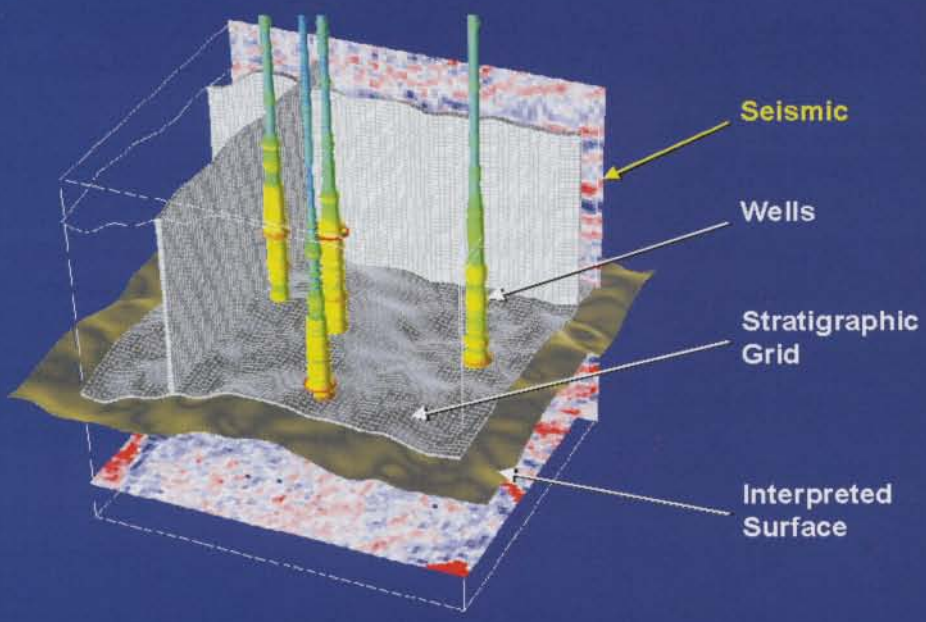
Stochastic inversion: geostatistical inversion, MCMC

Multiple realizations

SEG/EAGE DISC 2003

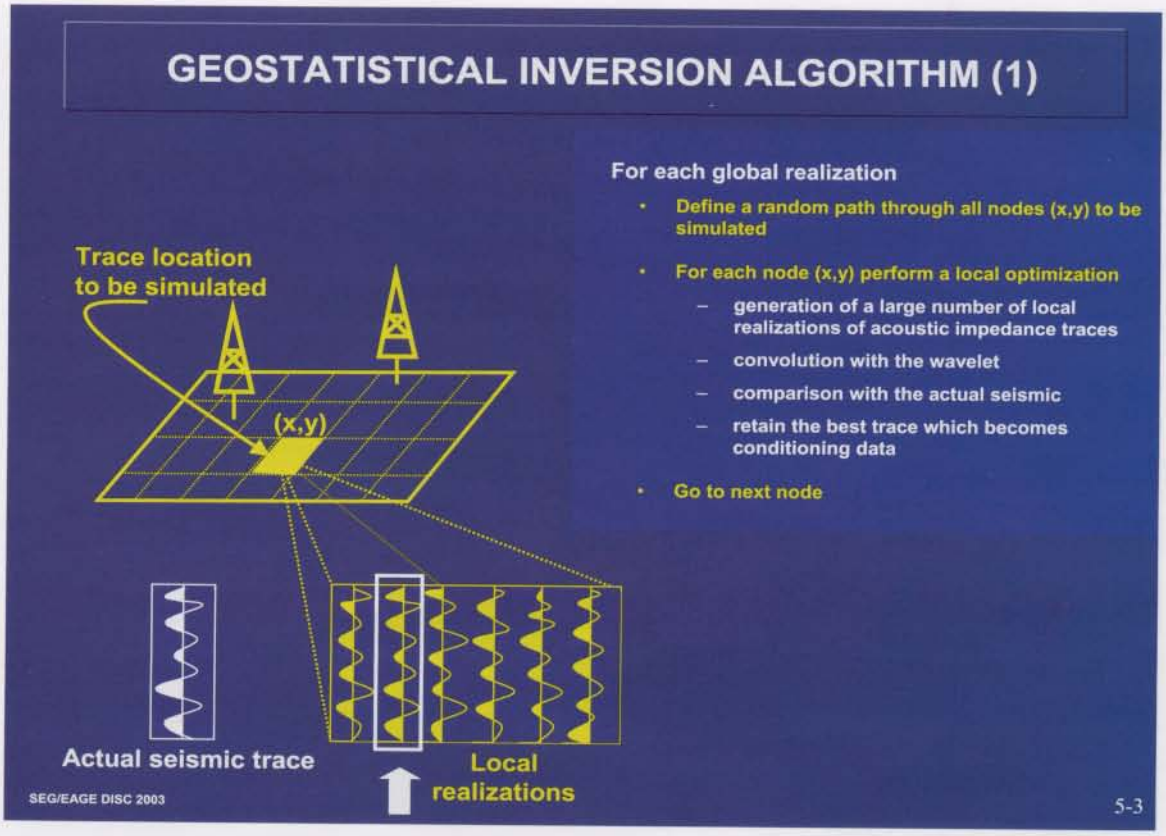
5-1

INPUT TO GEOSTATISTICAL INVERSION STUDY



SEG/E AGE DISC 2003

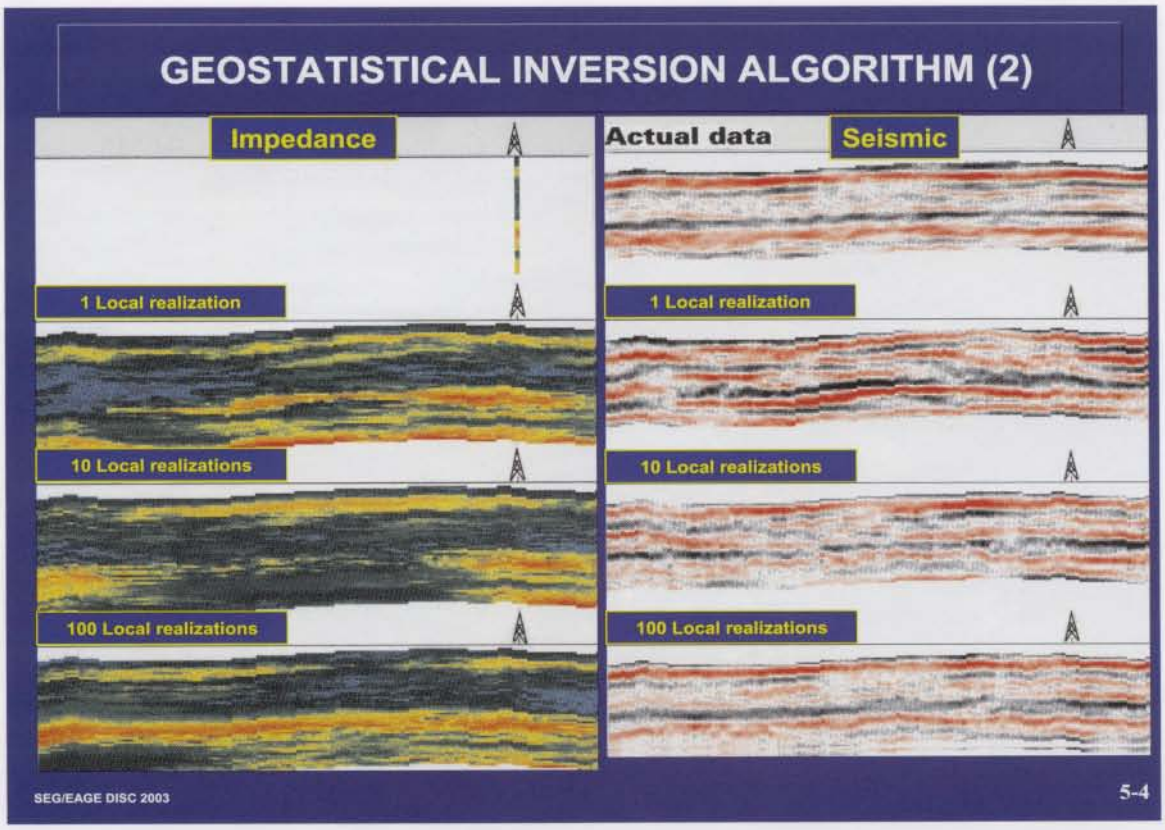
5-2

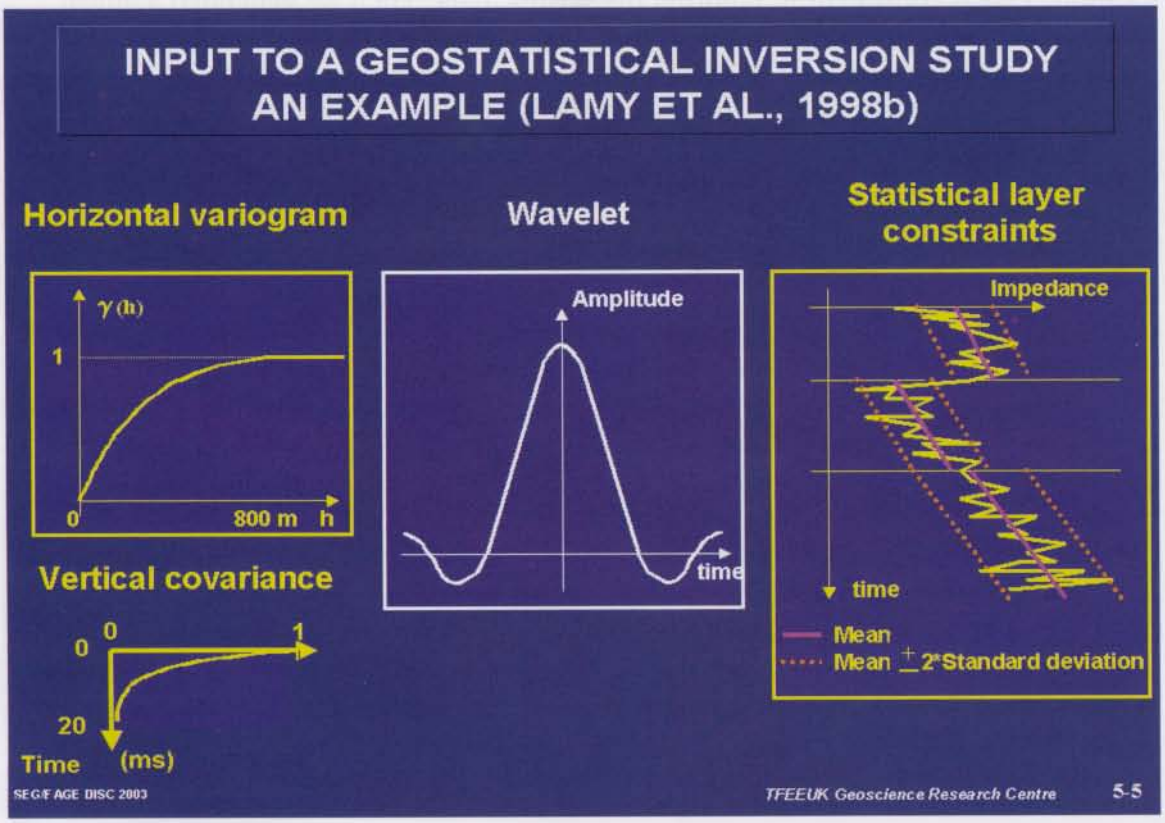


location. This is equivalent to ignoring the seismic data. Then we generated a global realization where only 10 local realizations were sampled at each location, and we compared the result with that obtained using 100 local realizations (Fig. 5-4). The seismic match improves significantly as the number of local realizations increases.

This match can be controlled in a number of ways. One is to increase the number of local realizations at each SGS step. Another is to keep generating local realizations until a certain threshold of the local objective function is reached. There is great flexibility in the choice of this objective function, which can be the absolute error, the mean squared error, or the correlation coefficient. It can also be a combination of those three, or it can vary with the location of the simulated trace. Common sense must be applied when matching a synthetic from a simulated trace with the actual seismic data. Often, well-to-seismic calibration does not produce very high correlation coefficients between actual and synthetic seismic. Between the wells, it would be computationally possible, but it would not make sense to ask for correlation coefficients higher than those obtained at the wells themselves!

A number of input statistical parameters must be fitted to the well and seismic data and used as input to GI, in order to constrain the different realizations (Fig. 5-5). These parameters, which are usually layer-dependent, are the same parameters as are in a standard geostatistical simulation study, plus the wavelet. They may be considered as the a priori geostatistical model. Fig. 5-6 shows another example of convergence of the algorithm, this time with layer-dependent parameters, and a display of the residuals between synthetic and actual seismic.


GEOSTATISTICAL INVERSION ALGORITHM (2)
Impedance
Actual data
Seismic
1 Local realization
1 Local realization
10 Local realizations
10 Local realizations
100 Local realizations
100 Local realizations
SEG/EAGE DISC 2003
5-4


INPUT TO A GEOSTATISTICAL INVERSION STUDY
AN EXAMPLE (LAMY ET AL., 1998b)
Horizontal variogram
γ (h)
1
0
800 m h
Wavelet
Amplitude
time
Statistical layer constraints
Impedance
time
Mean
Mean ± 2*Standard deviation
Vertical covariance
0
0
1
20
Time (ms)
SEG/EAGE DISC 2003
TFEEUK Geoscience Research Centre
5-5

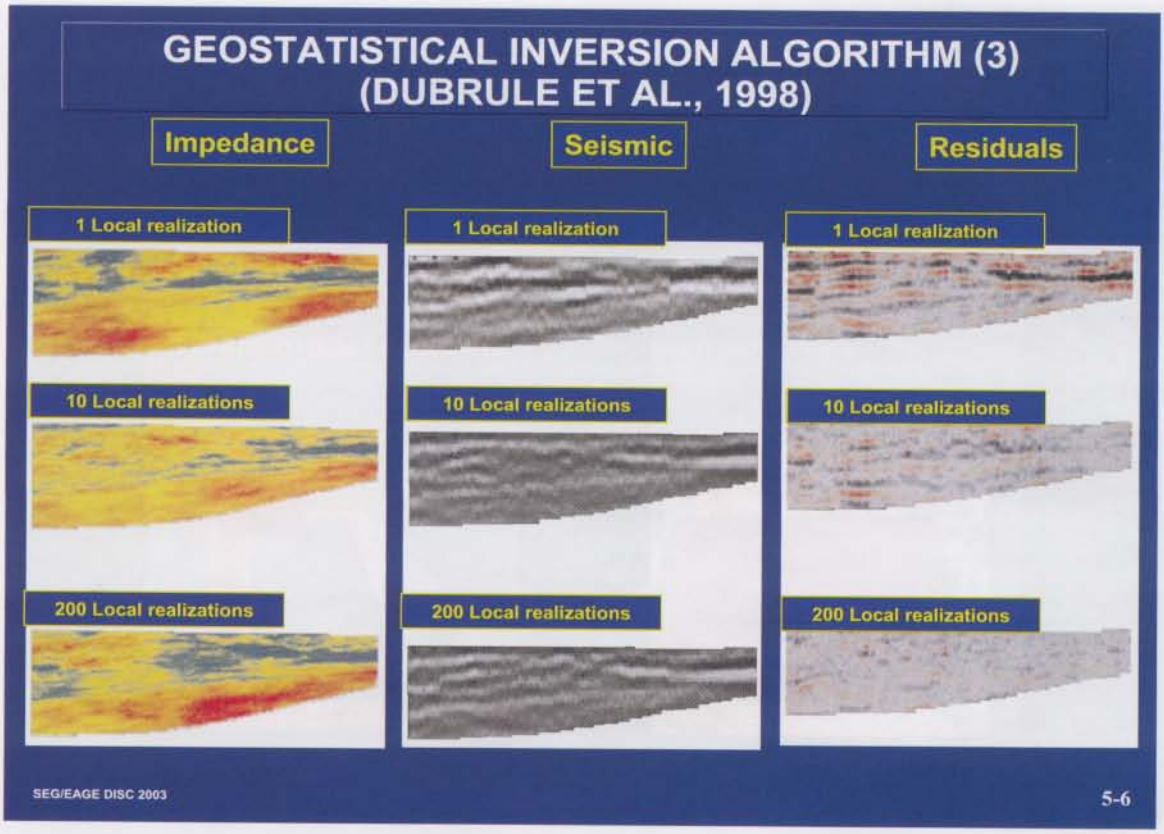


Fig. 5.7 compares two seismic slices from a North Sea example: one slice is through the actual seismic data, the other one is through the synthetic seismic block associated with the results of geostatistical inversion. The match is very good, although the synthetic block exhibits too much continuity because of the lateral correlation forced by the input horizontal variogram model.

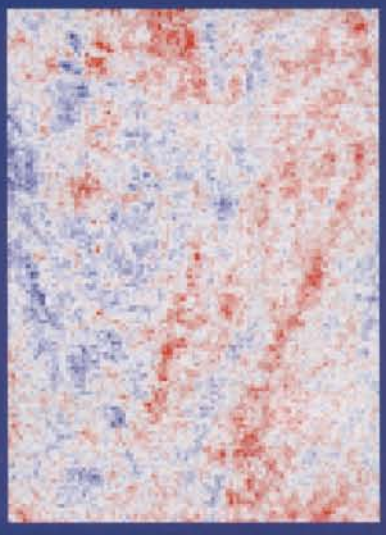
GI allows the generation of a large number of acoustic-impedance realizations honoring the 3D seismic data (see Fig. 5-8 for a North Sea example). Uncertainty is quantified by the variability from one realization to another. Why is there such uncertainty affecting the results of geostatistical inversion, considering all the input constraints that are injected into the a priori model? This is because GI is typically run on stratigraphic grids composed of about 2-ms-thick individual grid cells, that is, on a thickness that cannot be resolved by standard 3D seismic data sets. Because the seismic data can only control impedance variations within the seismic bandwidth, the higher frequency variations remain nonunique and vary from one realization to another. These higher frequencies are precisely controlled by the vertical variogram model and the frequencies it carries (see Fig. 5-9). On the other hand, lower frequencies are controlled by statistical constraints (mean and standard deviation) on the acoustic impedance distribution within each layer of the stratigraphic grid.

An important point to stress is that, because of the use of SGS, all realizations honor the well data (most standard acoustic-impedance inversion approaches do not honor the well data).

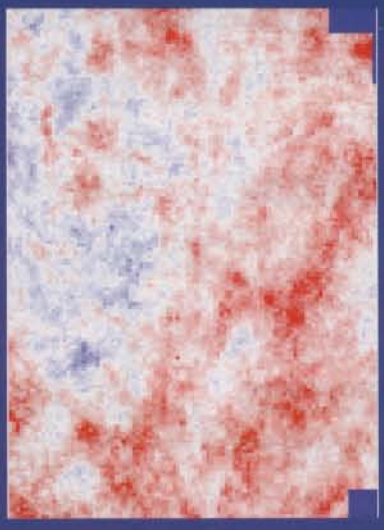
The lateral variogram applies to variations within the stratigraphic grid and is usu-

COMPARING TWO SEISMIC SLICES (LAMY ET AL., 1998b)

ACTUAL



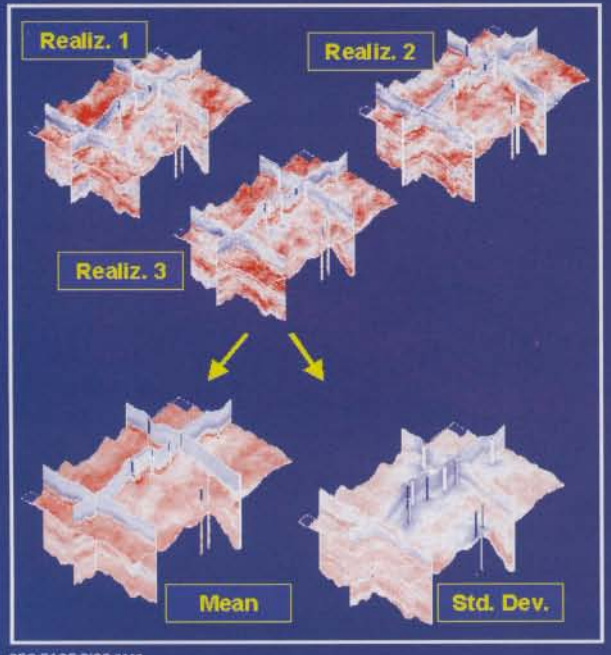
SYNTHETIC



SEG/EAGE DISC 2003

TFEEUK Geoscience Research Centre 5-7

MAIN FEATURES OF GEOSTATISTICAL INVERSION (LAMY ET AL., 1999)



Impedance realizations honor
- seismic data
- well data

Uncertainty on impedance results

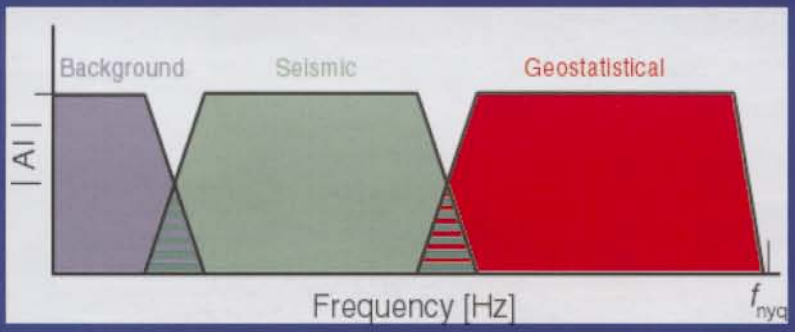
High vertical frequency of impedance realizations controlled by variogram model

SEG/EAGE DISC 2003

TFEEUK Geoscience Research Centre 5-8

SCHEMATIC VERTICAL FREQUENCIES CONTROL IN GEOSTATISTICAL INVERSION

- **Statistical layer constraints (background) control low frequencies.**
- **Seismic amplitudes control medium frequencies.**
- **Variogram model controls high frequencies.**



SEG/EAGE DISC 2003 L. Barens & E. Robein, TFEEUK Geoscience Research Centre 5-9

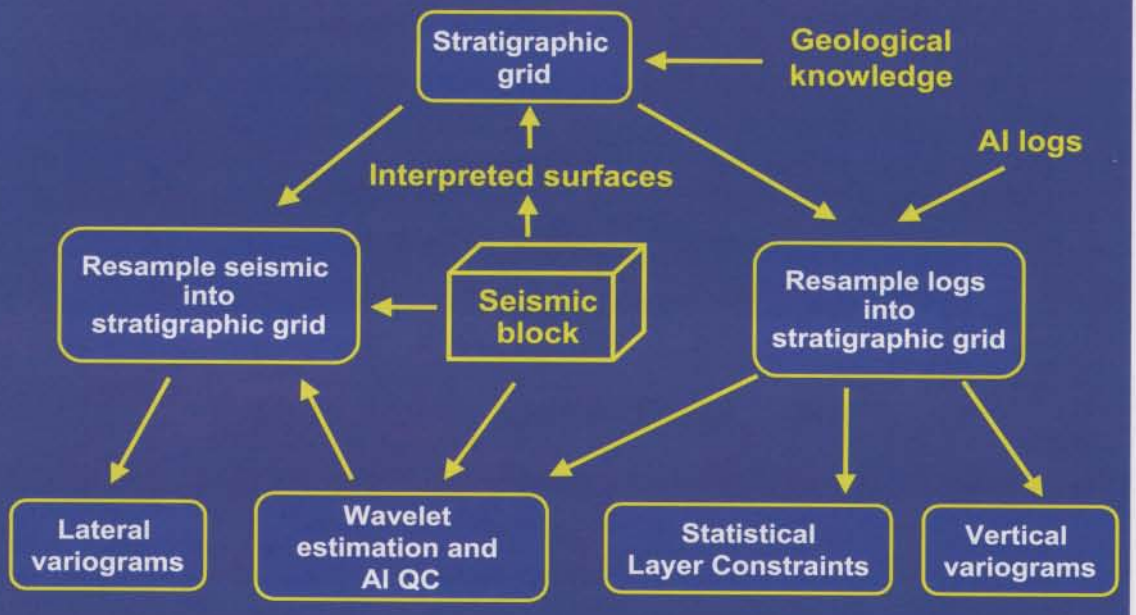
ally derived from both seismic information and well data, using considerations as described in Fig. 2-39.

Obviously, GI is not reduced to an algorithm, but requires a detailed workflow that is described in Fig. 5-10. In this figure, input data and information (including a priori geological knowledge) are indicated in yellow, whereas interpretation processes are in white.

Fig. 5-11 shows a cross-section resulting from a standard GI study (North Sea). The wells are honored by GI, and the lateral variogram ensures lateral correlation between traces. This implies that the mean of the realizations is continuous and that their standard deviation is zero at the wells (at least if there is no nugget effect in the variogram model; see Section 3.3).

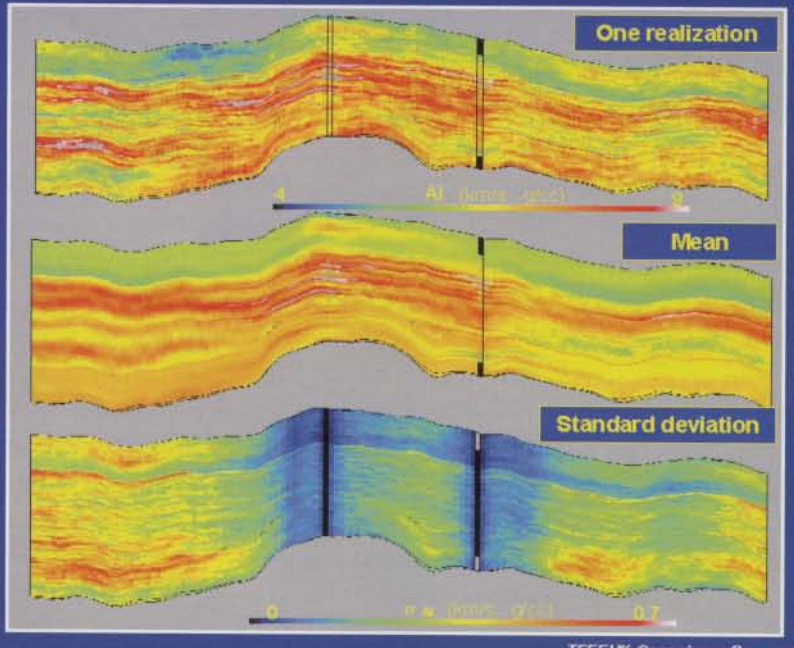
To help understand how different kinds of data are used to constrain the models, Fig. 5-12 compares the average 3D blocks obtained in four situations. We are using the same example as that used to introduce conditional simulations, in Chapter 4. Fig. 5-13 shows a view of the actual seismic data together with a synthetic block computed from one of the inverted acoustic-impedance realizations. If the realizations in Fig. 5-12 are conditioned neither by wells nor by seismic, the average block is simply an image of the input statistical constraints (mean and standard deviation) that were used as the input to the model. If only the wells are used to constrain the realizations, the average is simply equal to kriging from the wells only. The two other average blocks are constrained by seismic. Because of a compensation effect from one realization to another, the high-frequency variations are smoothed out, and the information contained in the mean

GEOSTATISTICAL INVERSION PRELIMINARY WORKFLOW (ROWBOTHAM ET AL., 2000)

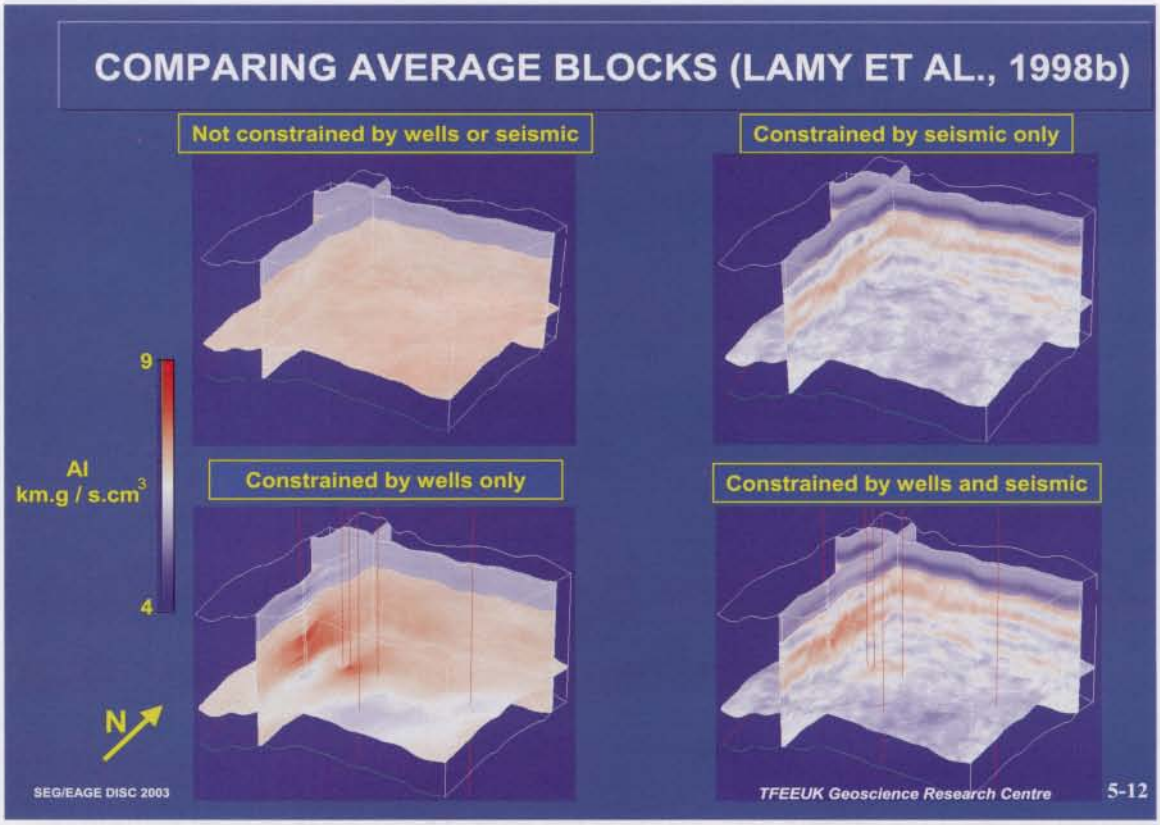


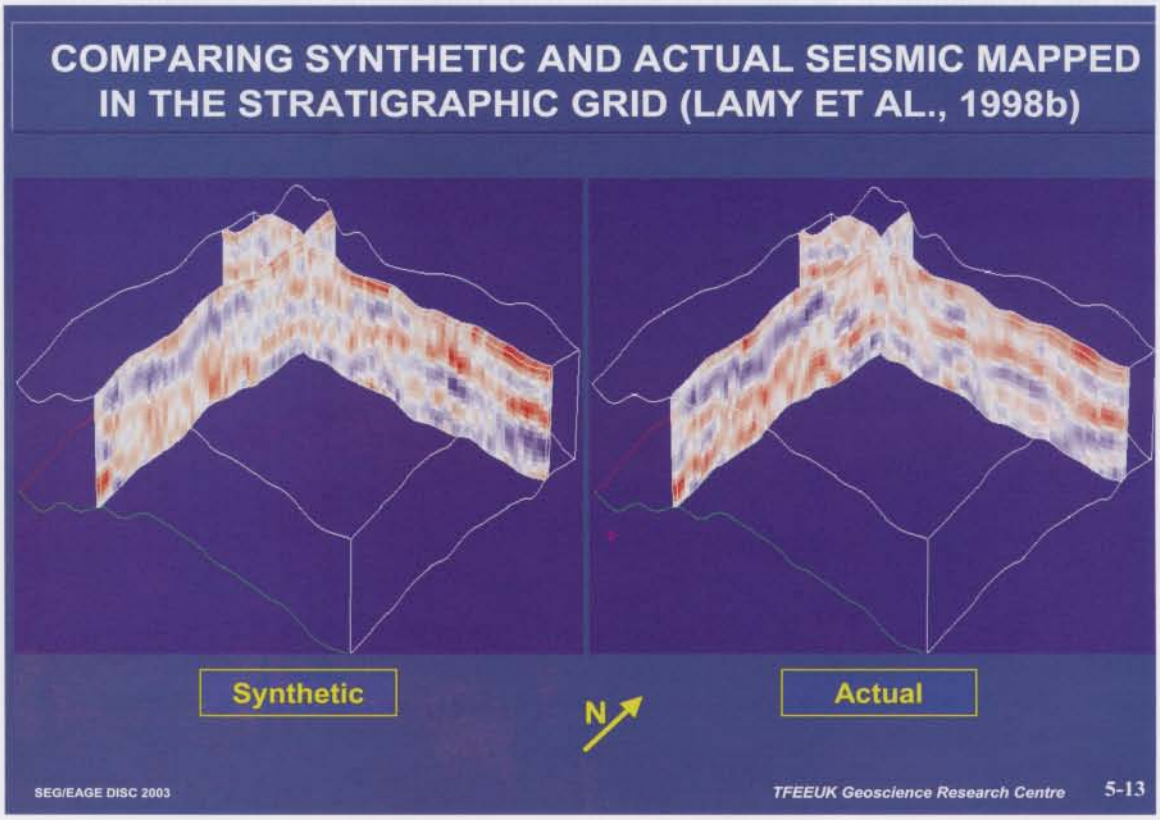
SEG/EAGE DISC 2003 TFEEUK Geoscience Research Centre 5-10

EXAMPLE OF GEOSTATISTICAL INVERSION RESULTS (LAMY ET AL., 1999)



SEG/E AGE DISC 2003 TFEEUK Geoscience Research Centre 5-11


COMPARING AVERAGE BLOCKS (LAMY ET AL., 1998b)
Not constrained by wells or seismic
Constrained by seismic only
Constrained by wells only
Constrained by wells and seismic
9
AI
km.g / s.cm3
4
N
SEG/EAGE DISC 2003
TFEEUK Geoscience Research Centre
5-12


COMPARING SYNTHETIC AND ACTUAL SEISMIC MAPPED IN THE STRATIGRAPHIC GRID (LAMY ET AL., 1998b)
Synthetic
N
Actual
SEG/EAGE DISC 2003
TFEEUK Geoscience Research Centre
5-13

comes from what is common to all realizations, that is, the seismic data control within their bandwidth. As expected, the difference between the two seismic-constrained blocks is limited to the area close to the wells.

GI generates a large number of realizations of acoustic-impedance blocks — say, 100. There is a real issue about the processing of these multiple realizations. How can this enormous amount of information be summarized? A first approach may consist of calculating the mean and the standard deviation of realizations. There may also be other interesting ways to process the data, such as that of Fig. 5-14. At each grid cell, the number of realizations above a certain threshold is counted and transformed into a probability. This can provide very useful information in situations where high or low impedances can be straightforwardly associated with the presence or absence of reservoir rocks.

5.3 Accounting for Faults

The GI algorithm is very flexible, because it is a local approach. Lamy et al. (1998a) and Rowbotham et al. (2000) took advantage of the flexibility of the algorithm, in the case of faulted reservoirs. The quality of seismic data close to the faults tends to be poorer than that away from these faults. Thus, the GI algorithm is first limited to the construction of acoustic-impedance traces away from the faults (Fig. 5-15). Once these traces have been generated, which are controlled by a better-quality seismic, traces closer to the faults are sampled. For these traces, the number of local simulations may be reduced to one if the seismic data quality is too poor. The number of local realizations can vary, depending on the quality of seismic data. This approach also guarantees that the traces away from the faults are optimal, because they are not affected by artifacts that may come from the fault traces.

Fig. 5-16 shows results obtained on a North Sea field by Rowbotham et al. (2000). The match between actual and synthetic seismic is of average quality, but the blind well test is rather satisfactory (Fig. 5-17), showing that the difference between the actual well log and the mean of the GI realizations is almost everywhere smaller than one standard deviation.

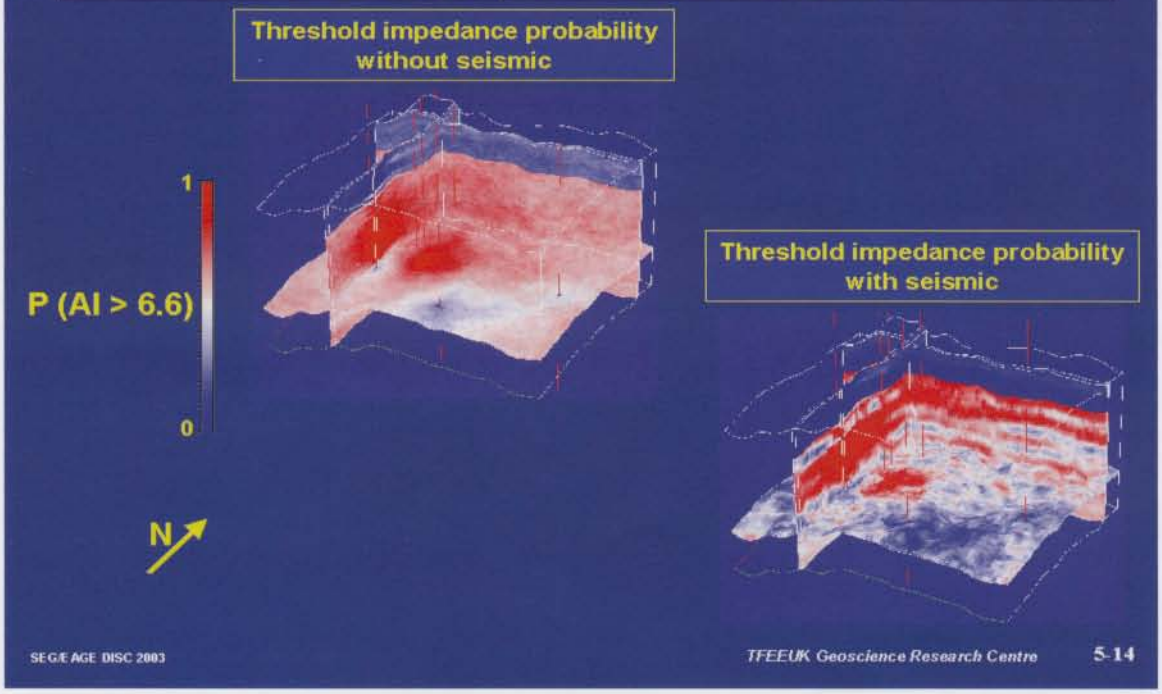
Shrestha and Boeckmann (2002) also take advantage of the flexibility of the geostatistical-inversion algorithm. Because the quality of seismic data is poor in areas influenced by salt, SGS is not permitted to generate more than one local realization in those areas, which means that the simulation is not constrained by seismic.

5.4 A Variety of Methods

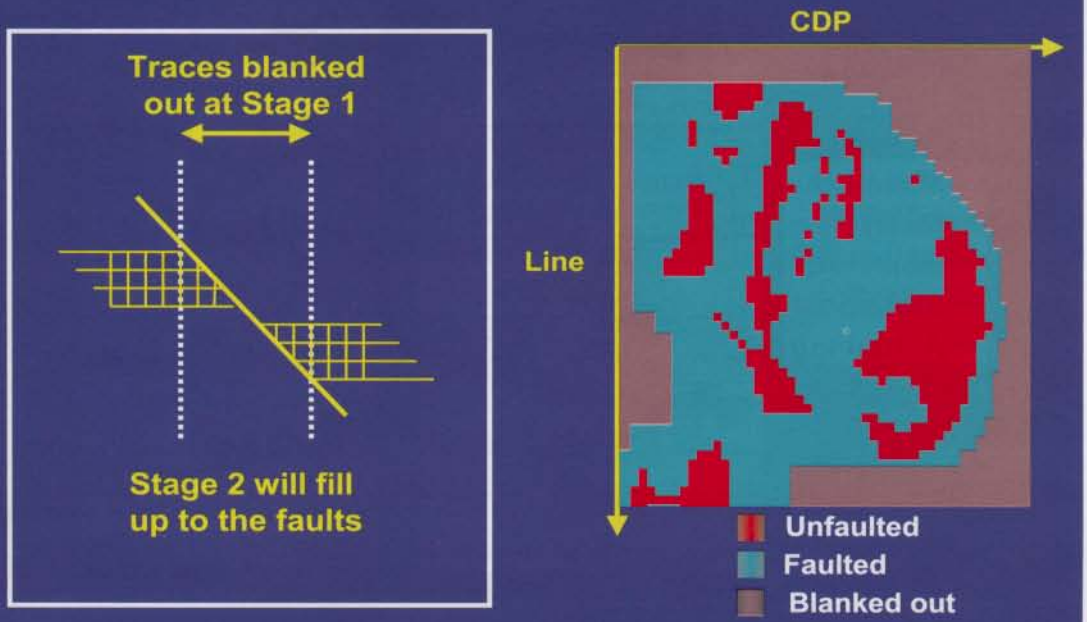
5.4.1 A different sampling algorithm

Following the publication of the first papers on GI, the method has been integrated into a number of commercial software packages, using a variety of algorithms. A popular one is that presented in Grijalba-Cuenca et al. (2000). The main difference it has with that discussed above lies in the fact that, instead of working trace by trace as the previous method does, Grijalba-Cuenca et al.'s method works grid cell by grid cell (Fig. 5-18). First, an initial realization of acoustic impedance is generated, based on well data.

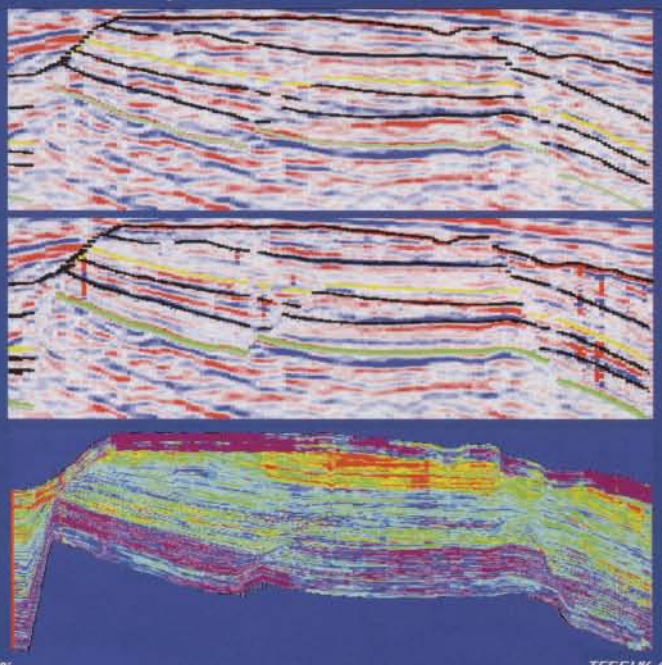
HOW TO SUMMARIZE GEOSTATISTICAL INVERSION RESULTS? (LAMY ET AL., 1998b)



DEALING WITH FAULTS IN GEOSTATISTICAL INVERSION (ROWBOTHAM ET AL., 2000)

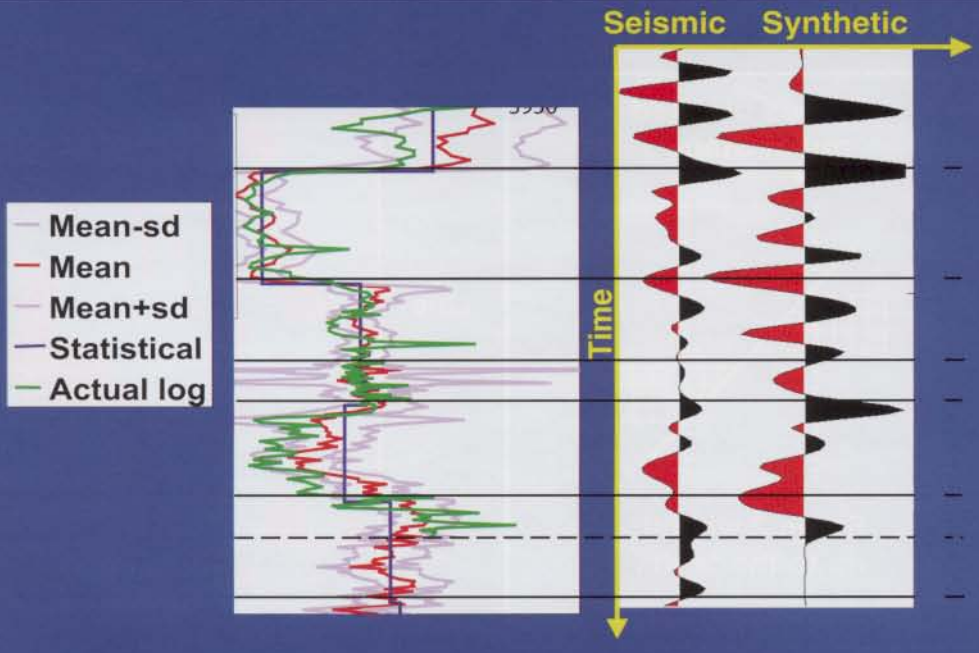


FAULTED GEOSTATISTICAL INVERSION EXAMPLE (ROWBOTHAM ET AL., 2000)

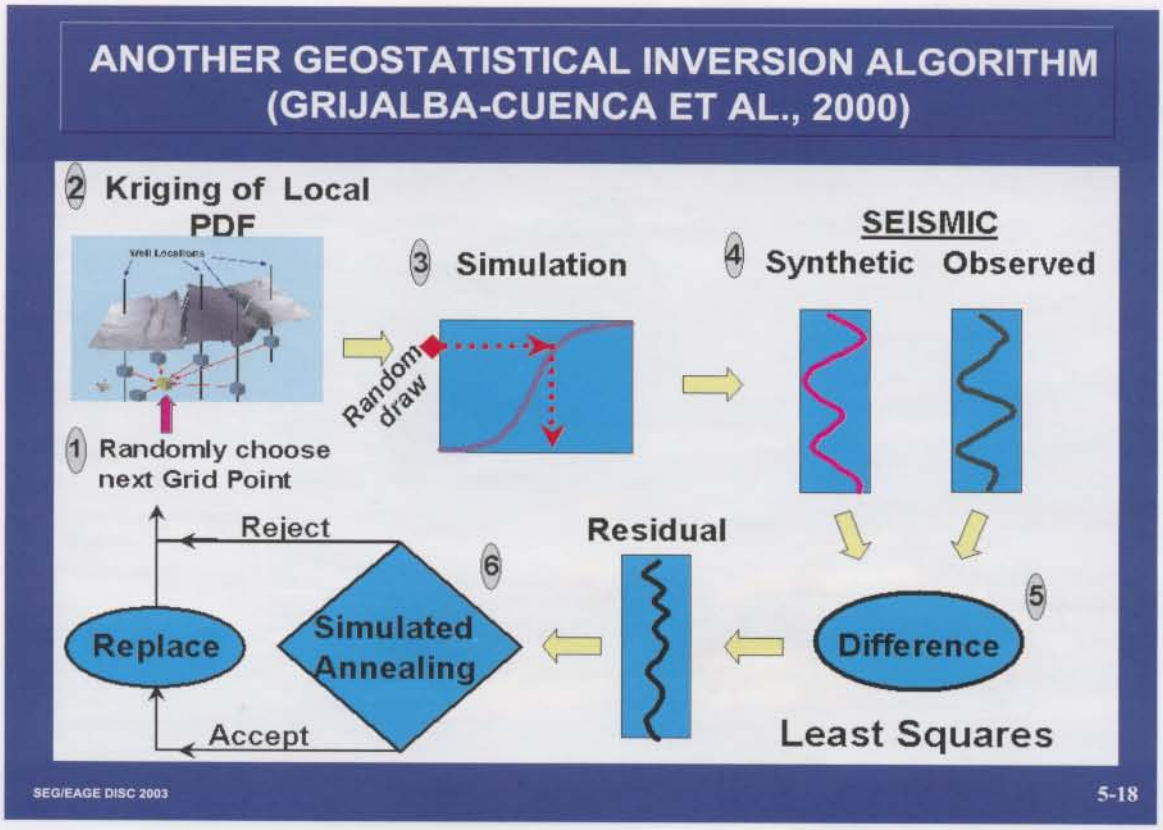


SEG/EAGE DISC 2003 TFEEUK Geoscience Research Centre 5-16

BLIND WELL RESULTS (ROWBOTHAM ET AL., 2000)



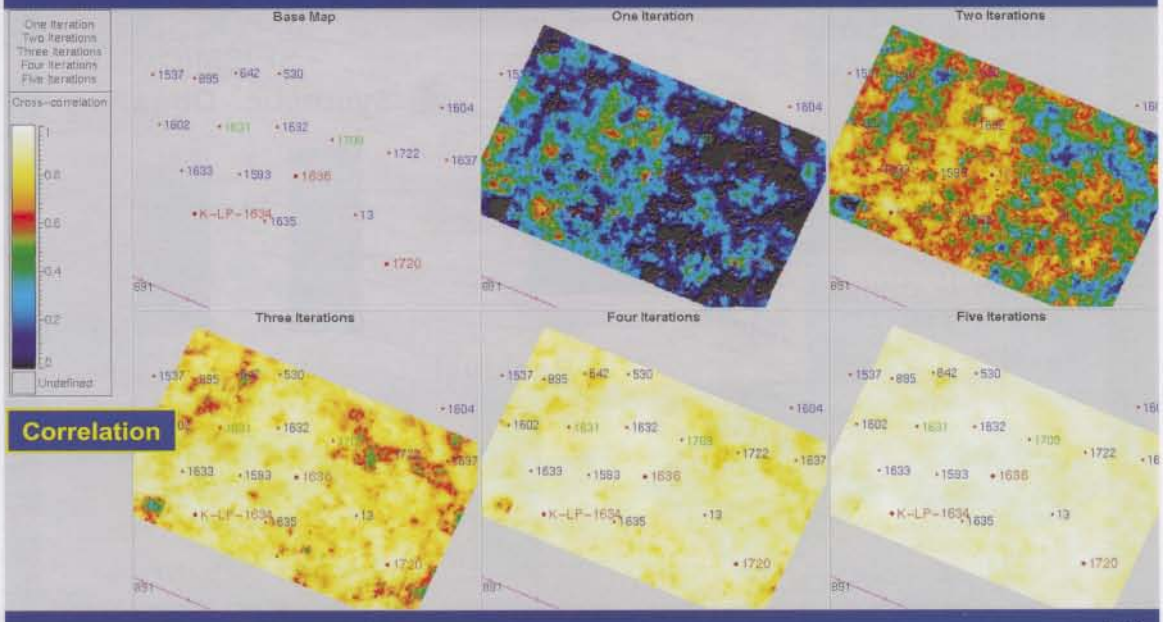
SEG/EAGE DISC 2003 TFEEUK Geoscience Research Centre 5-17



Then, each grid cell is sampled randomly, and each time, the new objective function measuring the match between synthetic and seismic trace is calculated. If it is small enough (according to a user-defined criterion), the search moves to another point. If not, the value is resampled using an acceptance criterion similar to that of Metropolis (Fig. 4-24). Fig. 5-19 shows, in an example from Argentina, how the map of the correlation coefficient between synthetic and actual seismic varies as a function of the number of global iterations (a global iteration corresponds to an iteration where all individual grid cells have been visited once). There is a computer time issue, because it takes more time to iterate on individual grid cells than on individual traces. In order to reduce computer time, Grijalba-Cuenca et al. do not restart from scratch for each 3D seismic-constrained realization. Instead, they generate the different realizations at each local simulation step, which amounts to using the same search path for all realizations. There is a risk, however, that not sampling the random path randomly may lead to realizations that do not span the actual range of uncertainty.

The method is quite popular and has had many successful applications. Fig. 5-20 shows a recent high-resolution geostatistical inversion performed by Torres-Verdin et al. (C. Torres-Verdin, et al., personal communication, 2002), who went as high as a 0.5-ms sampling rate on the inverted realizations (Fig. 5-21). The central well was used as a blind well to validate the method, whereas the two other wells were used as input data points. The three SP well logs are displayed on each cross-section, because the project was looking for sands that were below seismic resolution. Acoustic impedance was useful, in that high values of acoustic impedance correlated with clean sands (low volumes

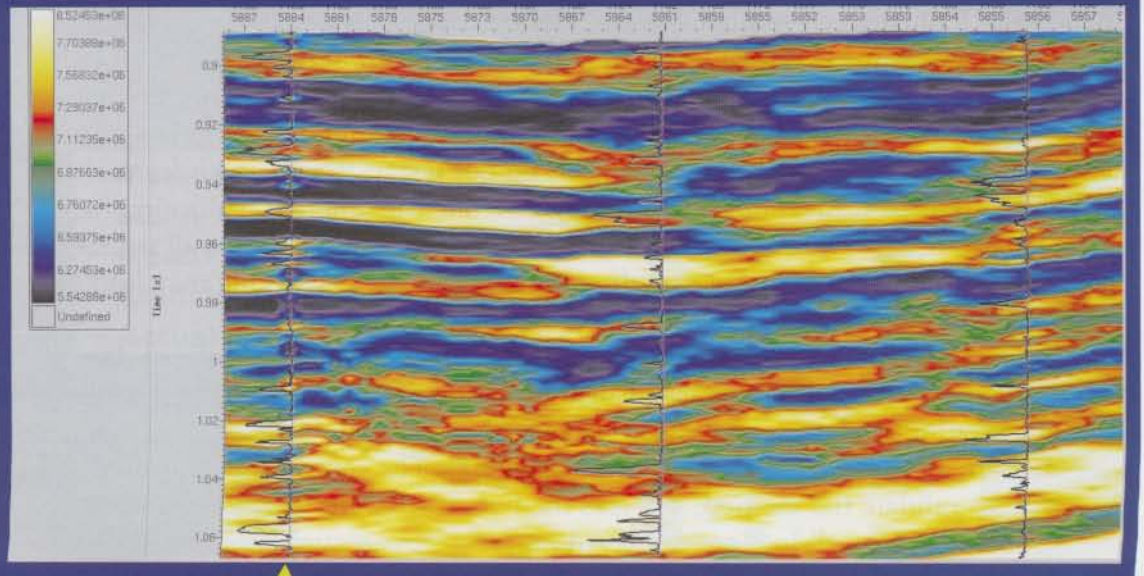
SEISMIC MATCH AS A FUNCTION OF GLOBAL ITERATIONS 1 MS GRID SIZE (TORRES-VERDIN ET AL., PERS. COMM., 2002)



SEG/EAGE DISC 2003

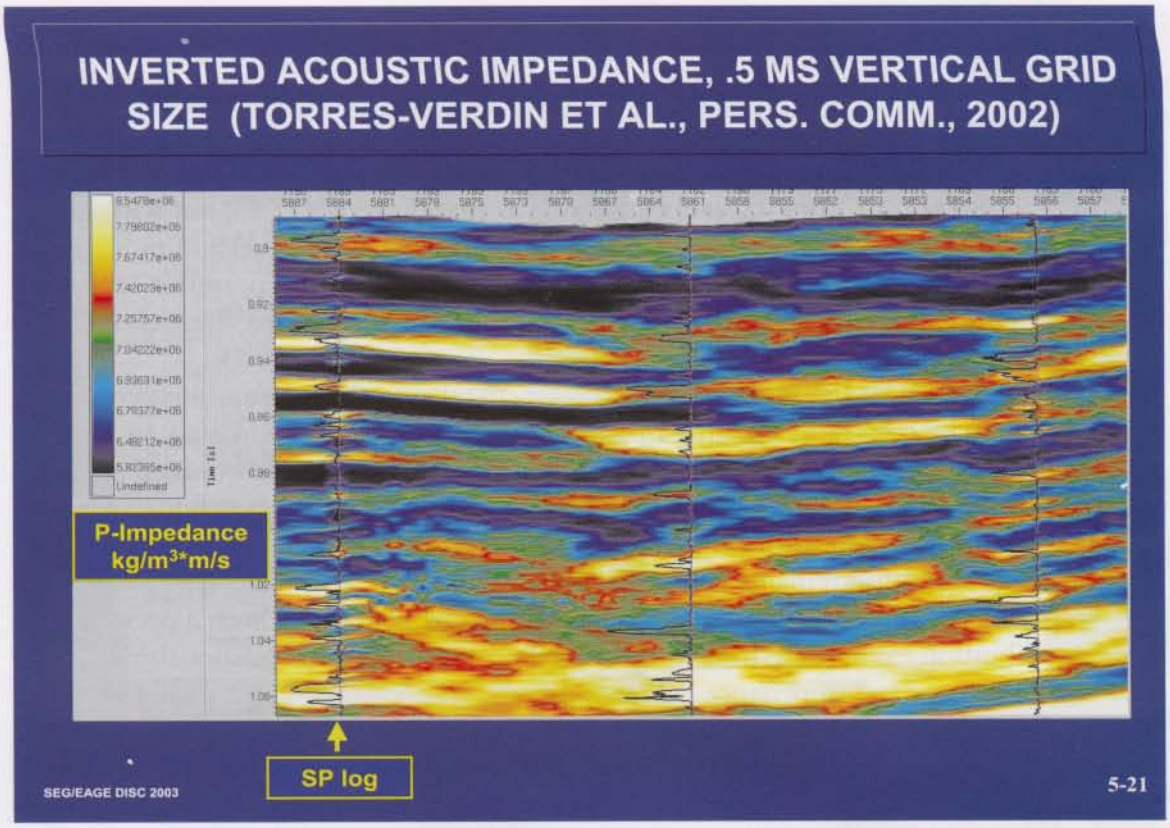
5-19

INVERTED ACOUSTIC IMPEDANCE, 1 MS VERTICAL GRID SIZE (TORRES-VERDIN ET AL., PERS. COMM.,2002)



SEG/EAGE DISC 2003

5-20



of shales). As expected, the match between the two input-well SP logs and the inverted section is excellent, but it is also very satisfactory at the blind well. At which vertical resolution should the inversion be performed? The higher the resolution, the more heterogeneities will be captured by the model, but the more uncertainty will be attached to these heterogeneities. A trade-off must be found by running flow simulations and choosing the resolution that is closest to that of the seismic but still provides realistic flow simulations. This is an area that demands more investigation.

5.4.2 *Geostatistical inversion based on fractals*

Gunning and Paterson (1999) use a similar approach to that of Bortoli et al. (1992), in that they sample the 3D volume trace-by-trace. But there are many other differences between the two approaches. First, Gunning and Paterson work with fractal models, which they claim to be more general than those used by Bortoli et al., because they are nonstationary. It can be objected that fractal models, if they are certainly nonstationary, are also quite specific and cannot be applied to all reservoirs. Gunning and Paterson also use, as do Grijalba-Cuenca et al., the same random path for each global realization, which saves a lot of time. A Cholesky decomposition and the subsequent inversion of the covariance matrix is performed for the first global realization only, then reused for other global realizations, leading to repeat simulations that are at least 100 times faster than the first. But, as mentioned for the method described in the previous section, this may also decrease the variability from one realization to another and thereby underestimate the uncertainties. As an objective function, Gunning and Paterson propose to use

a normalized sum of squares of the differences between the synthetic and the recalibrated true seismic trace. They accelerate the convergence by using an algorithm similar to that of Metropolis-Hastings.

5.4.3 Analytical approach

Eide et al. (1997a) start from the same approach as that of Bortoli et al., but with a view to solve the problem analytically. They formalize the geostatistical inversion problem in a Bayesian-Gaussian framework and calculate the expression for the posterior reflection coefficient pdf, that is, the prior covariance-based model is updated by the well and seismic information.

Eide et al. address the problem in the very general situation of an unknown seismic wavelet, to be determined during the inversion process. They show that, in the situation where this wavelet is unknown, the posterior pdf cannot be determined analytically. As a result, the sampling of the realizations is not trivial and is best addressed using the Metropolis algorithm (Fig. 4-24), which converges toward the required pdf. However, this algorithm remains very time-consuming.

In the more-frequent situation in which the wavelet is known (from a previous seismic calibration exercise performed by the geophysicist) and the relationship between reflectivity and impedance is linearized, the posterior pdf of acoustic impedance is Gaussian and its sampling is trivial, thanks to such approaches as SGS.

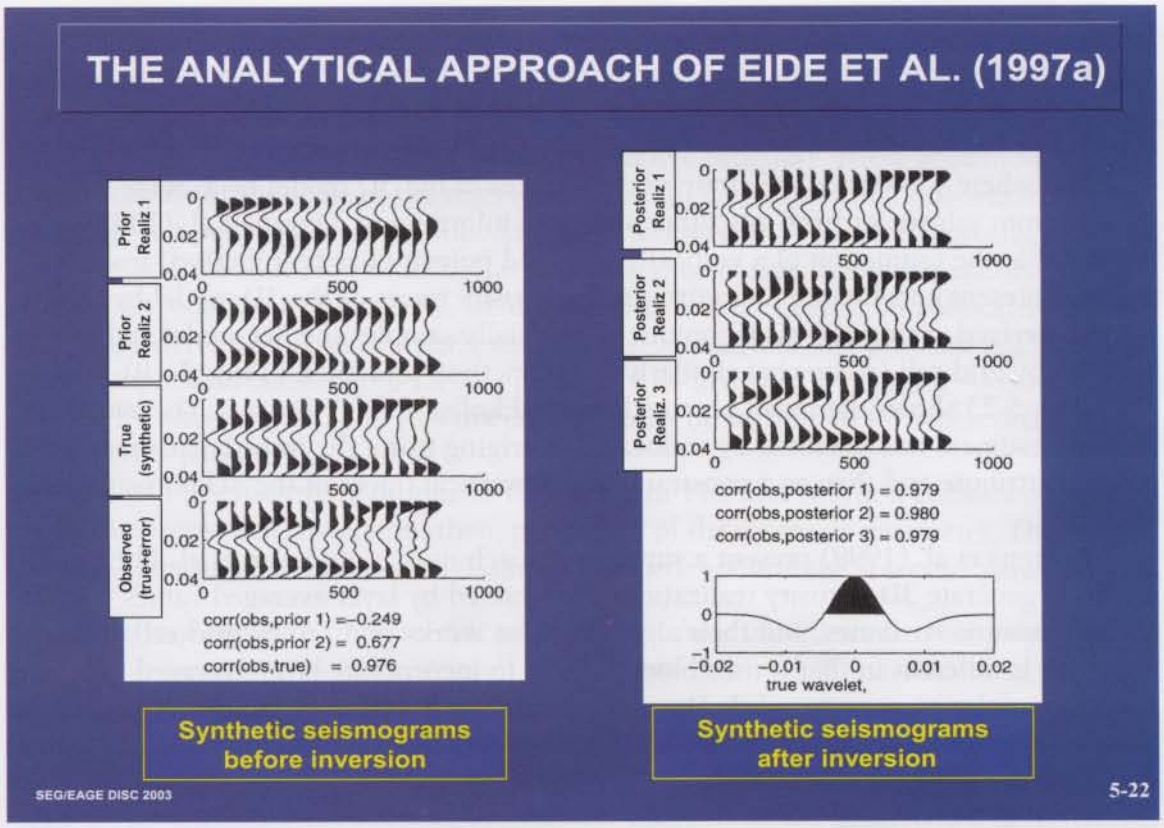
Buland and Omre (2003) also propose a Bayesian AVO inversion technique where the solution is given in an analytical form. The model parameters are *P*-wave velocity, *S*-wave velocity, and density, which are assumed to follow lognormal distributions (Fig. 1-27). The inversion method is based on a weak contrast approximation to seismic reflectivity, as proposed by Aki and Richards (1980). Thanks to these simplifying assumptions, the posterior distribution of the three model parameters is multivariate normal (Fig. 1-38) and can be calculated analytically. Realizations from these distributions can easily be sampled using previously discussed techniques. However, an important limitation of the model is that it is a single-trace inverse algorithm, because there is no lateral correlation between vertical traces.

The mathematical developments of Eide et al. are somewhat cumbersome and the computer implementation is slow, but the approach has the merit of formalizing the problem consistently. As a result, the theoretically correct posterior distribution allows a proper assessment of uncertainties. The approach seems to provide satisfactory results on synthetic cases (Fig. 5-22).

5.4.4 Emerging techniques

- Wavelet inversion

A. Buland and H. Omre (personal communication, 2001) have developed a Bayesian method for wavelet inversion. The method works both on stacked data and on prestack data in the form of angle gathers. Seismic noise, errors in log data, and also possible mis-ties between the seismic and well time axis, can be incorporated in the model. The solution is not analytical, but is obtained by MCMC. Uncertainty in the estimated wavelet is also quantified in this process.



- Processing larger and larger seismic data sets, including AVO data

The algorithms used for geostatistical inversion will certainly evolve, because the generation of a large number of realizations constrained by seismic data remains very time-consuming, and because the size and information content of seismic data sets is likely to grow in the future.

The inversion of prestack seismic data is around the corner. Buland et al. (2003) propose an AVO inversion technique incorporating spatial correlation between model parameters. As in Buland and Omre (2003), the inverted parameters are *P*-wave velocity, *S*-wave velocity, and density, and the inversion method is based on a weak contrast approximation to seismic reflectivity. In the Fourier domain, the spatially correlated parameters can be decoupled, and the inversion problem can be solved independently for each frequency component. This may be one of the most interesting ways to reduce computer time, as shown by the promising results obtained on the inversion of a 3D data set from the Sleipner field represented by three angle stacks on a grid with four million grid cells.

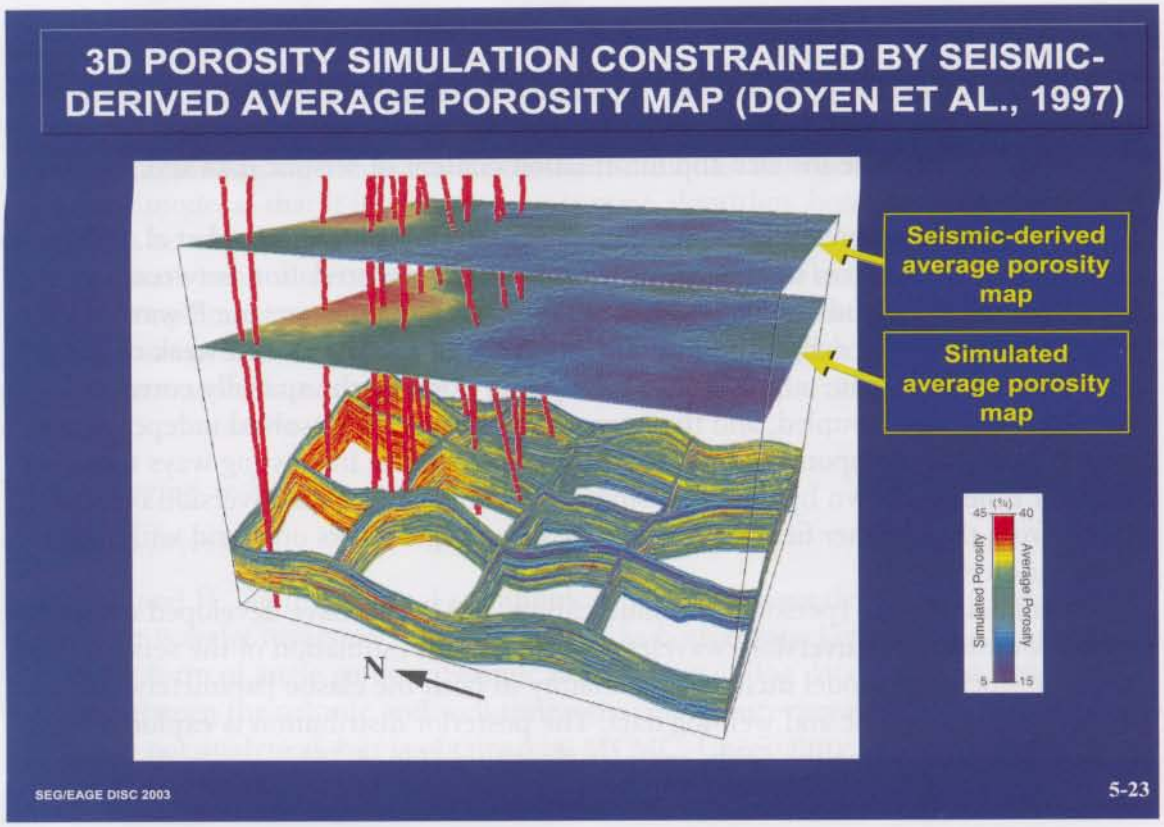
Buland and Omre (personal communication, 2002) also have developed a Bayesian method for joint AVO inversion, wavelet estimation, and estimation of the seismic noise level. The stochastic model includes uncertainty in both the elastic parameters, the wavelet, and the seismic and well-log data. The posterior distribution is explored by MCMC simulation using the Gibbs sampler algorithm (Fig. 4-25).

5.5 A Generalized Downscaling Approach

So far, geostatistical inversion has been presented as a method for constraining 3D acoustic-impedance models by seismic data, using a convolution model. However, GI can also be considered as a general "downscaling" technique, and can be applied in the situation where we wish to constrain vertical traces of the 3D model by average values derived from seismic or from any other source of information. Doyen et al. (1997) assume that the estimation of a vertically averaged porosity has been derived from seismic and present a technique for constraining porosity traces of the 3D model by these seismic-derived averages. This technique sequentially samples the 3D model and works grid cell by grid cell (somewhat similarly to the method presented in Fig. 5-18).

Fig. 5-23 shows the result obtained on the Ekofisk field (Norway). The average porosity estimate was obtained by collocated cokriging using the impedance map as a guiding attribute and then as a constraint on the vertical traces of the 3D porosity realization.

Behrens et al. (1998) present a similar approach to that of Doyen et al. Their goal is also to generate 3D porosity realizations constrained by layer-averaged values predicted from seismic attributes, and their algorithm also works grid cell by grid cell. The approach is different in that it uses block kriging to incorporate layer-averaged values in the SGS simulation process, while Doyen et al. use the Bayesian formalism. The Bayesian formalism appears to provide more flexibility for handling uncertainties associated with the layer-averaged values.



5.6 Going Further with Geostatistical Inversion Results

Earlier in this chapter, we saw that geostatistical inversion led to the generation of 3D acoustic-impedance realizations all matching the seismic data up to a certain degree, as measured by the value of the trace-by-trace objective function. These acoustic-impedance traces are at the scale of the reservoir model. In favorable situations where there is a relationship between acoustic impedance and a reservoir parameter (porosity, facies, net/gross and the like), the next logical step is to predict this reservoir parameter from acoustic impedance. In Chapter 4, we described techniques based on collocated cosimulation for deriving a reservoir parameter from seismic. Here, the problem is a little bit more complicated because the seismic parameter — acoustic impedance — is now itself affected by uncertainties, and this uncertainty must be accounted for when predicting the reservoir parameter.

We will discuss two solutions to the problem. The first solution consists of working in two steps: first, inversion, then, prediction of the reservoir parameter. The second solution consists of simultaneously sampling impedance and the reservoir parameter.

5.6.1 Two-step approach: from seismic to impedance, from impedance to other properties

- Generalization of collocated cokriging

Lamy et al. (1999) argue that the prediction of reservoir parameters must be performed in two steps: first, geostatistical inversion of acoustic impedance, and then, prediction of reservoir parameters from acoustic impedance. This allows the geoscientist to have a careful look at the outcome of the first step before embarking on the second one.

The example used by Lamy et al. is that of Fig. 5-8. There is a linear relationship between acoustic impedance and V_{shale}, and their goal is to predict the value of V_{shale} at each location by combining the results of geostatistical inversion with the information provided by the V_{shale} well logs. If there were no uncertainty affecting the acoustic impedance derived from geostatistical inversion, the problem would be addressed by collocated cokriging, which combines the V_{shale} kriging estimate with the estimate derived from acoustic impedance, using the correlation coefficient between impedance and V_{shale} as a weighting factor (Fig. 3-75).

Now, because acoustic impedance is affected by uncertainty, the weighting factor must be changed. Acoustic impedance must have less weight than in the situation where it is not affected by uncertainty. Lamy et al. derive a new formula (Fig. 5-24) for the correlation coefficient, which is now location-dependent. Logically, this new coefficient is smaller than the one used in the situation where uncertainty affecting acoustic impedance is ignored, and both are equal if the variance affecting the result of inversion is zero. The smaller the ratio in the denominator, that is, the greater the gain in variance resulting from inversion, the closer the new coefficient is to the old one. Fig. 5-25 shows the results obtained by Lamy et al. on the same case study as that of Fig. 5-9.

- Generalization of collocated cosimulation

Marion et al. (2000) generalize the approach of Lamy et al. (1999) to the generation of stochastic realizations. In their case study, they have a good relationship between

MODIFIED CORRELATION COEFFICIENT FOR PREDICTING V_{SHALE} (LAMY ET AL., 1999)

Correlation coefficient associated with cross-plot between acoustic impedance and V_{shale}

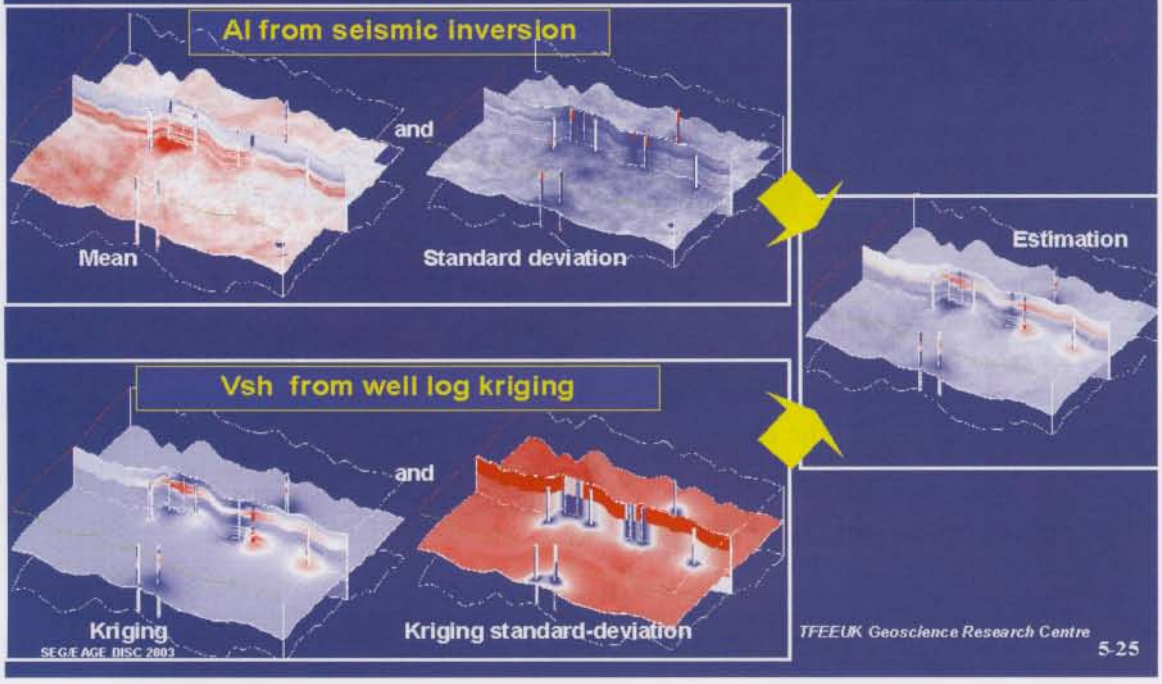
$$\rho^2(\mathbf{x}) = \frac{\rho^2}{1 + \frac{\sigma^2_{GI}(\mathbf{x})}{\sigma^2_{apriori}}}$$

Variance associated with inverted acoustic impedance at location x

Acoustic impedance a priori variance

SEG/EAGE DISC 2003 TFEEUK Geoscience Research Centre 5-24

USING RESULTS OF GEOSTATISTICAL INVERSION TO PREDICT Vshale (LAMY ET AL., 1999)



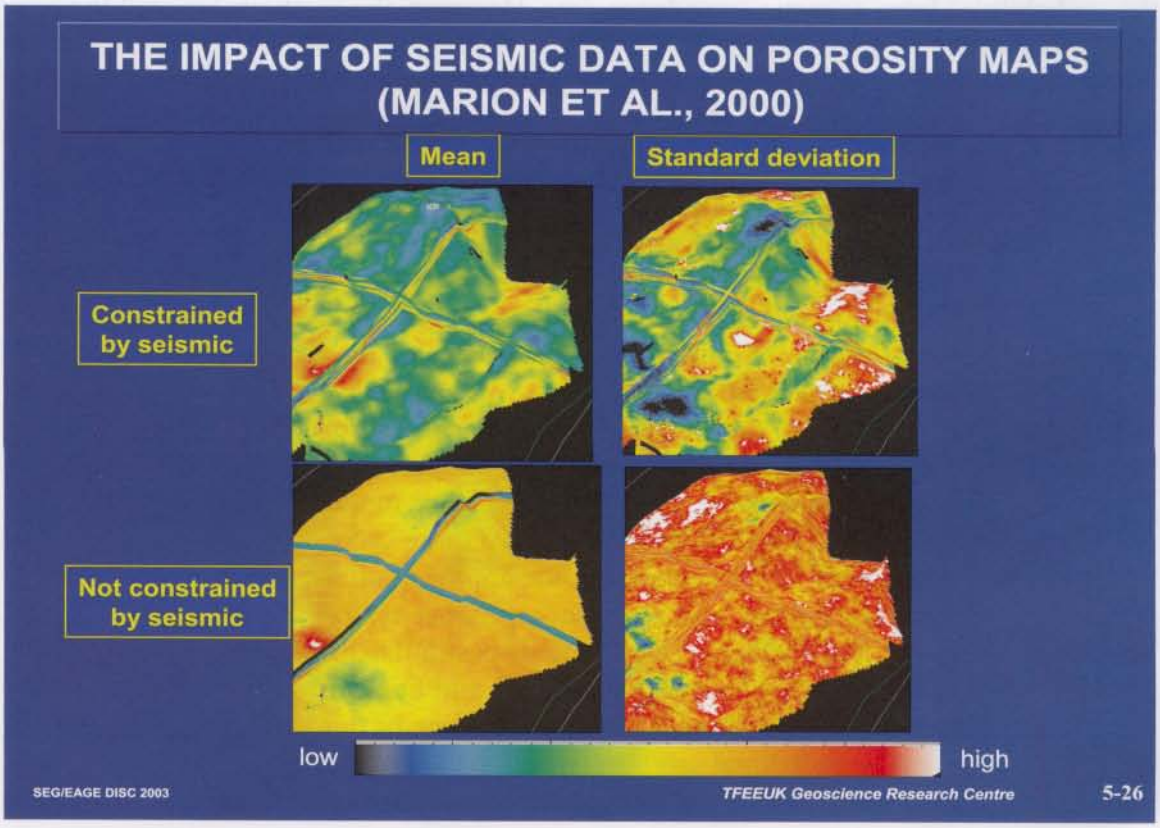
porosity and acoustic impedance. This leads to about 50 porosity realizations, all constrained by seismic data, wells, and a sedimentary model produced by the geologist. Obviously, the mean of the porosity realizations is smoother for the realizations unconstrained by seismic (Fig. 5-26), because, having no seismic constraint away from the wells, the average of realizations — which is equivalent to kriging — simply returns to the global mean. There is also a clear difference between the color of the two maps, because the use of seismic for constraint leads to a degradation of porosity realizations.

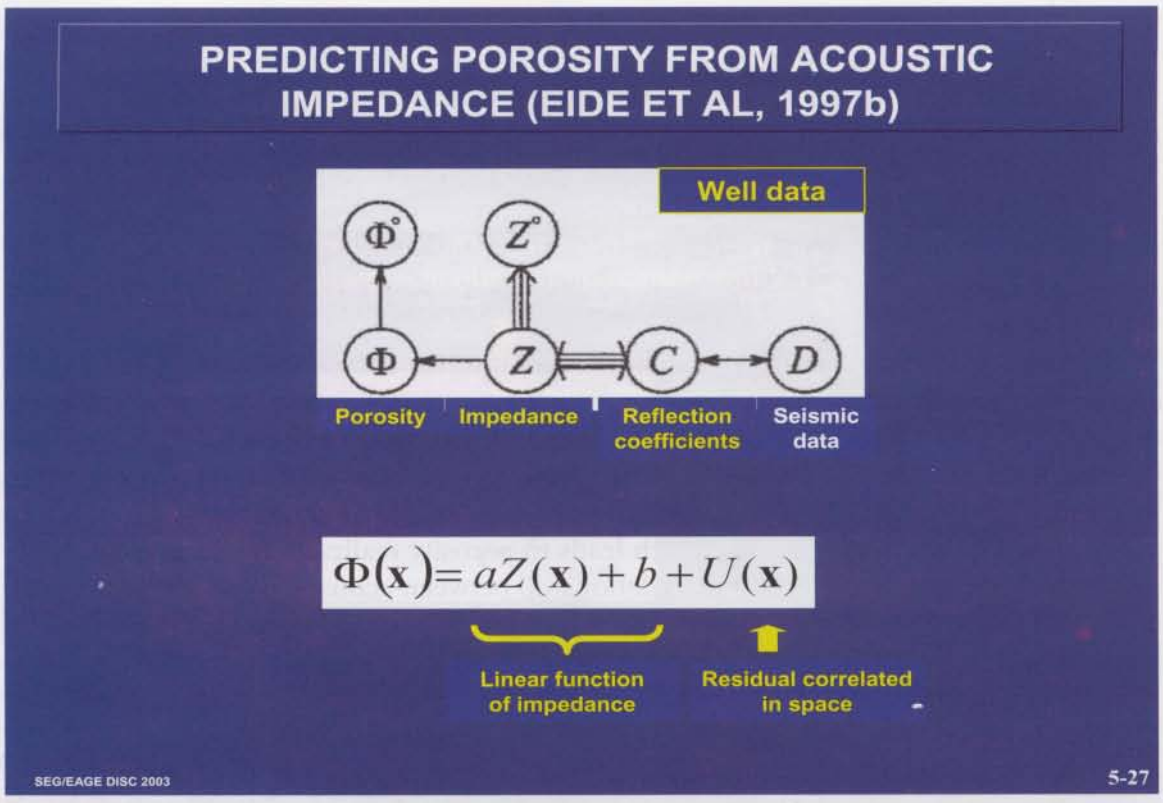
5.6.2 *Predicting porosity during the acoustic-impedance inversion process*

Eide et al. (1997b) propose an approach that accounts for the uncertainty affecting acoustic-impedance realizations when predicting a parameter such as porosity. In the synthetic example they present, the model they use (Fig. 5-27) assumes that porosity is equal to a deterministic function of acoustic impedance plus a random term correlated in space. They construct impedance realizations constrained to seismic data, then they predict porosity using this model, which leads to porosity realizations that account not only for the uncertainty affecting the relationship between acoustic impedance and porosity, but also for the uncertainty affecting inverted acoustic impedance itself.

5.6.3 *Predicting facies during the acoustic-impedance inversion process*

In Chapter 4, we discussed methods for constraining discrete facies models by seismic data (Fig. 4-53). We saw that Approach 1 consisted of deriving probabilities from seismic and then constraining facies models by these probabilities. Most methods derived





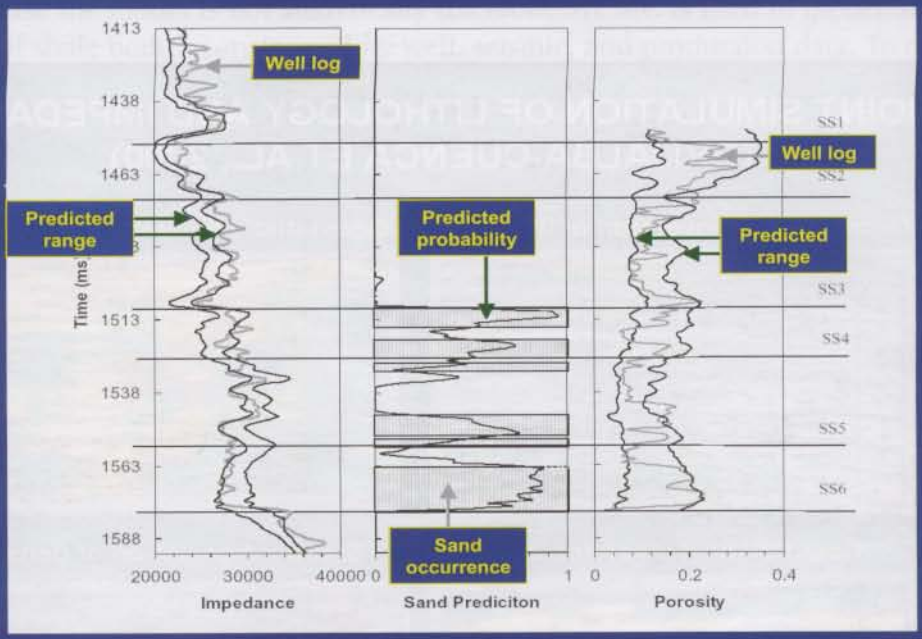
the probabilities from acoustic-impedance data. However, the methods ignored the uncertainty affecting these inverted data. We are now going to discuss methods associated with Approach 2, wherein facies and impedance models are jointly simulated.

- Joint sequential simulation of lithology and acoustic impedance

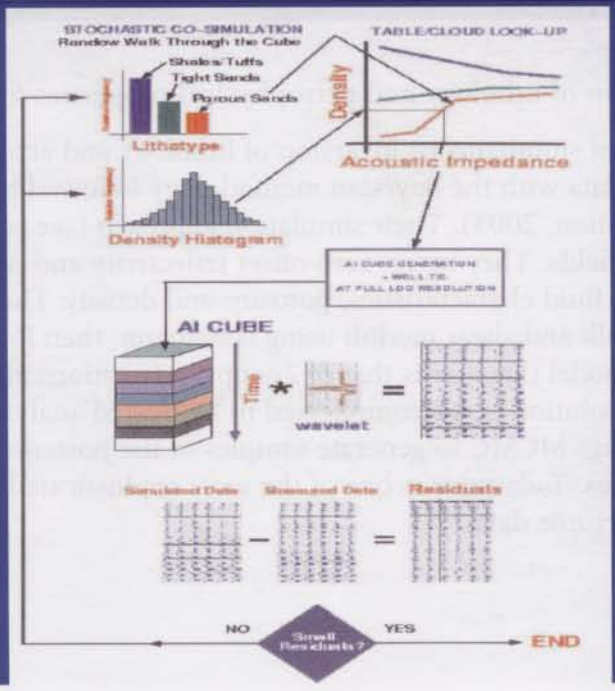
Sams et al. (1999) propose a promising approach whereby facies and acoustic-impedance realizations are simultaneously inverted. An initial realization of lithology and associated impedance is generated that is consistent with the geostatistical input and the well data. Then this model is iteratively updated at each grid cell by geostatistically sampling new values of lithology and impedance, such that the match between synthetic and actual seismic is improved. This approach was applied with a vertical sampling of 1 ms (about 1 m) in a field of the Central Sumatra Basin, and proved capable of resolving sand units that could not be resolved with other approaches. Porosity distribution was then predicted from each of the lithology–impedance model realizations. Validation of the approach at one blind well (Fig. 5-28) proved quite successful.

Grijalba-Cuenca et al. (2000) generalize the approach of Sams et al. to the joint simulation of lithofacies, density, and acoustic-impedance realizations in a field in Argentina (Fig. 5-29). For each grid cell, lithology is simulated first, then density is sampled from the lithology-dependent distribution, and finally acoustic impedance is sampled from the bivariate distribution of impedance versus density. Once again, the objective was to resolve individual sand units, and the simulations were performed with a vertical resolution better than 2 ms. The outcome consisted of high-resolution cubes

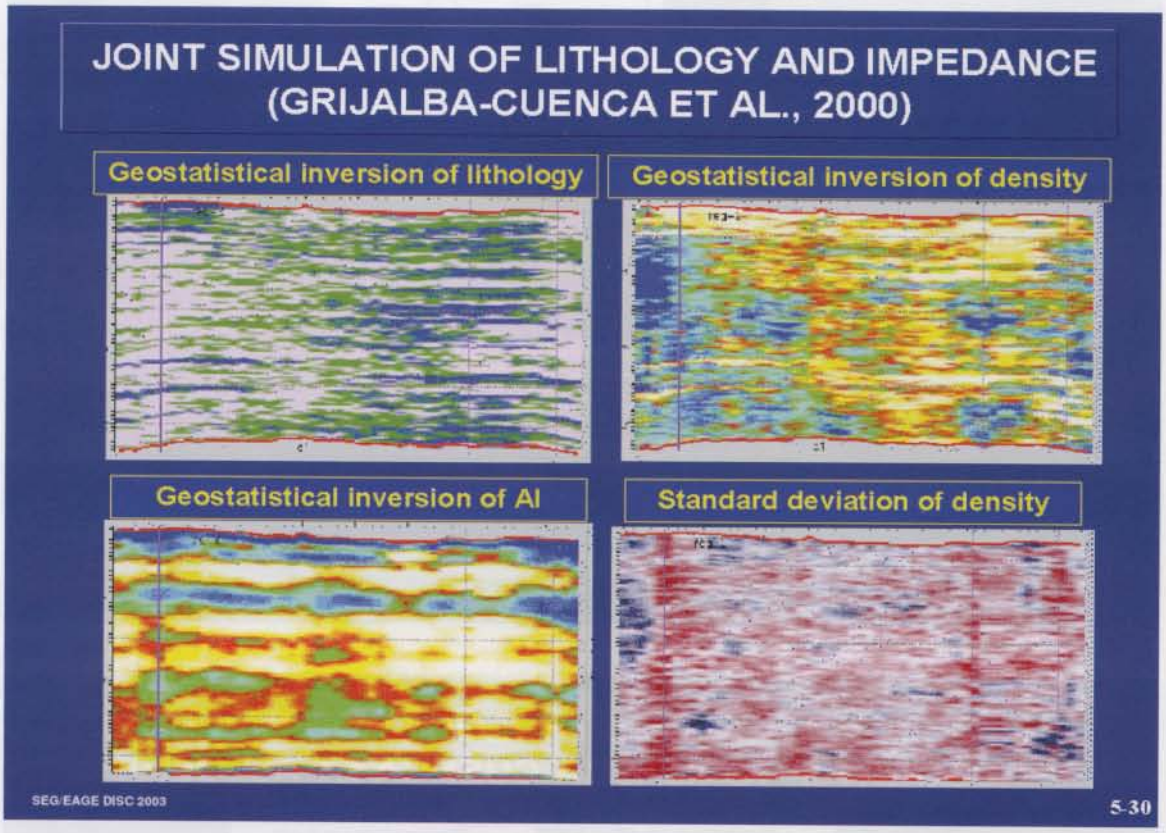
JOINT SIMULATION OF LITHOLOGY AND IMPEDANCE BLIND WELL TEST (SAMS ET AL., 1999)



JOINT SIMULATION OF LITHOLOGY AND IMPEDANCE ALGORITHM (GRIJALBA-CUENCA ET AL., 2000)



of lithology and density, which were fully consistent with the 3D seismic data, thanks to the use of acoustic impedance as a variable in the process. Fig. 5-30 represents the average of six realizations, which show the good consistency between the different variables.



- Joint simulation of lithology and petrophysical properties for AVO inversion

This philosophy of simultaneous inversion of lithology and acoustic impedance is generalized to AVO data with the Bayesian methodology followed by J. Eidsvik et al. (personal communication, 2003). Their simulation approach (see section 4.5.2) is based on Markov random fields. They invert zero-offset reflectivity and AVO gradient by jointly simulating facies, fluid characteristics, porosity, and density. They derive from this joint simulation the bulk and shear moduli using Gassmann, then *P*-wave and *S*-wave velocity. The forward model they use is that of Zoeppritz equations and approximations by Shuey. Since the solution is too complicated to be treated analytically, they use Metropolis-Hastings MCMC to generate samples of the posterior distribution of the reservoir properties. Today, this is one of the most sophisticated applications of stochastic inversion to seismic data.

- Direct inversion of object-based models

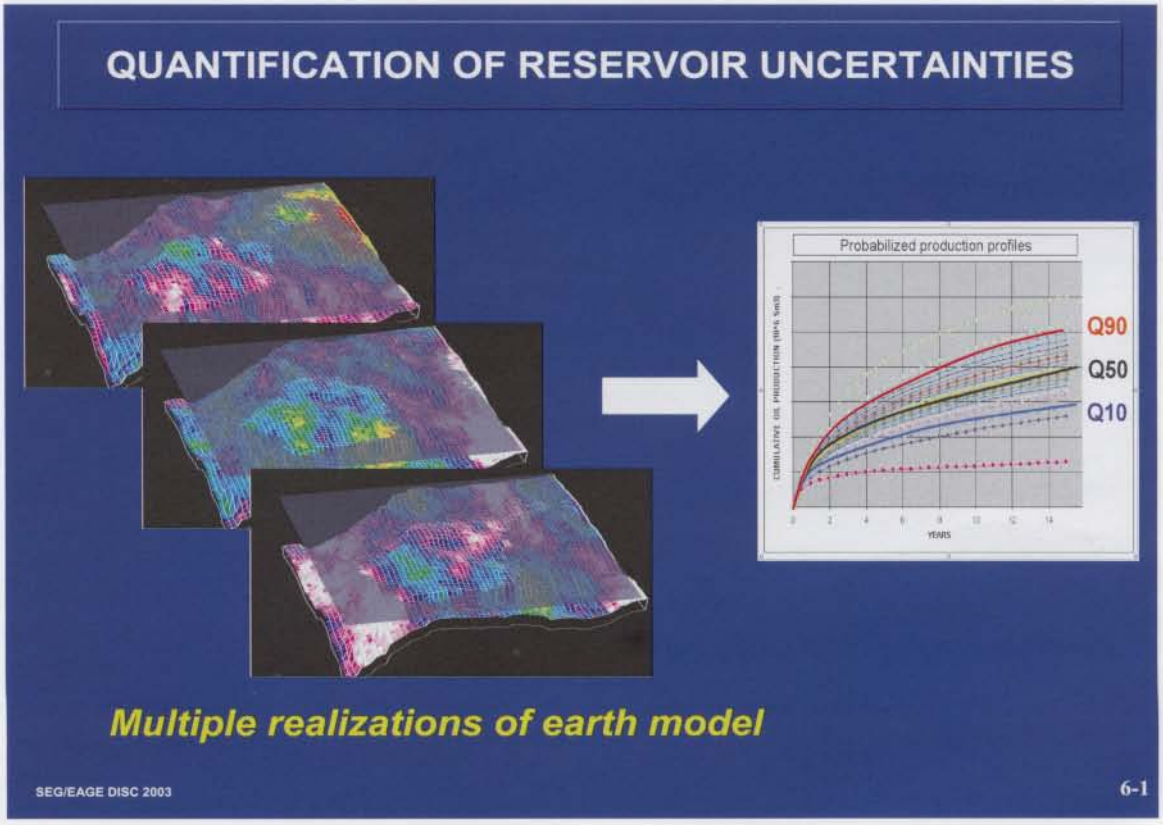
Tjelmeland and Omre (1997) present a general approach for constraining an object-based model of the distribution of shales to well, seismic, and production data. Once again, because the model is not analytically tractable, MCMC is used to model 3D distributions of shale bodies constrained by well, seismic, and production data. To our knowledge, industry applications along those lines have remained limited.

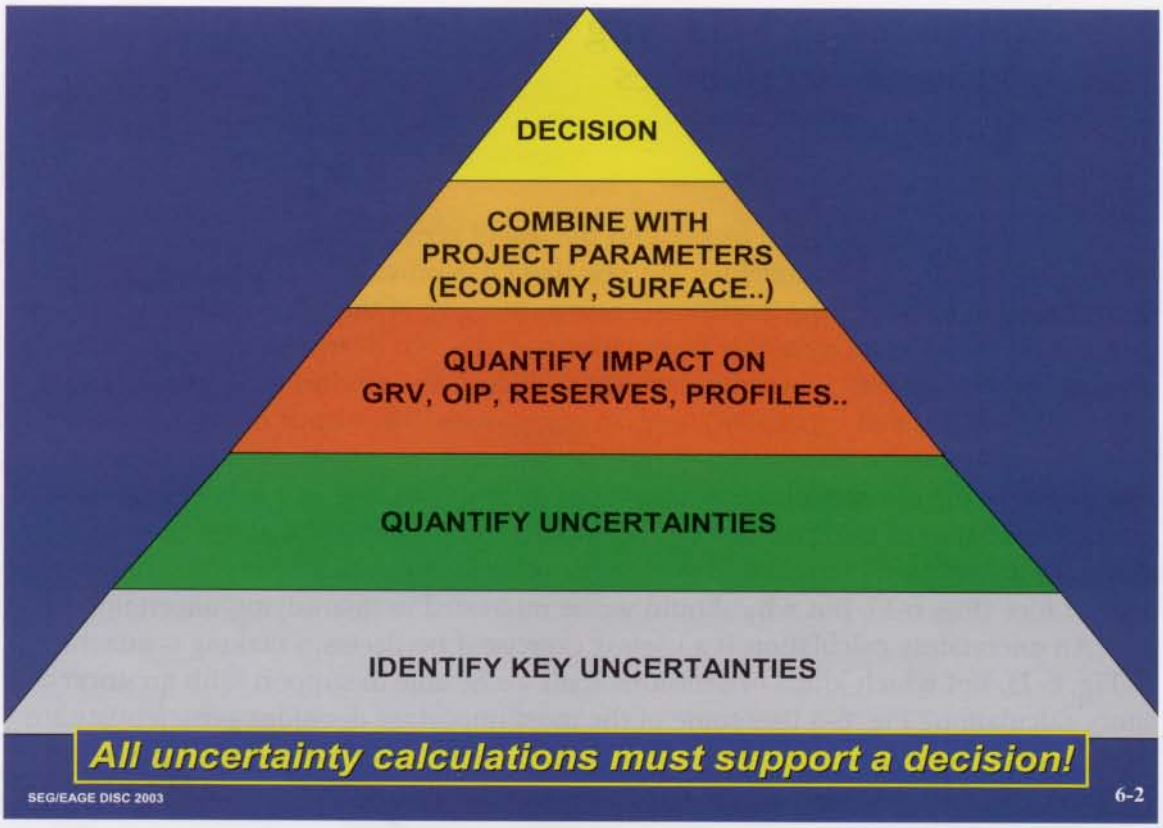
6 Stochastic Earth Modeling That Integrates All Subsurface Uncertainties

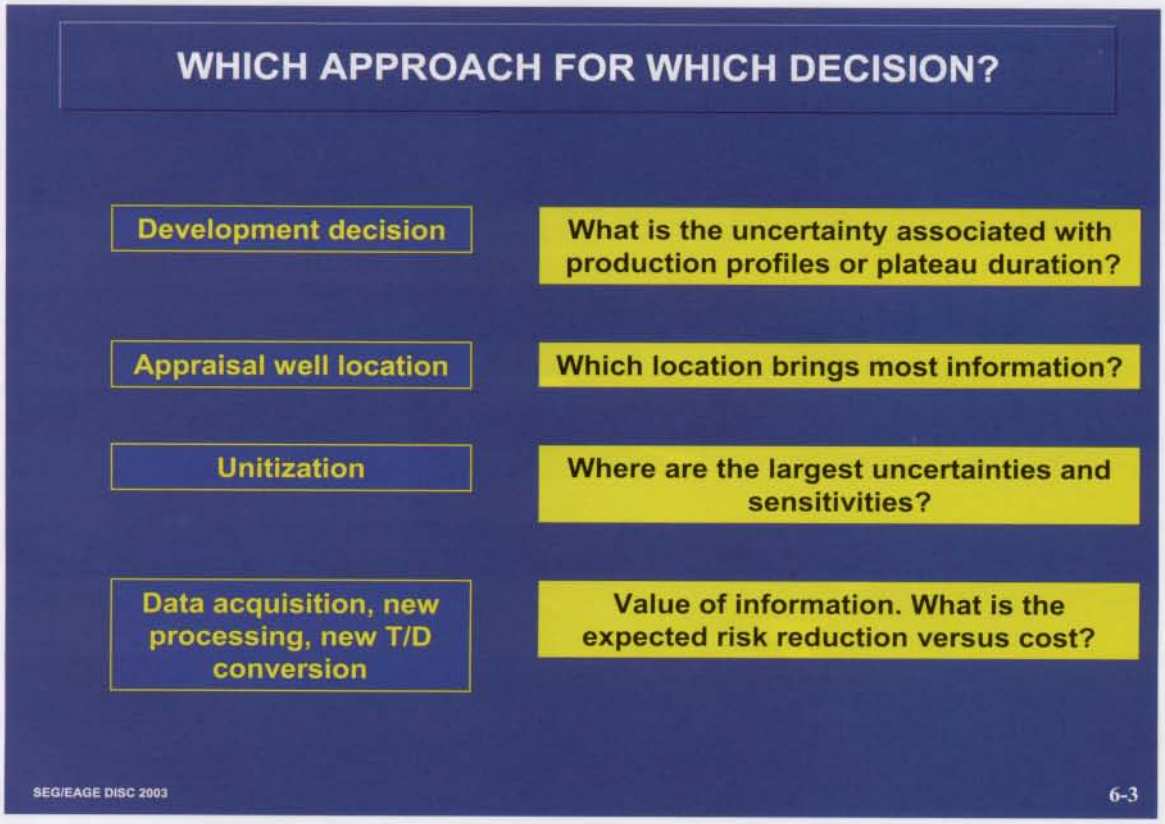
6.1 Introduction

In the previous pages, we have focused on the use of geostatistical conditional simulation for 3D heterogeneity modeling. We saw that GCS provided a satisfactory solution to the problem of generating realistic 3D representations of the subsurface. We also saw that, thanks to GCS, we were able to generate not one, but a large number of realizations, all of which were compatible with the well data, the a priori geostatistical constraints (histogram and variogram), and, in many cases, the seismic data. The variability from one realization to another was a representation of the remaining uncertainty left after constraining our models by all this input information. We will now discuss how this quantification of uncertainties can be applied to all parameters of the earth model to lead to uncertainties attached to gross-rock volume, oil-in-place, reserves, or production profiles (Fig. 6-1). But why should we be interested in quantifying uncertainties?

An uncertainty calculation is a useless exercise if no decision making is attached to it (Fig. 6-2). But which kinds of decisions shall we be able to support with an uncertainty calculation? Fig. 6-3 lists some of the most important decisions geoscientists are led to support with their uncertainty studies (see examples in Tyler et al., 1996 and Charles et al., 2001). Usually, these decisions are related to a significant financial invest-




DECISION
COMBINE WITH PROJECT PARAMETERS (ECONOMY, SURFACE..)
QUANTIFY IMPACT ON GRV, OIP, RESERVES, PROFILES..
QUANTIFY UNCERTAINTIES
IDENTIFY KEY UNCERTAINTIES
All uncertainty calculations must support a decision!
SEG/EAGE DISC 2003
6-2

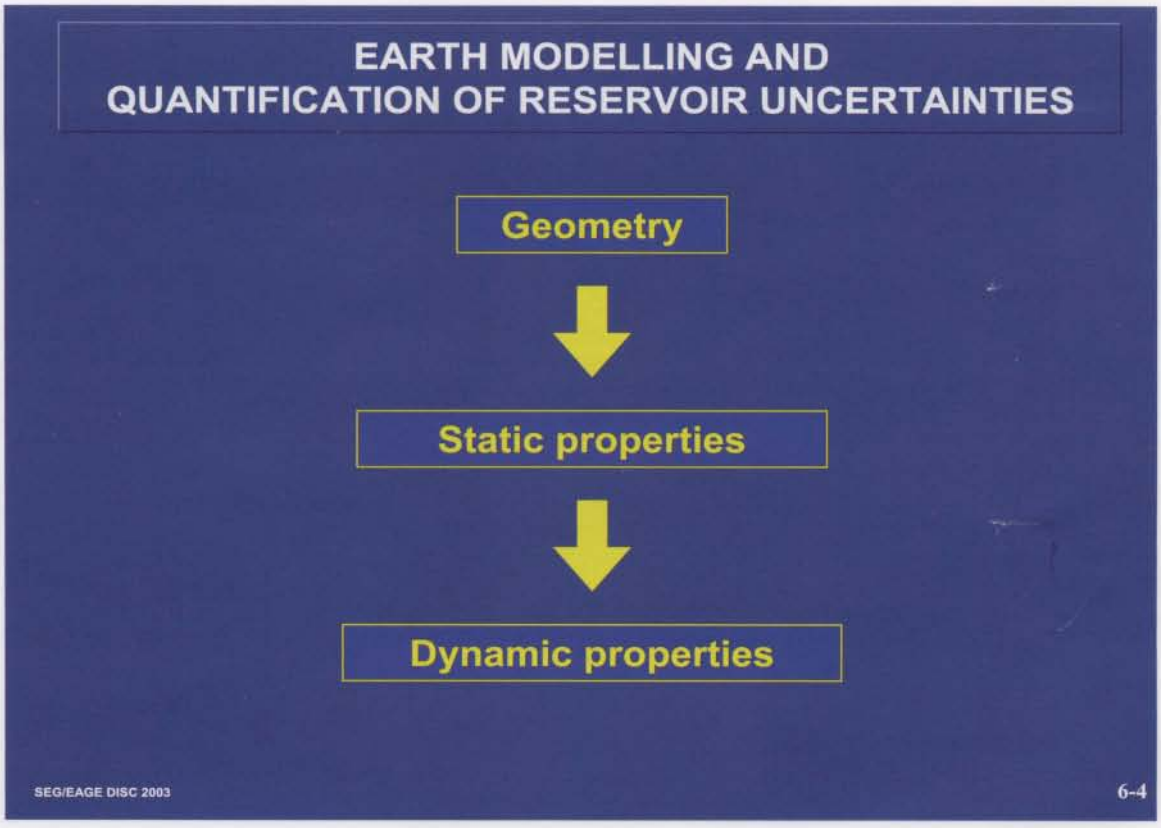

WHICH APPROACH FOR WHICH DECISION?
Development decision
What is the uncertainty associated with production profiles or plateau duration?
Appraisal well location
Which location brings most information?
Unitization
Where are the largest uncertainties and sensitivities?
Data acquisition, new processing, new T/D conversion
Value of information. What is the expected risk reduction versus cost?
SEG/EAGE DISC 2003
6-3

ment. Instead of one production profile, a typical uncertainty study will produce a family of production profiles or the field reserves pdf.

In the past, the geoscientist or reservoir engineer was asked to produce just one reserves figure or one production profile, in spite of the great uncertainty affecting the results. In a way, the geoscientist was asked, by choosing one scenario over many other possible ones, to substitute himself/herself for the decision maker. It can be said that an uncertainty-quantification approach, by attaching a risk to each possible decision, will put the decision back in the hands of the decision maker.

The Norwegian School has been a pioneer in the quantification of earth model uncertainties. See, for instance, the work of Sandsdalen et al. (1996), Damsleth and Omre (1997), and Lia et al. (1997). The Lia paper can be regarded as a classic that showed it was possible to combine all the uncertainties affecting the different building blocks of a 3D earth model and quantify their impact on production profiles. Hegstad and Omre (2001) show the progress that has been made in less than five years, by developing an earth model uncertainty-quantification approach constrained by seismic and dynamic data.

Nowadays, most earth model uncertainty studies are performed in three steps (Fig. 6-4). Although each step corresponds more or less to the quantification of uncertainties associated with a different discipline (geophysics, geology/petrophysics, and reservoir engineering), the combination of all the uncertainties in the earth model provides a fantastic multidisciplinary integration tool. The direct quantification of the impact of structural uncertainties on fluid flow, an issue that was often ignored in the past, can now be addressed.



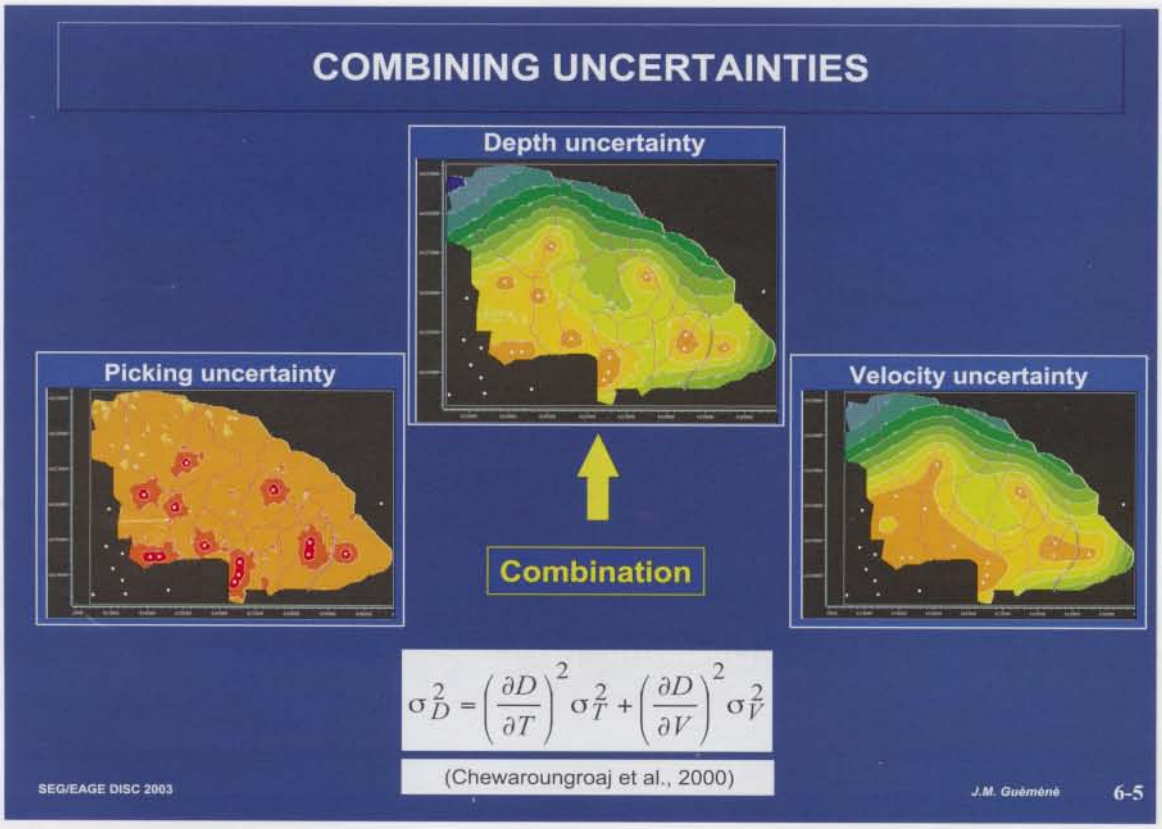
6.2 Geometrical Uncertainties

We present two different models for quantifying geometrical uncertainties and their impact on gross-rock-volume uncertainties. The first approach is used extensively at TotalFinaElf (Samson et al., 1996; Guéméné et al., 2002; Thore et al., 2002), and has been designed to let the seismic interpreter give as much input as possible to the uncertainty quantification, by producing separate picking- and time-to-depth conversion-uncertainty maps. The second approach, used for a long time by the Norwegian School (see Abrahamsen et al., 1991 or Abrahamsen et al., 2000), relies on the Bayesian kriging formalism.

6.2.1 Two-step approach

- Quantifying picking and time-to-depth conversion-uncertainty maps

In most cases, a depth map is the result of the combination of a time map and a velocity map. The seismic interpreter is best placed for quantifying the picking uncertainties affecting his/her interpretation. The velocity model may be the result of a great variety of computations, depending on the velocity data available to start with. The basis of the two-step approach is to produce two separate and independent uncertainty maps, then combine them as shown in Fig. 6-5 to obtain a global depth-uncertainty map (see Thore et al., 2002 for a complete discussion).



- *Uncertainty maps*

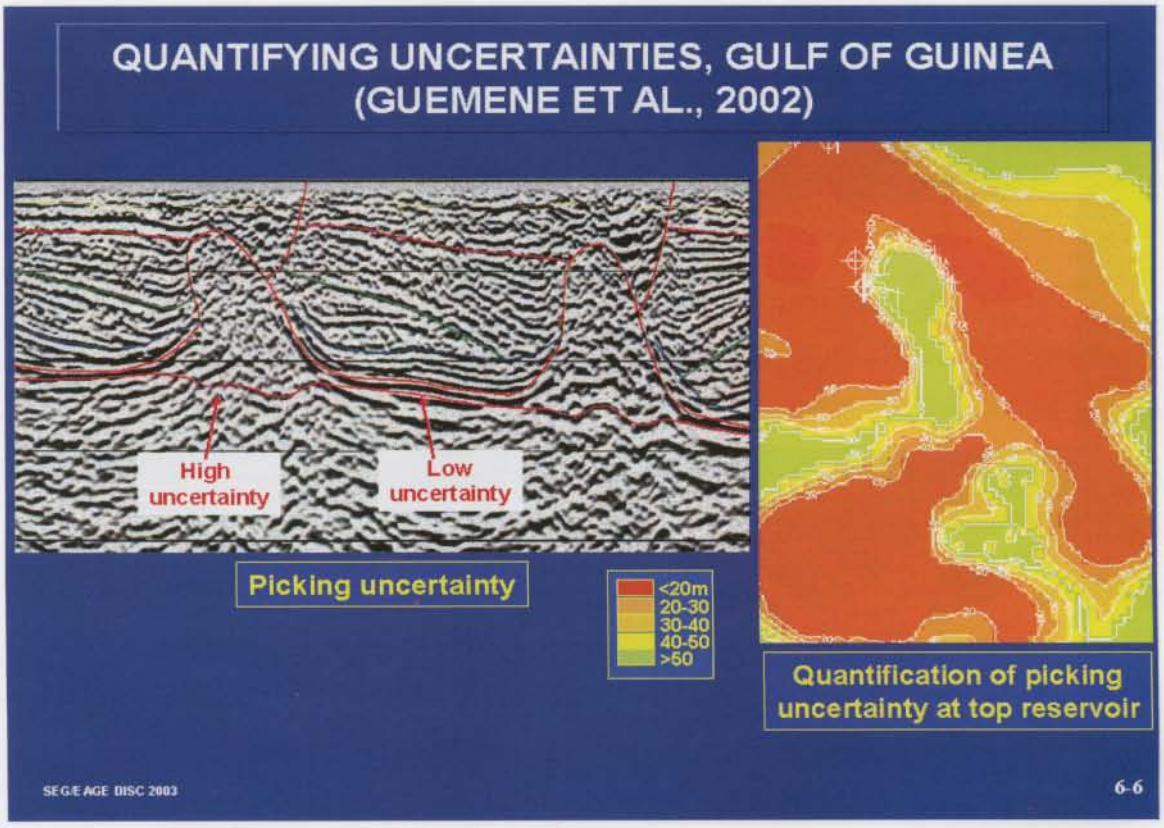
What do we exactly mean by "uncertainty map"? Let us take, for instance, the velocity uncertainty map of Fig. 6-5. For practical purposes, we will assume that if the velocity is V and if the uncertainty is ΔV at a given map location, this means that we have a 95% chance that the actual unknown velocity value falls within the interval $[V - \Delta V, V + \Delta V]$. We will also assume that errors are normally distributed. If we remember what was said in Fig. 1-26, this implies that the uncertainty is equal to twice the standard deviation of the errors. This is to be related to the discussion of the kriging standard deviation in Section 3.2.2.

- *Picking uncertainty*

The picking uncertainty map must be built by the seismic interpreter. In the example from Guéméné et al. (2002) that is displayed in Fig. 6-6, the interpretation was based on the results of PSDM, and performed in depth. The analysis of all the seismic sections resulted in the mapping of areas of poor, fair, and good seismic quality. The confidence interval around the picked marker was estimated at 50 m in the areas of poor quality (below salt domes, ends of lines) and 20 m where the seismic image was better. The absolute magnitude of the uncertainty is a matter of experience and judgment from the interpreter.

- *Time-to-depth conversion uncertainty*

In the example of Fig. 6-6, the PSDM was performed with a velocity model com-



puted from focalization analysis on each 2D line. The main uncertainty on velocity was associated with the lateral instability of the velocity pick and to its approximation by a smooth mathematical curve. The impact on depth of the velocity uncertainty was estimated to vary between about 30 and 60 m (Fig. 6-7). A third uncertainty was that related to the interpolation of the depth picks between 2D seismic lines, and was derived from the depth kriging standard deviation, which increases away from seismic lines.

- *Combining the uncertainties*

All these depth-uncertainty maps were combined to obtain the total uncertainty map of Fig. 6-7. Because all uncertainties were already expressed in depth, the partial derivative term of Fig. 6-5 was not needed. Let us stress here that confidence intervals cannot be added to each other. Only the variance of the sum of independent errors is equal to the sum of their individual variances (Fig. 1-15). This is why the formula of Fig. 6-5 applies to squares and not to absolute values. Because confidence intervals are proportional to standard deviations, only squares of confidence intervals (which are proportional to variances) can be added to each other.

- Translating structural uncertainties into realizations

Our ultimate goal is to combine various uncertainties affecting parameters of the earth model in order to quantify their joint impact on GRV, STOOIP, reserves, or production profiles (Fig. 6-2). To reach this goal, we will need to combine realizations of depth maps with realizations of petrophysical parameters, as shown in the three realizations of

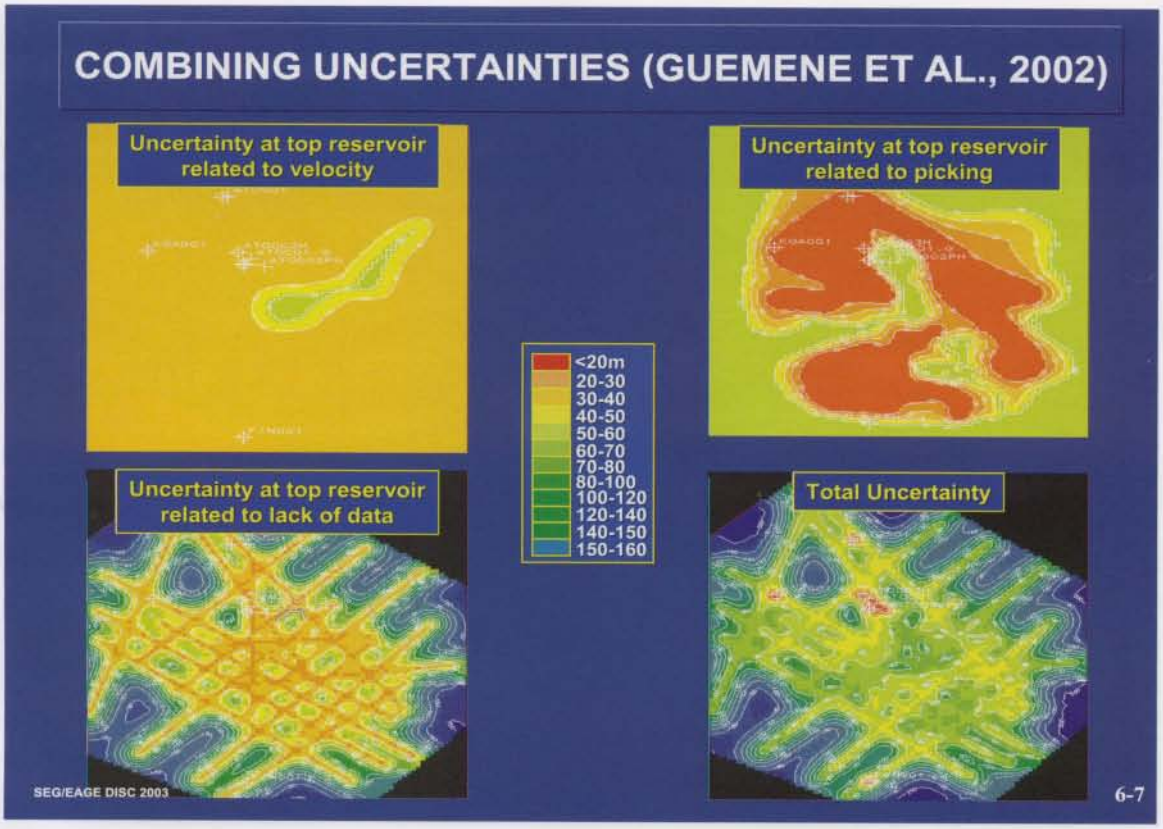
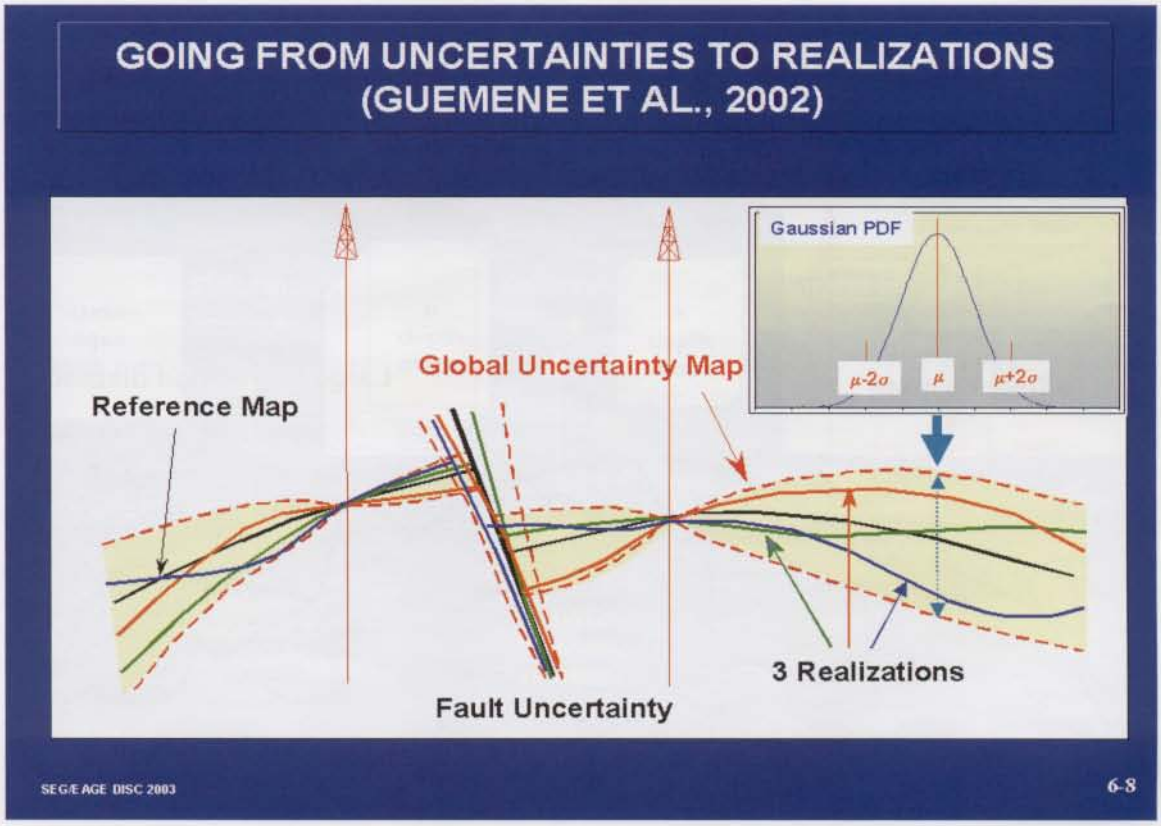
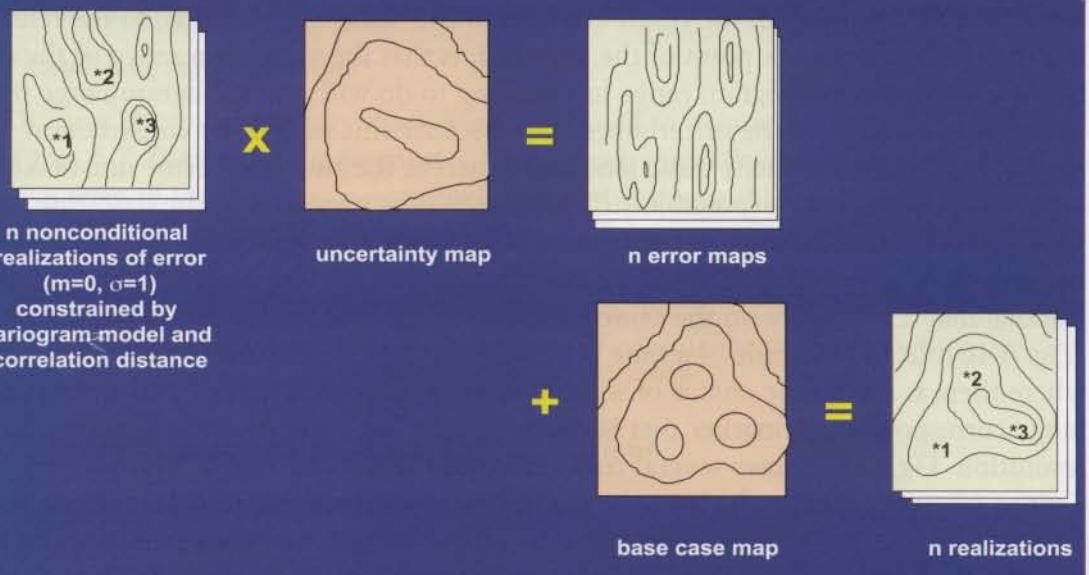


Fig. 6-1. This means that we must be able to generate multiple realizations of picked surfaces, as shown in Fig. 6-8.

Thanks to our definition of uncertainty, the task of generating depth realizations will be easy. The approach is explained in Fig. 6-9. Nonconditional simulations of normally distributed "errors" of mean zero and standard deviation 1 are multiplied by the uncertainty map and then the result is added to the "base case," that is, to the reference interpretation. One weak point of the approach is that the variogram range of the errors is not known. This variogram range has nothing to do with the variogram range that would be derived from the uncertainty map, because this map is, by construction, always positive (and usually quite smooth), whereas the simulated error must take negative and positive values (possible structural scenarios vary around the base-case interpretation). Thus, the choice of the range is left to the interpreter and the "feeling" he or she may have about the pattern of variation of the error. A large range will result in realizations that tend to stay on the same side of the base case longer than a short range (Fig. 6-10). In the example of a time pick, a large-range error corresponds to a situation where there is a multiple choice of seismic loops to pick, whereas a small variogram range corresponds to a marker that is unambiguous but fuzzy because of poor seismic resolution. Fig. 6-11 shows several realizations of the top of a formation in a Gulf of Guinea reservoir. Obviously, once the realizations have been generated, they can be associated with hydrocarbon-water-contact (HWC) values, and the corresponding GRV can be calculated (Fig. 6-12). After eliminating unrealistic realizations, such as those not showing closure, we can proceed to a pdf of the GRV.



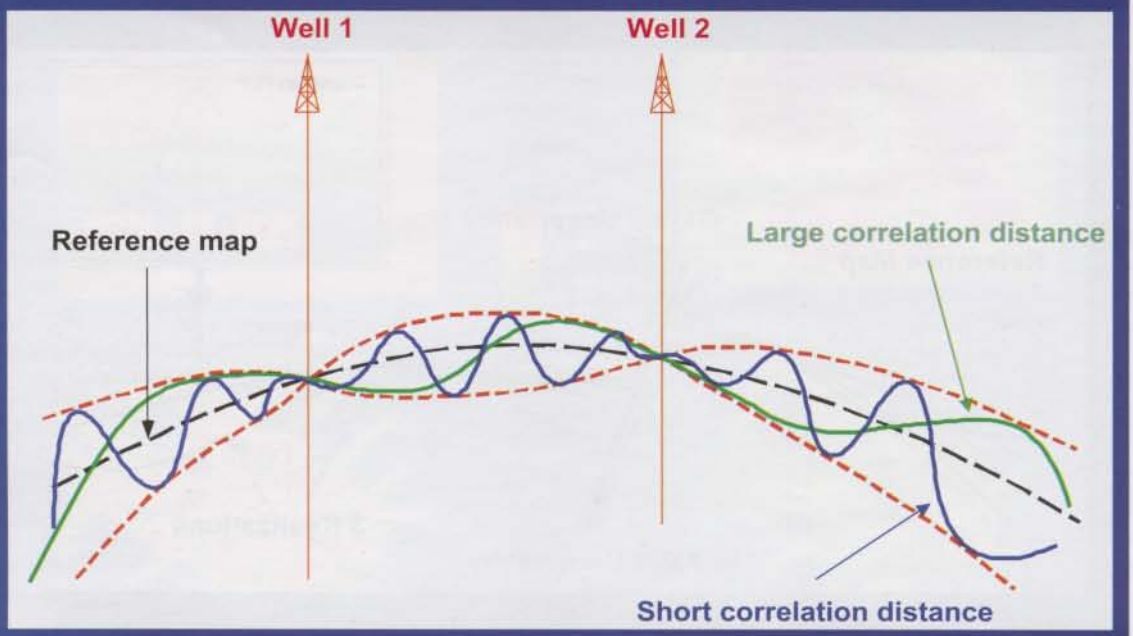
GOING FROM UNCERTAINTIES TO REALIZATIONS FLOWCHART (THORE ET AL., 2002)



SEG/EAGE DISC 2003

6-9

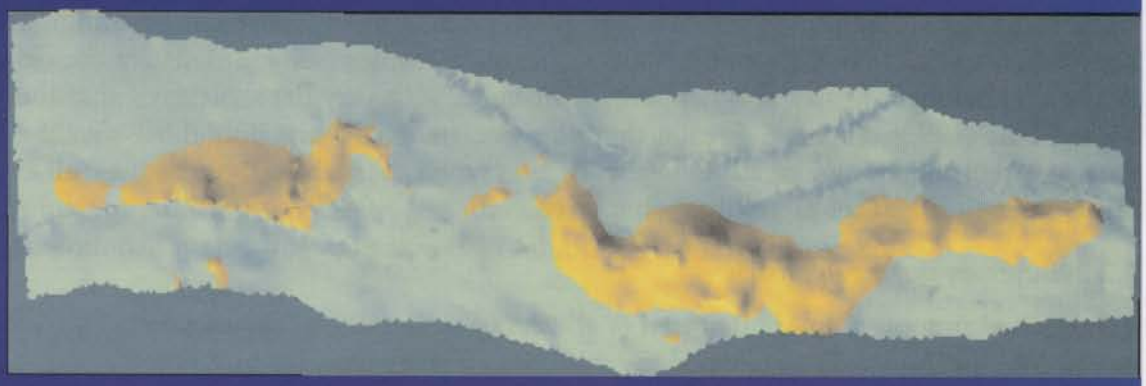
GOING FROM UNCERTAINTIES TO REALIZATIONS: IMPACT OF CORRELATION DISTANCE (SAMSON ET AL., 1996)



SEG/EAGE DISC 2003

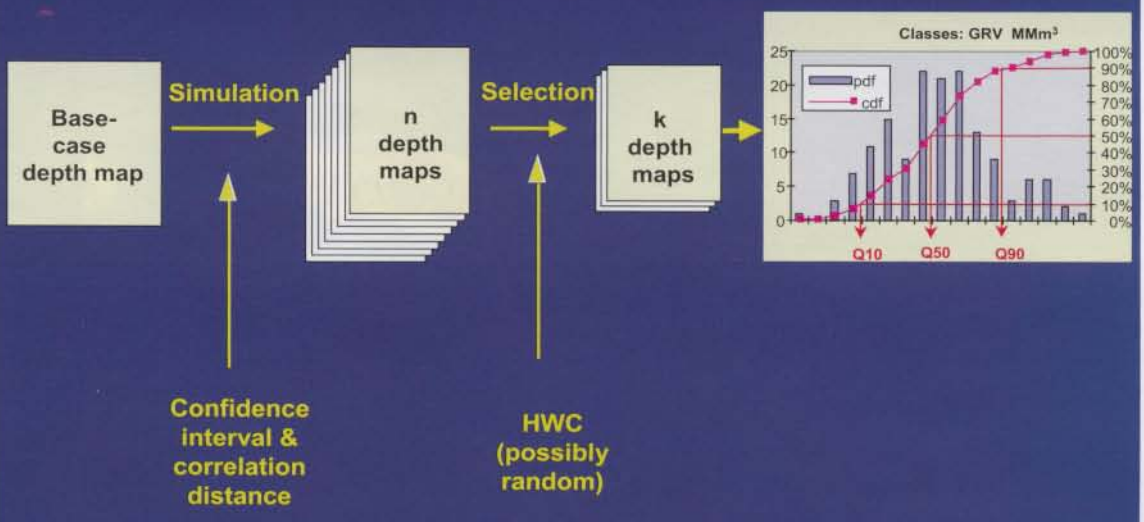
6-10

DIFFERENT REALIZATIONS OF THE TOP OF A TURBIDITE CHANNEL COMPLEX, GULF OF GUINEA



Visualization of the impact of the geophysical uncertainties on the reservoir shape and on the spatial fluid distribution is very important!

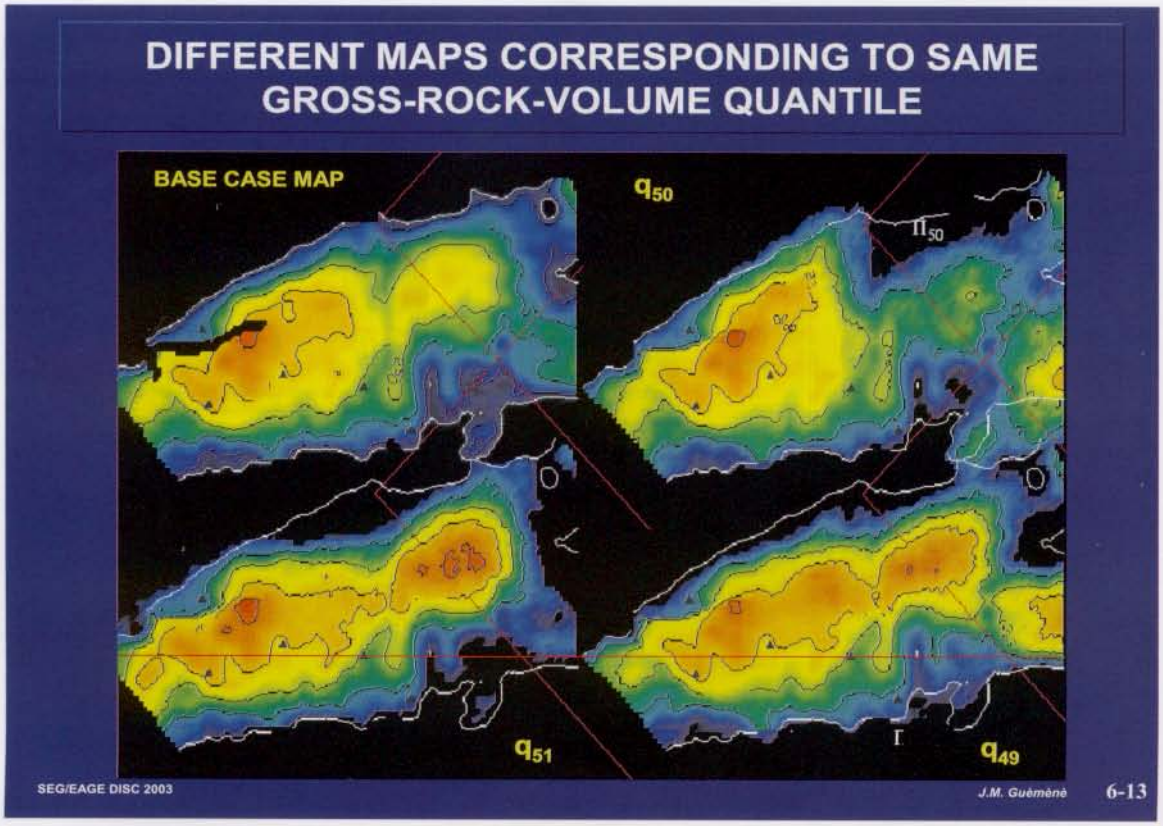
STRUCTURAL UNCERTAINTIES WORKFLOW (THORE ET AL., 2002)



This method of quantifying GRV uncertainties is quite fast, and hundreds of realizations can be generated. Some realizations will be very different from each other but will still be associated with the same GRV. For instance, Fig. 6-13 shows, on an actual case study, the base-case map and three realizations associated with a very similar volume, which happens to be the median of the GRV pdf. This means that there is no such thing as a single "median" or "most likely" map. There are only maps corresponding to the median value of the GRV pdf or to its q_{10} or q_{90} quantiles (see Fig. 1-24 for the definition of quantiles). It may also happen, for instance with very flat structures, that the mean of the GRV pdf is much greater than the base-case GRV. The only thing we are sure of is that the average of all the realizations is equal to the base case, because all the error realizations cancel each other when averaged.

6.2.2 *Bayesian kriging approach*

The method proposed by Abrahamsen et al. (2000) is just one of the most recent applications of a methodology developed over the last 10 years (see, for instance, Abrahamsen et al., 1991). The model is an extension of that already presented in Fig. 3-65. The main difference between this method and the previous one lies in the way the time and velocity uncertainty models are treated. Guéméné et al. (2002) consider the uncertainty maps for time and velocity to be an input of the method, in order to let the interpreter control the whole process. Then time and velocity realizations are simulated independently. Abrahamsen et al. do not use a velocity or time uncertainty map as



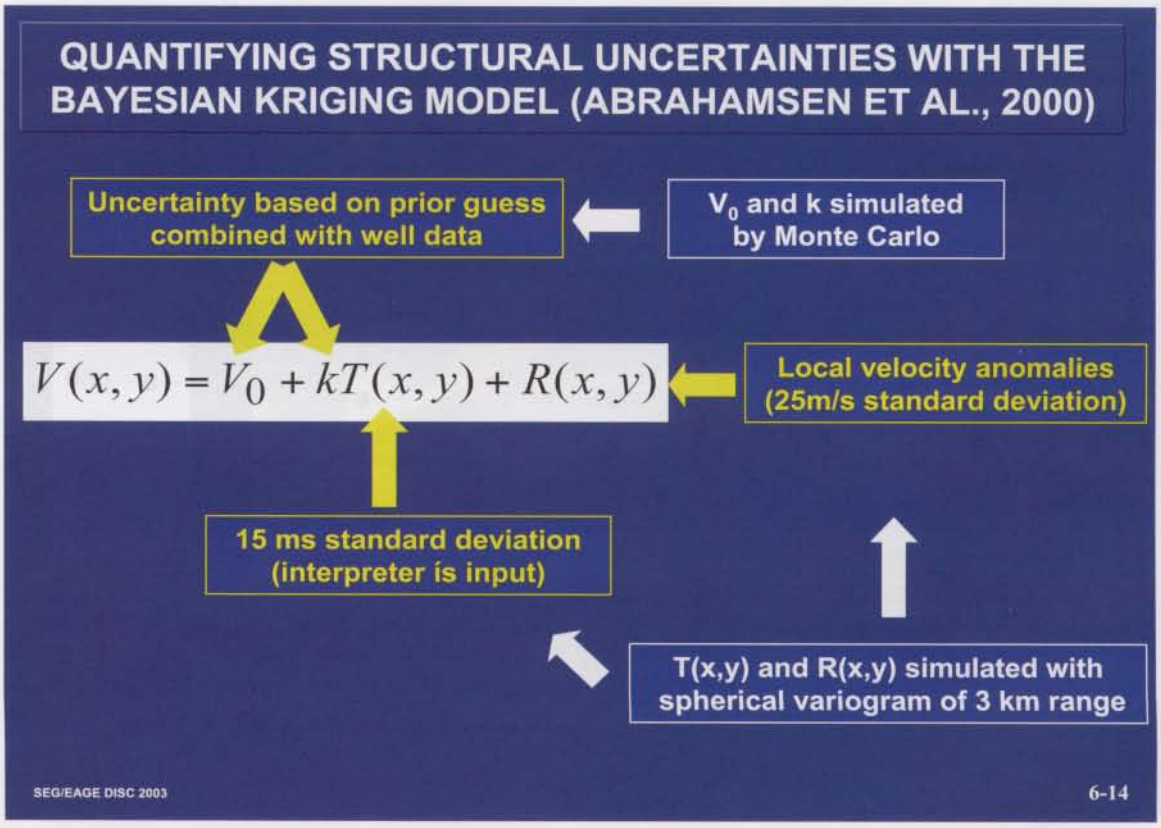
input. Their model incorporates a global (independent of location) standard deviation for time and for velocity anomalies (Fig. 6-14). Time realizations are simulated first, then velocity realizations are derived from time by sampling a velocity anomaly and a value for the compaction coefficients k and V_0.

Fig. 6-15 shows three realizations, which are ranked on the basis of their trapped GRV. The map on the left corresponds to the smallest trapped volume, whereas the map on the right does not show closure. The latter realization must be eliminated during the simulation exercise.

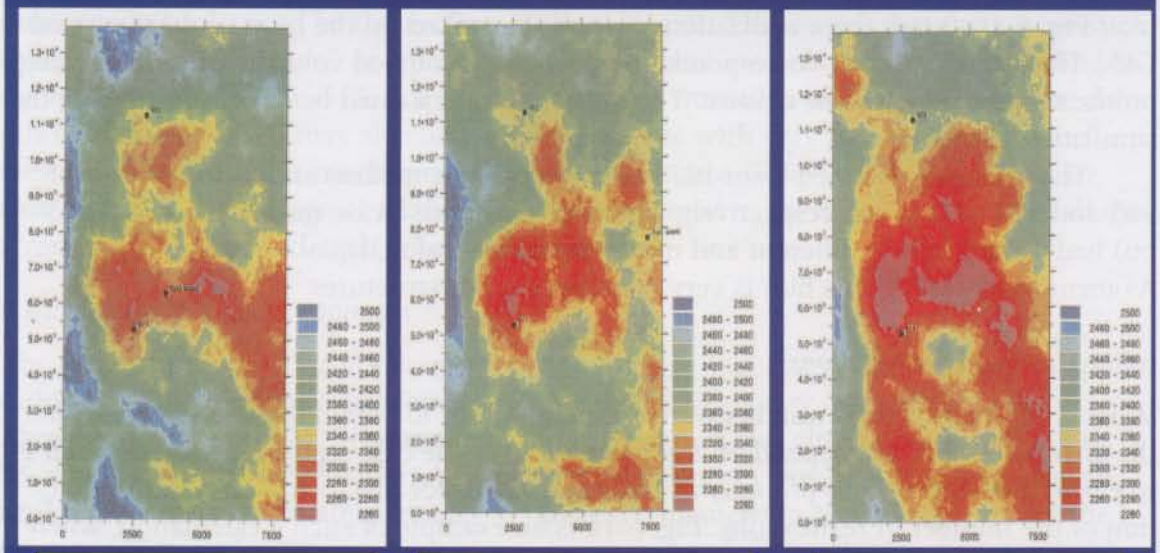
The GRV pdf obtained from 197 realizations had a median and mean volume of 965 and 880 million m^3, respectively. The GRV of the base-case map (center of Fig. 3-66) had a significantly different and more pessimistic value, equal to 652 million m^3. As mentioned earlier, this bias is very common for flat structures.

6.2.3 *How many realizations?*

What is the "reasonable" number of realizations to run to obtain reliable statistics on the GRV pdf? This will depend on the complexity of the case study. A good approach is to evaluate how the main property of interest — GRV, for instance — varies as a function of the number of realizations. Fig. 6-16 is an example of such a calculation, where we see that the quantiles become stable after a few hundred realizations.



NORTH SEA STRUCTURAL UNCERTAINTY QUANTIFICATION CASE STUDY (ABRAHAMSEN ET AL., 2000)



Depth realization with minimum trapped GRV

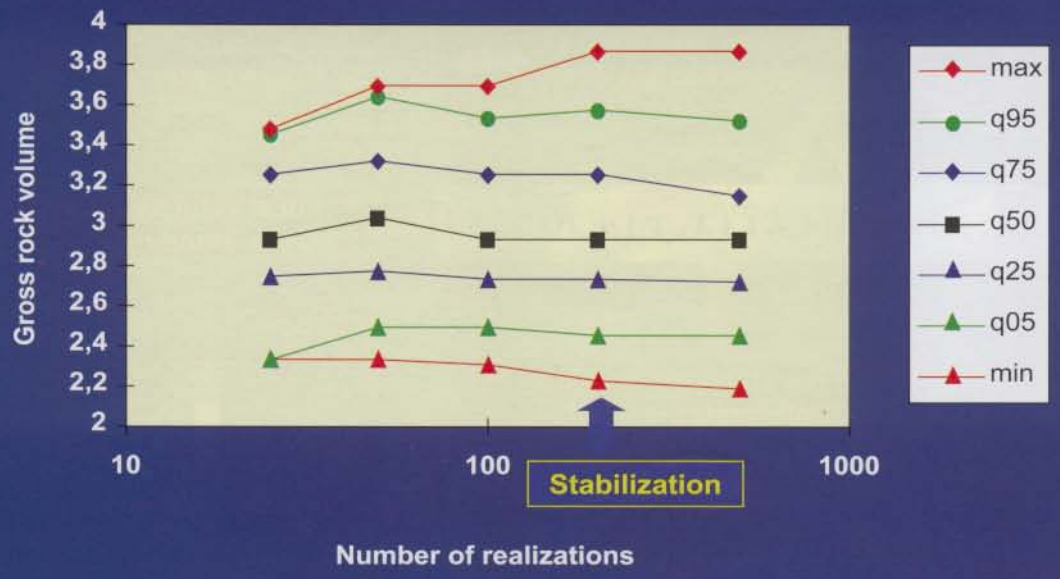
Depth realization with median trapped GRV

Unlimited depth realization

SEG/EAGE DISC 2003

6-15

HOW MANY REALIZATIONS?



SEG/EAGE DISC 2003

P. Biver 6-16

6.3 Static and Dynamic Model Uncertainties

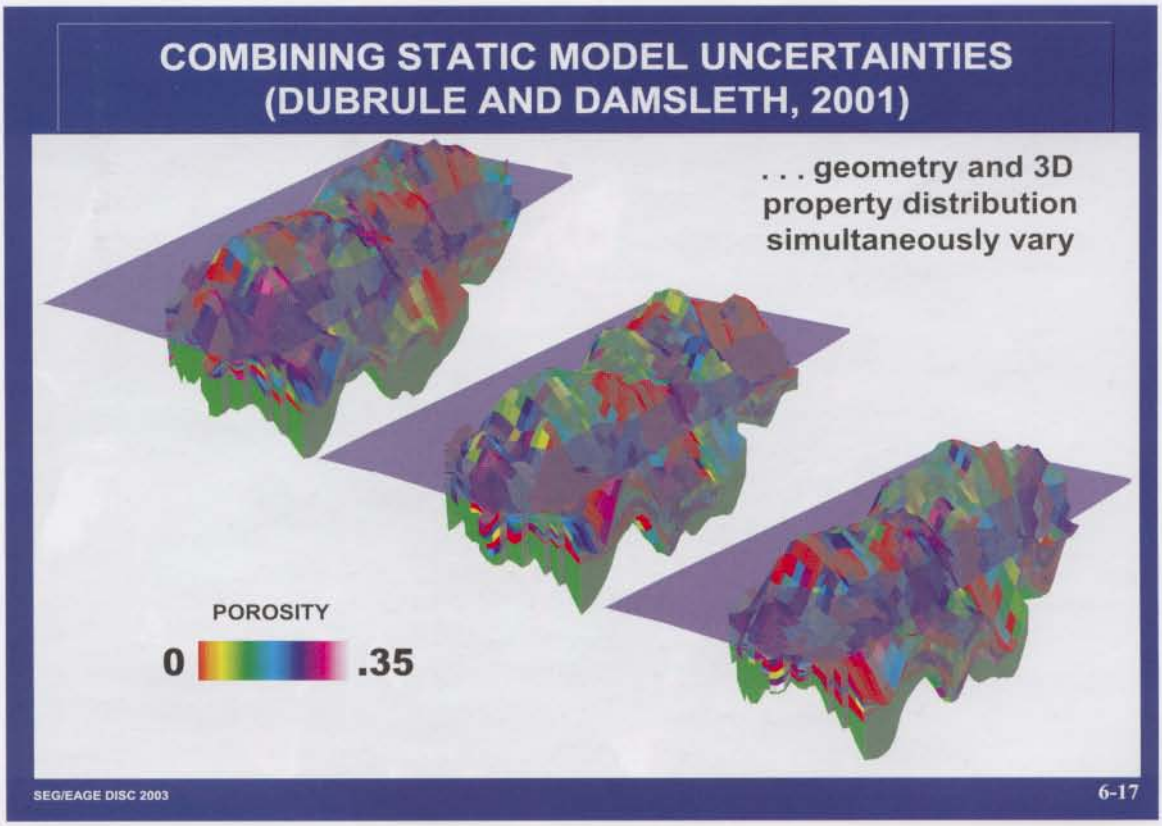
Once a number of realizations of the geometry have been generated, the next step will consist of evaluating other earth model uncertainties, in order to quantify their impact with relation to that of geometrical uncertainties (Fig. 6-17).

6.3.1 Static model uncertainties

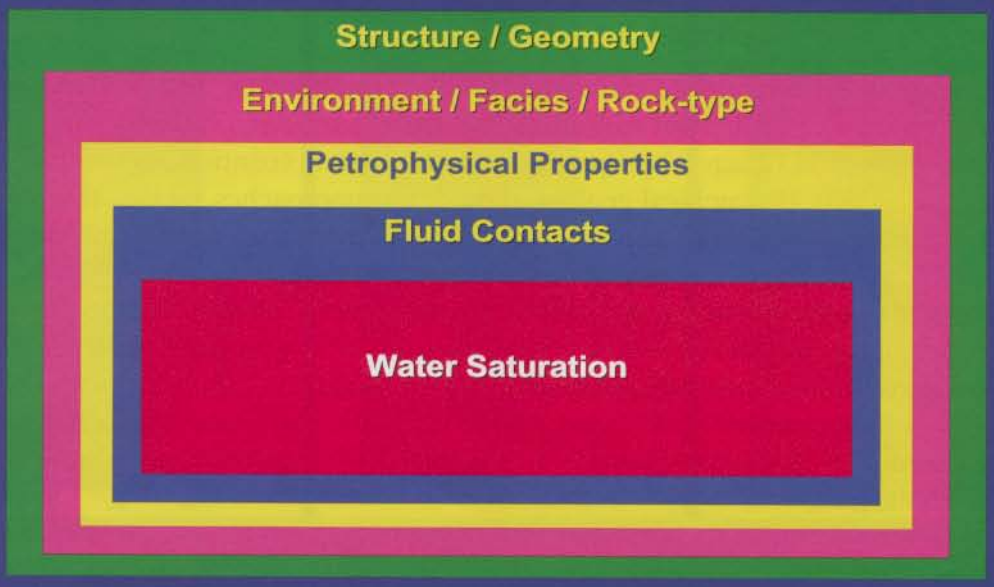
New generation earth modeling software offers the possibility to "hang" a stratigraphic grid from each new geometrical realization and fill this stratigraphic grid with 3D realizations of geological (discrete) properties or petrophysical (continuous) properties (Corre et al., 2000). Hierarchical geological modeling approaches such as those discussed in Section 4.5.5 can be applied. Realizations may or may not be constrained by seismic data, using one of the approaches described in chapter 4 or 5.

The hydrocarbon-water contacts (HWC) may also be affected by uncertainties, for example, if we are dealing with an "oil-down-to" or a "water-up-to" situation. Once the HWC have been positioned, water-saturation realizations consistent with these contacts can be generated. Thus, quantification of static earth model uncertainty is usually based on a hierarchical approach (Fig. 6-18) from geometrical to water-saturation uncertainties.

Fig. 6-19 is an example of six realizations resulting from an actual North Sea case study. Just one well was available to construct the model, and each realization is based on a different structural map, a different HWC, and a different 3D porosity distribution.



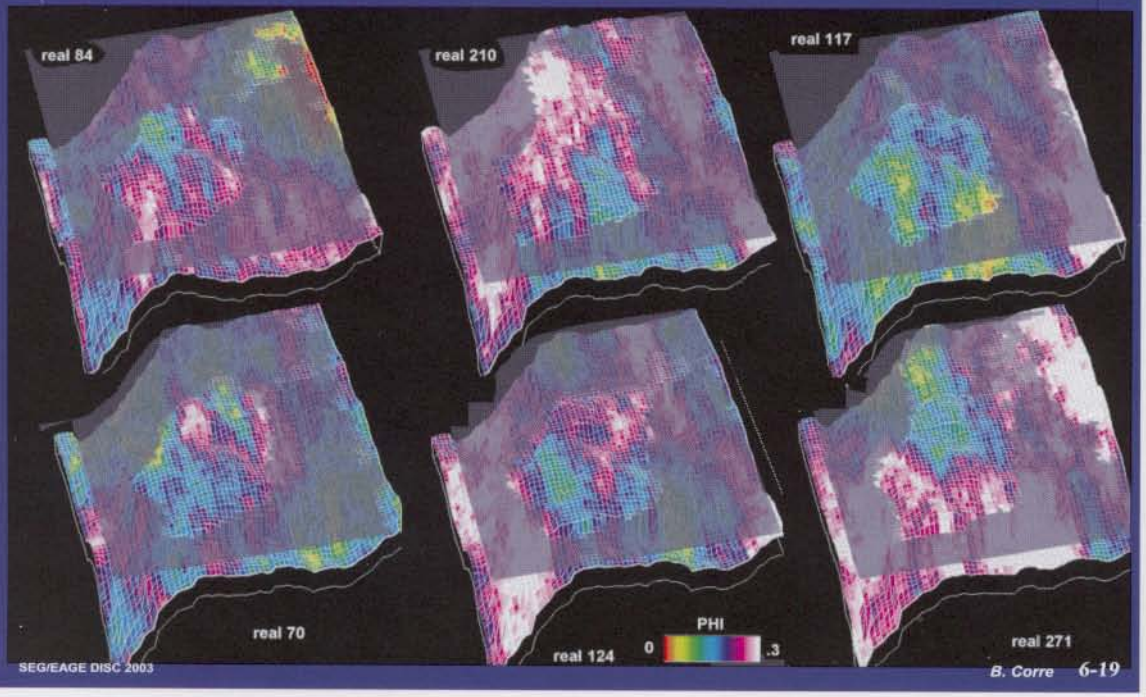
COMBINING STATIC MODEL UNCERTAINTIES (CORRE ET AL., 2000)



SEG/EAGE DISC 2003

6-18

GEOSTATISTICAL REALIZATIONS IN 3D EARTH MODEL



SEG/EAGE DISC 2003

B. Corre 6-19

Figs. 6-20 and 6-21 are two realizations from an uncertainty quantification study on a turbidite channel complex in the Gulf of Guinea. Four variables (N/G, V_{shale}, porosity, and permeability) are generated for each realization and are strongly correlated with each other. The structure map is also different for each realization.

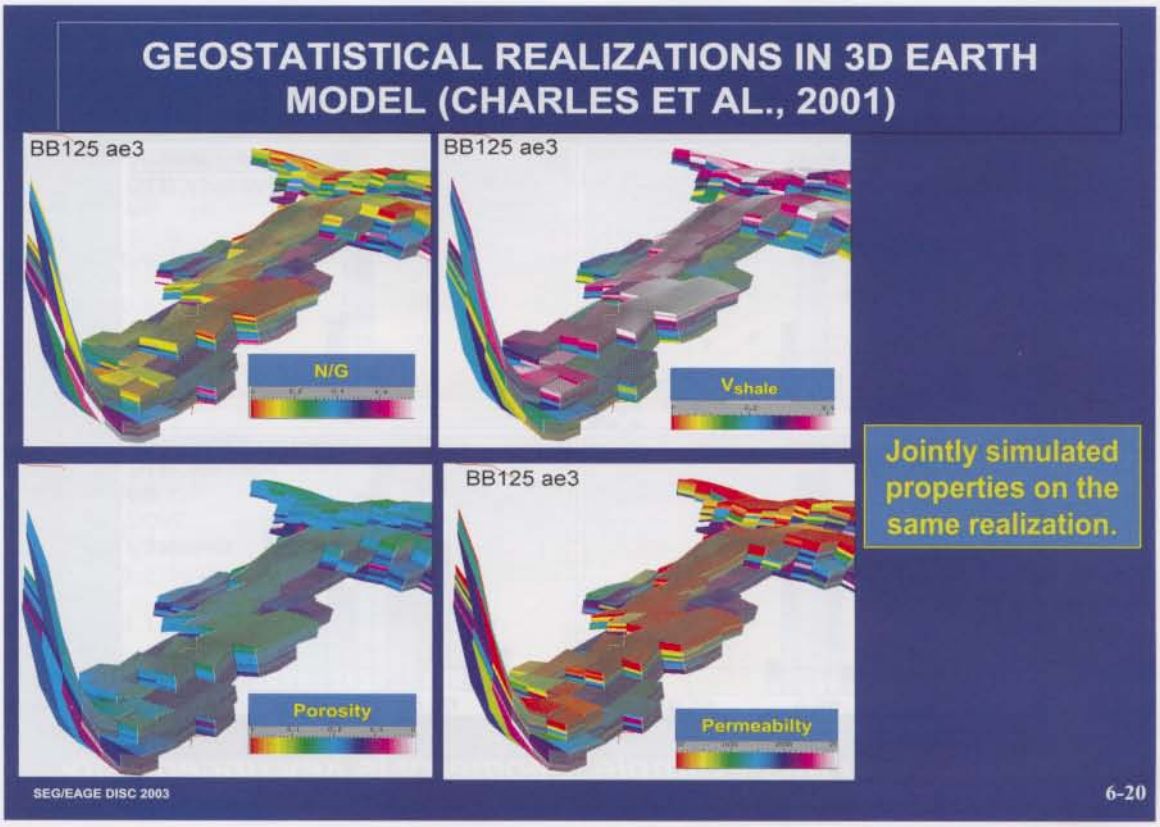
A number of parameters of interest can be calculated from such earth model realizations. An important one is the oil-in-place volume (OIP). This volume will be different for each realization, and the spread of the OIP pdf will represent the impact of all individual uncertainties on the final volume.

In Fig. 6-22 we have histograms, corresponding to an actual case study, that illustrate the power of such an uncertainty study. As we move from left to right and from top to bottom, more and more uncertainties are taken into account. It is clear from the example that most of the uncertainty comes from the geometry, as is often the case when few wells are available and the seismic is not top quality. Following such a study, the decision might be to drill an extra appraisal well in order to reduce the structural uncertainty.

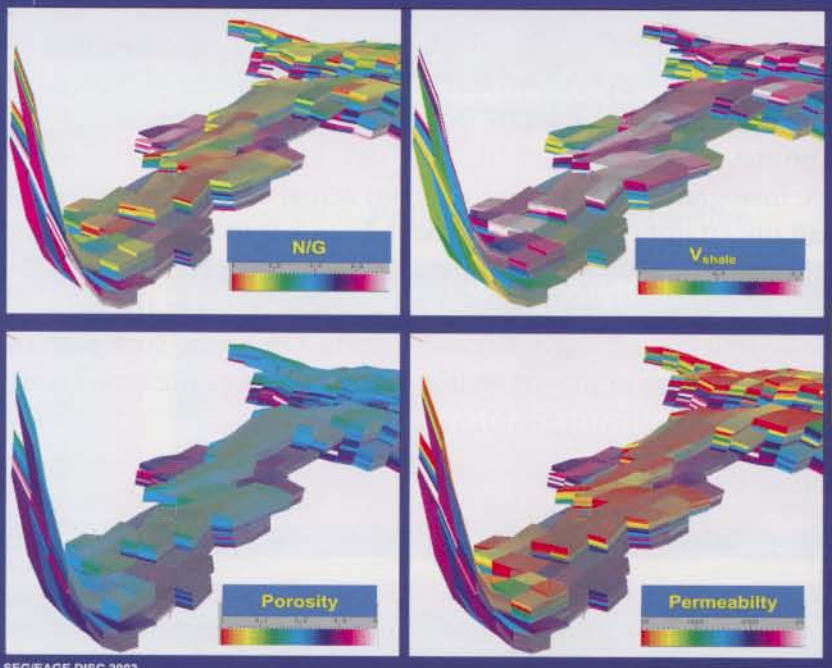
6.3.2 Link with dynamic flow simulation

- Impact of static-model uncertainties on predicting production profiles

In the Fig. 6-22 example, only OIP uncertainties were addressed. We may also wish to quantify the uncertainty on reserves by running flow simulations on a number of realizations. When we discussed geometrical uncertainties, we saw that a minimum or



GEOSTATISTICAL REALIZATIONS IN 3D EARTH MODEL (CHARLES ET AL., 2001)

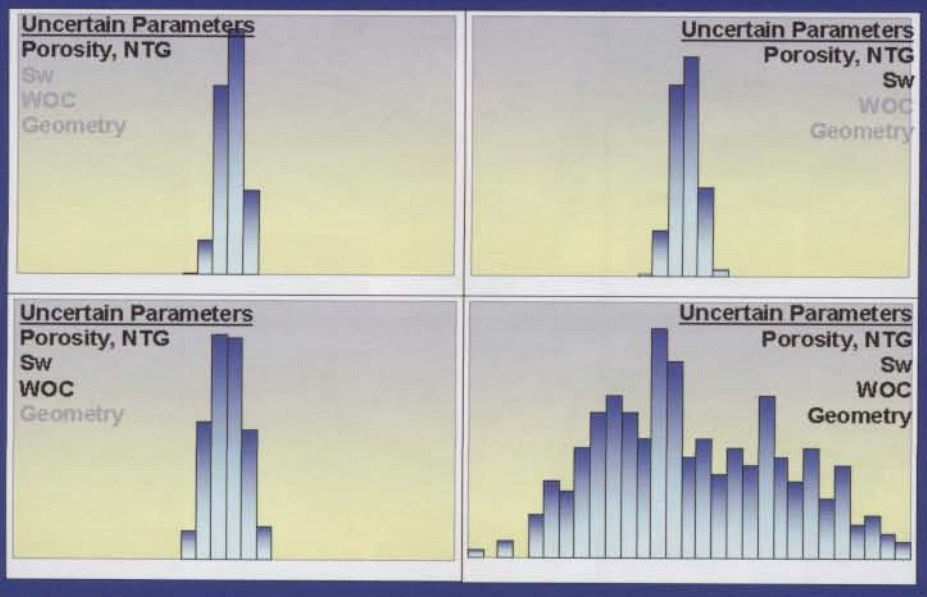


Jointly simulated properties on the same realization.

SEG/EAGE DISC 2003

6-21

EXAMPLE OF INTEGRATED UNCERTAINTY ANALYSIS



In this example, geometry is key uncertainty

SEG/EAGE DISC 2003

G. Vincent

6-22

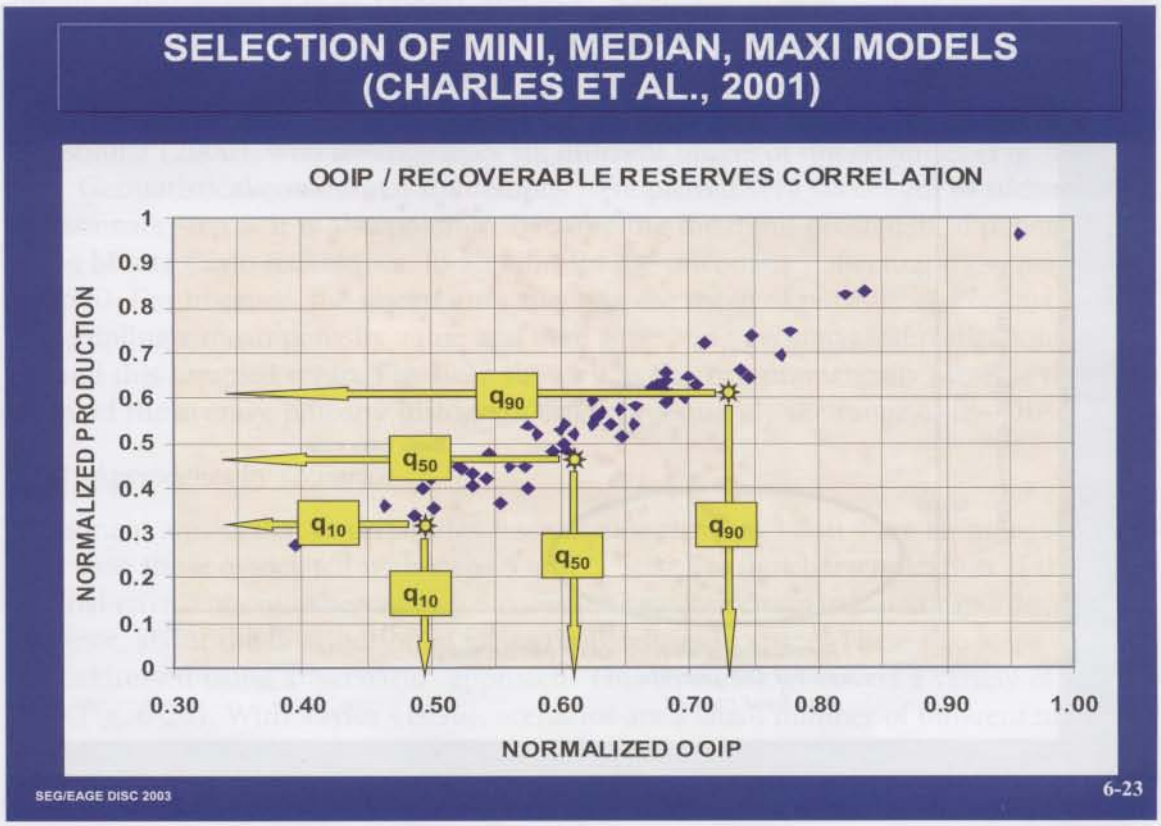
maximum realization could only be defined with relation to a parameter of interest, such as GRV or OIP. A more powerful way to pick such a realization is to use the cross-plot of reserves versus OIP (Fig. 6-23) and pick the realizations that position themselves as quantiles of interest on both the OIP and reserves distributions.

- Accounting for dynamic-parameter uncertainties

So far, we have only discussed the impact of static-model uncertainties on production profiles. Dynamic-parameter uncertainties must also be accounted for, in order to address such questions as: What is the impact of the uncertainty affecting fault transmissibility? or What is the impact of relative permeability uncertainty? Usually, dynamic parameters are easier to address, because they are real numbers, such as irreducible water saturation or the transmissibility multiplier. Damsleth et al. (1992) and Corre et al. (2000) explain how to use the statistical method of experimental design to derive information on the uncertainty in production profiles that is caused by dynamic-parameter uncertainties.

- Addressing structural uncertainties in the history-matching process

The applications of the approach described above go far beyond the quantification of uncertainties. They can also boost multidisciplinary integration, all the way to the history-matching exercise. Until recently, the construction of the reservoir model was very sequential, and the structural map was never used as a matching parameter, in



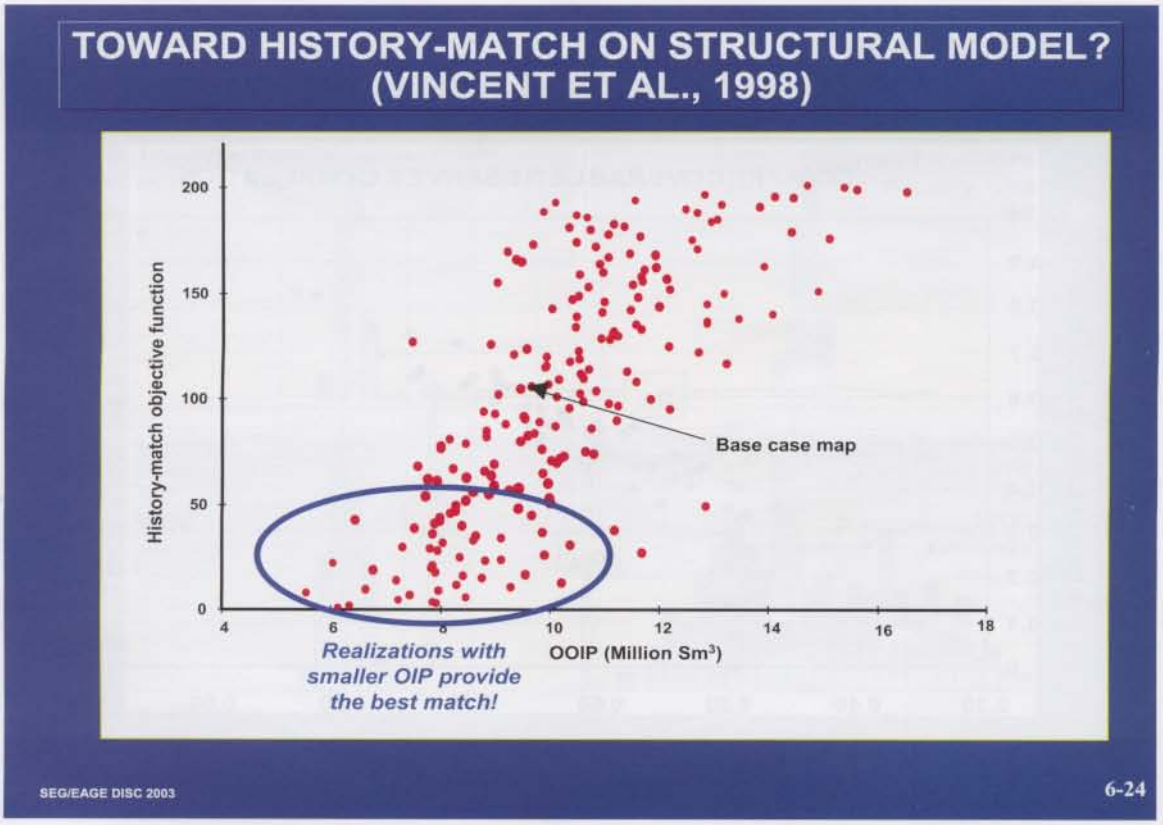
spite of the fact that structural uncertainties were known to have a major impact on results. This is now changing. In the example of Fig. 6-24 (Vincent et al., 1998), dynamic-flow simulations have been run on a large number of earth model realizations. Each point of the crossplot is the result obtained with one realization. Clearly, those realizations corresponding to a smaller value of OIP are those that provide the best match with production data. Thus, in this specific example, those realizations associated with deeper geometrical models are more likely to be representative of the actual unknown surface.

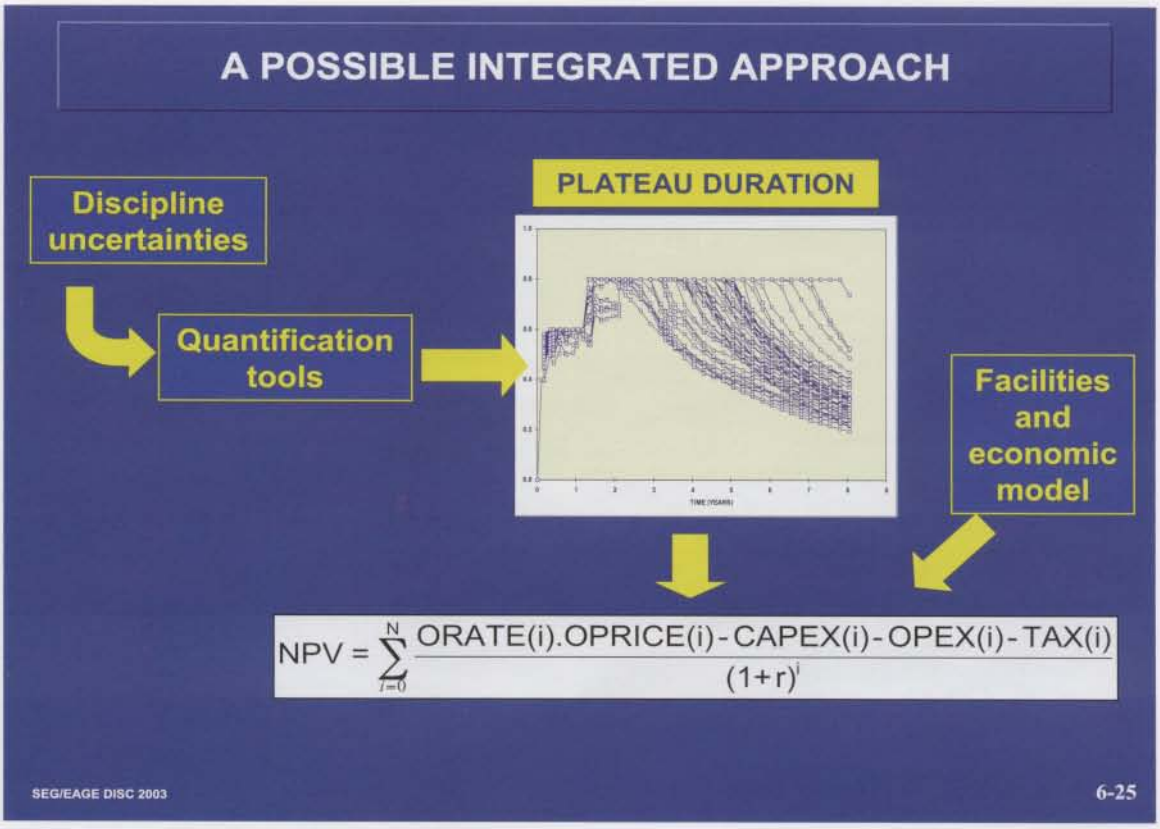
- History matching under uncertainty

Hegstad and Omre (2001) formalize the problem of history matching under uncertainty, using the Bayesian framework. They show how realizations of the posterior earth model — that is, the model constrained by all data, including wells, production data, and seismic — can be obtained using the Metropolis-Hastings algorithm. Although their solution appears to be quite general and has been tested successfully, it remains very cumbersome to run. As computing power develops in the future, it may prove more and more interesting.

6.4 Multirealization Uncertainty-quantification Approach: A Panacea?

On the basis of what has just been discussed, one would be tempted to generalize the approach into "multirealization economic analysis" (Fig. 6-25), as discussed, for



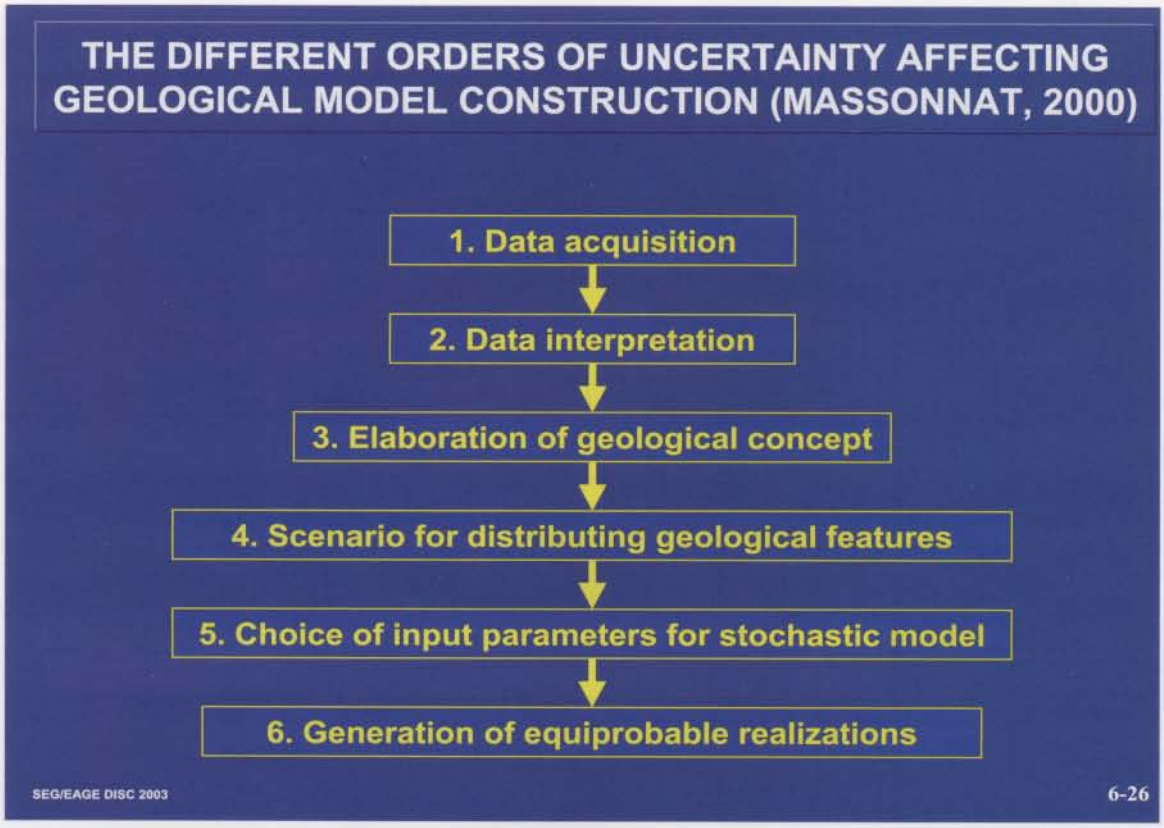


instance, by Ovreberg et al. (1992). However, a number of geoscientists remain sceptical about uncertainty quantification via geostatistical multirealization techniques (Dubrule et al., 1996). Many argue that actual uncertainties are of a different order from those that are quantified by geostatistical realizations. This point has been discussed by Massonnat (2000), who distinguishes six different orders of uncertainties (Fig. 6-26).

Geostatistical conditional simulations have proved very successful in addressing Massonnat's step 6. It is also possible, by sampling the input geostatistical parameters using Monte Carlo techniques, to account for the uncertainty affecting these parameters (step 5). For instance, the uncertainty affecting the mean of porosity can be quantified by sampling a mean porosity value and then generating geostatistical realizations around this sampled mean. Fig. 6-27 shows that this can dramatically increase the range of the average porosity histogram (and consequently the range of the OIP pdf).

6.4.1 Approaches by scenarios

Massonnat argues that uncertainties associated with steps 3 and 4 are far more significant than those associated with steps 5 and 6. Step 3 is the characterization of the depositional environment, whereas step 4 covers the major stationarity assumptions, for instance, about the distribution of facies probabilities in space. These two steps are usually addressed using a "scenario" approach. This terminology covers a variety of methods (Fig. 6-28). With Taylor (1996), scenarios are a small number of different models associated with possible assumptions — usually min, median, max — about some of the input parameters, such as N/G, sand body dimensions or fault sealing (steps 4 and



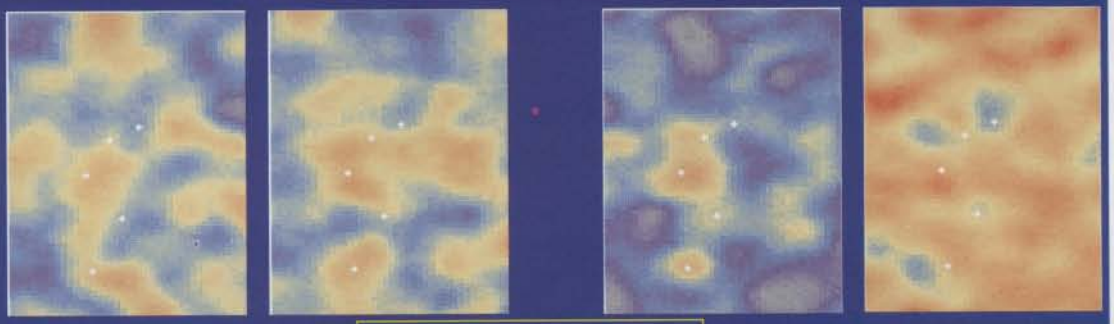
5). Dubrule and Damsleth (2001) discuss a possible approach where different scenarios are obtained from different geoconcepts (steps 3 and 4). Corre et al. (2000) associate scenarios with different assumptions (Fig. 6-29) about the size and number of sand bodies (steps 4 and 5).

Scenarios differ from geostatistical conditional simulations in that the former are discrete (leading to multimodal discrete histograms), whereas the latter result in a continuum of models (leading to unimodal continuous histograms). Usually, the number of scenarios is rather small, from two to, say, twenty, which allows the geoscientist to better control the geological meaning of each possible representation of the reservoir. In some cases, the approach by discrete scenarios is mandatory, because we deal with clearly different alternatives (e.g., choice of depositional environment, choice of seismic loop on an interpretation) with no possible intermediate situations. However, constructing multiple, discrete scenarios may be very time-consuming. The use of multiple, discrete scenarios may also create the illusion that only two or three situations are possible, when in reality all the intermediate ones may also happen. In this case, it may be better to sample geostatistical parameters in such a way that the discrete scenarios are incorporated into a range of realizations. The sampling approach, based on Monte Carlo and geostatistics (Fig. 6-27), can be more efficient to run, because it does not require a manual reconstruction of the model.

6.4.2 Combining scenarios and geostatistical realizations

Still, if we come back to Massonnat's (2000) classification, the discrete scenario-based

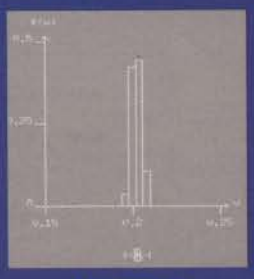
ACCOUNTING FOR UNCERTAINTY ON THE MEAN



without uncertainty on the mean

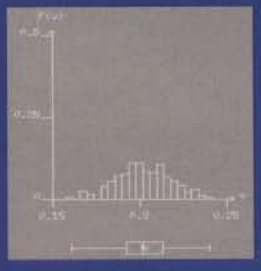
Φ MAPS (scale: 0.1 to 0.3)

with uncertainty on the mean



Average Φ

Uncertainty on the mean crucial if few data (5 wells) and small correlation distances



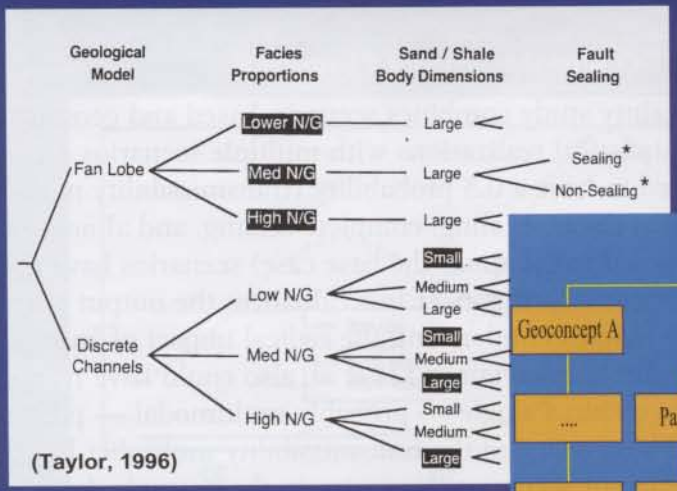
Average Φ

P. Biver

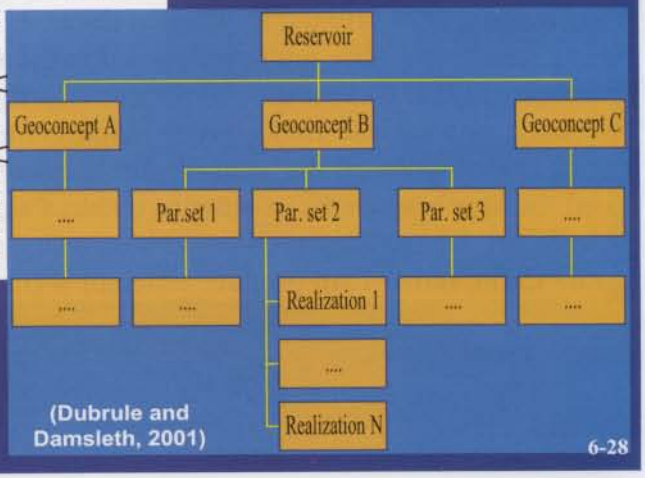
SEG/EAGE DISC 2003

6-27

APPROACHES BY SCENARIOS



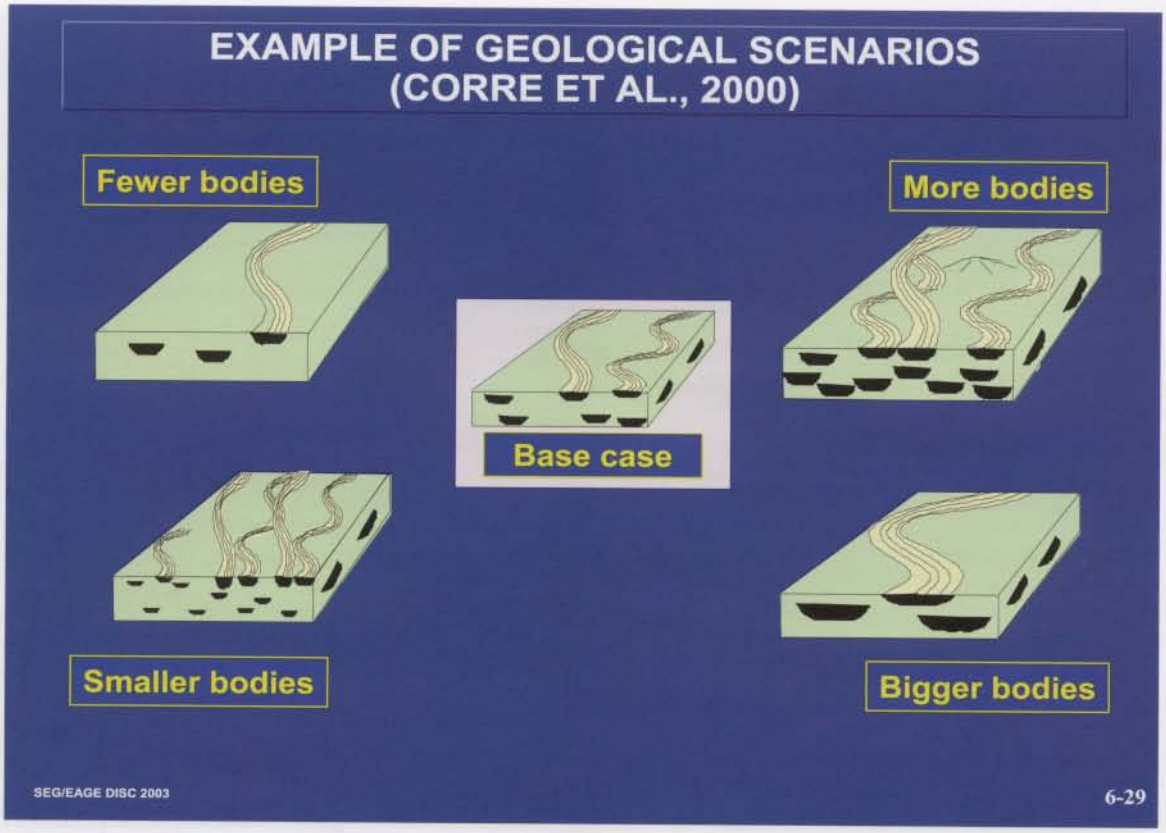
(Taylor, 1996)



(Dubrule and Damsleth, 2001)

SEG/EAGE DISC 2003

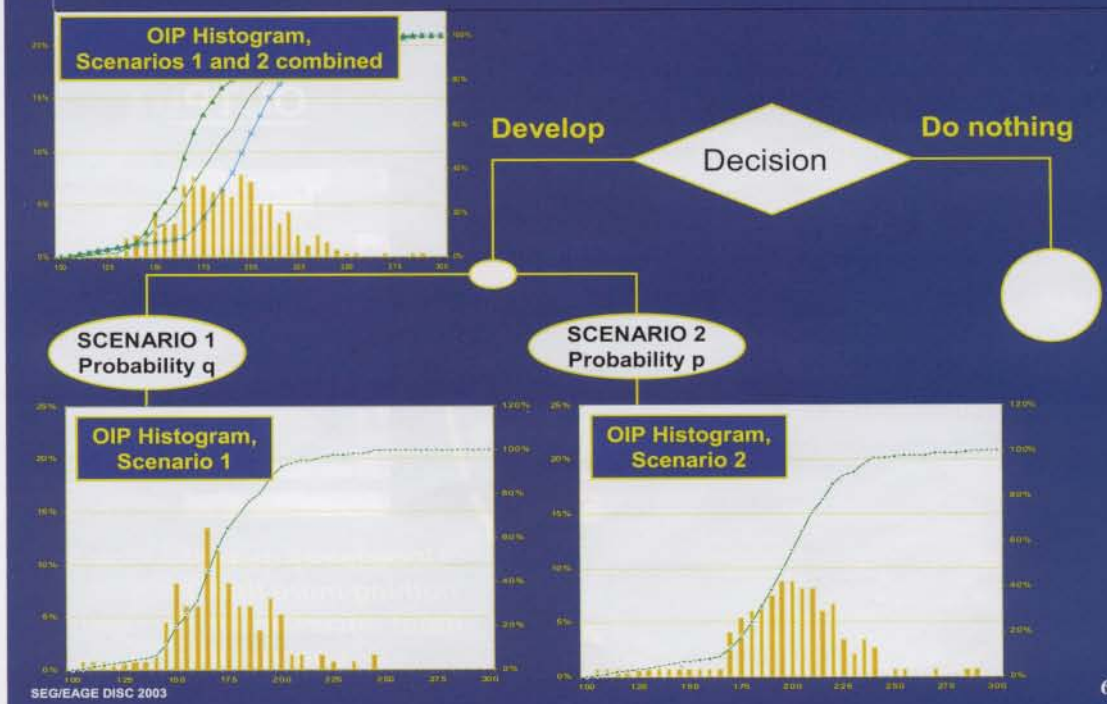
6-28



approaches for steps 3 and 4 can be combined with geostatistical ones for steps 5 and 6, as described in Fig. 6-30, where multiple geostatistical realizations have been produced for each scenario and then recombined using the probability of each scenario occurring. This requires a quantification of this probability, which may be a difficult but useful exercise. Corre et al. (2000) and Charles et al. (2001) discuss a number of applications of uncertainty quantification combining scenarios and geostatistical realizations.

Lia et al.'s (1997) classic uncertainty study combines scenario-based and geostatistical approaches. They combine geostatistical realizations with multiple scenarios on fault sealing. The base case is assumed to have a 0.5 probability (transmissibility multipliers ranging from 0 to 0.05), whereas the no-sealing, complete-sealing, and almost-sealing (transmissibility multipliers equal to 0.1 times the base case) scenarios have a probability of 0.2, 0.15, and 0.15, respectively. Fig. 6-31 individualizes the output pdf associated with each scenario, which helps us understand the radical impact of fault-sealing assumptions on the recovery-factor uncertainty. Lia et al. also could have recombined the four histograms in order to obtain a single — probably multimodal — pdf for the recovery factor. They even could have sampled the transmissibility multiplier by Monte Carlo, rather than produce different scenarios. However, in the example, keeping the scenarios separate helps us understand the actual impact of fault sealing. So, the choice of working by scenarios or by Monte-Carlo sampling of some of the input parameters may depend on the geological meaning and the sensitivity of the final results to these parameters.

COMBINING SCENARIOS AND GEOSTATISTICS

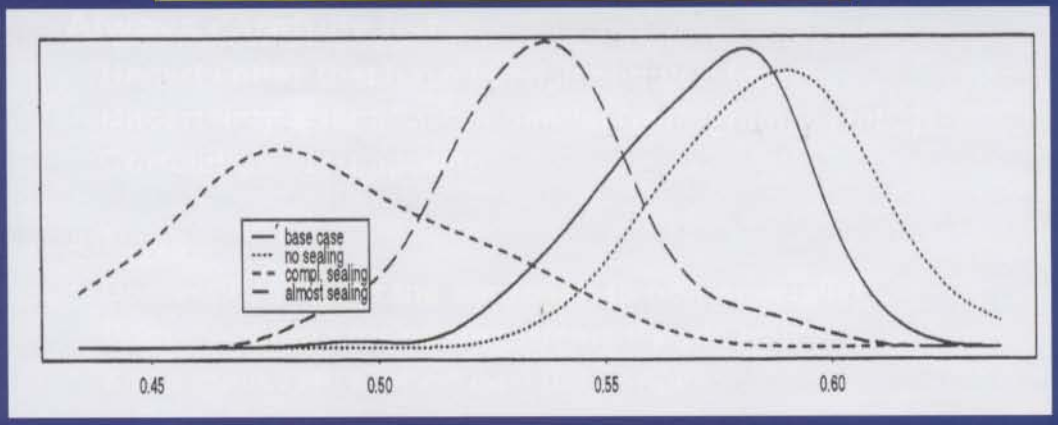


SEG/EAGE DISC 2003

6-30

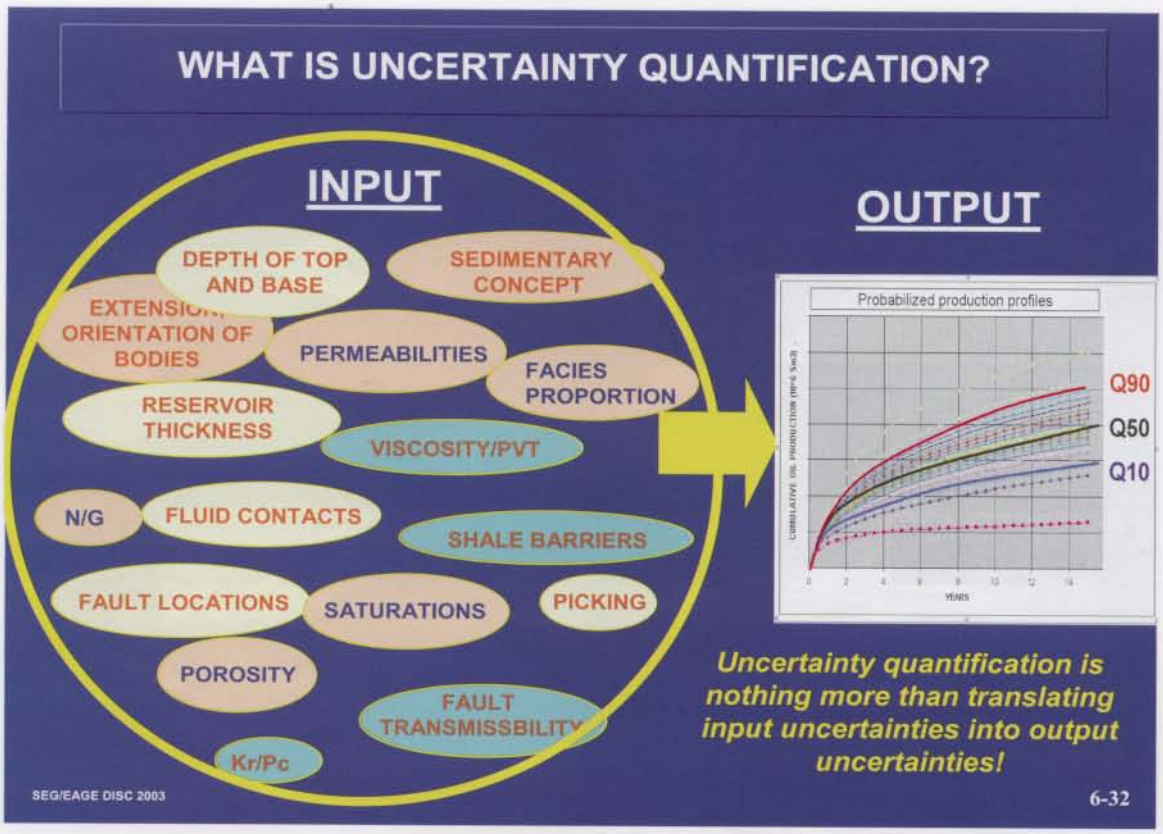
THE IMPACT OF FAULT SEALING SCENARIO ON RECOVERY FACTOR (LIA ET AL., 1997)

Recovery Factor, Mobile Oil (after 20 years)



SEG/EAGE DISC 2003

6-31



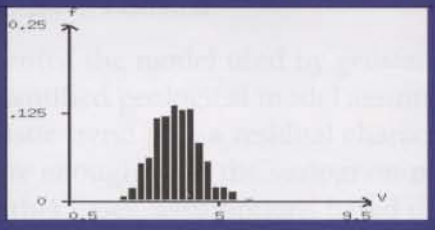
6.5 Conclusion on Uncertainties: A Word of Caution

The approaches presented above are merely a way to combine the uncertainties quantified by the various disciplines, in order to evaluate their joint impact on parameters that have an economic importance (Fig. 6-32). If some important uncertainties are ignored in the input phase, the uncertainty evaluation, whether quantified with geostatistics or with a scenario approach, will be wrong.

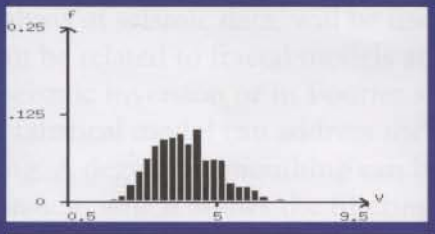
As with Monte-Carlo simulation, the importance of the assumptions made about parameter correlation cannot be stressed too much. As mentioned before, making no assumption about the correlation between parameters is the same as assuming independence. We saw earlier that independence means mutual compensation, which usually leads to an underestimation of global uncertainties (Fig. 6-33). The guidelines provided in Fig. 6-34 must absolutely be followed to avoid transforming uncertainty quantification into a black box.

The main value of uncertainty quantification may lie not in the results — the final pdfs — but in the process that it involves among a team of geoscientists and possibly other disciplines. This is nicely expressed by Spencer et al. (1998): "By having a clear expectation of the range of parameters expected, clear accountability of the estimates, a feedback process which trains us in our judgements, and an atmosphere in which the sharing of mistakes and learning is encouraged, the Company benefits in multiple ways."

IMPACT OF CORRELATION ASSUMPTION ON UNCERTAINTY

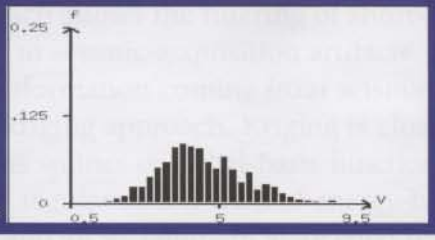


OIP histogram (MMm^3) if no correlation



OIP histogram (MMm^3) if intermediate correlation

$\rho(ntg\text{-}\phi)= .6$, $\rho(Swi\text{-}\phi)=-.5$ $\rho(Logk\text{-}\phi)=.7$



OIP histogram (MMm^3) if perfect correlation

RECOMMENDATION

Always carefully list:

- Uncertainties that have been accounted for
- Uncertainty quantification model (variogram, trend...)
- Correlation assumptions

Allows:

- Demystification of the approach to non-specialist
- Documentation for post-mortems ("audit trail")
- Keep model alive when new data appear

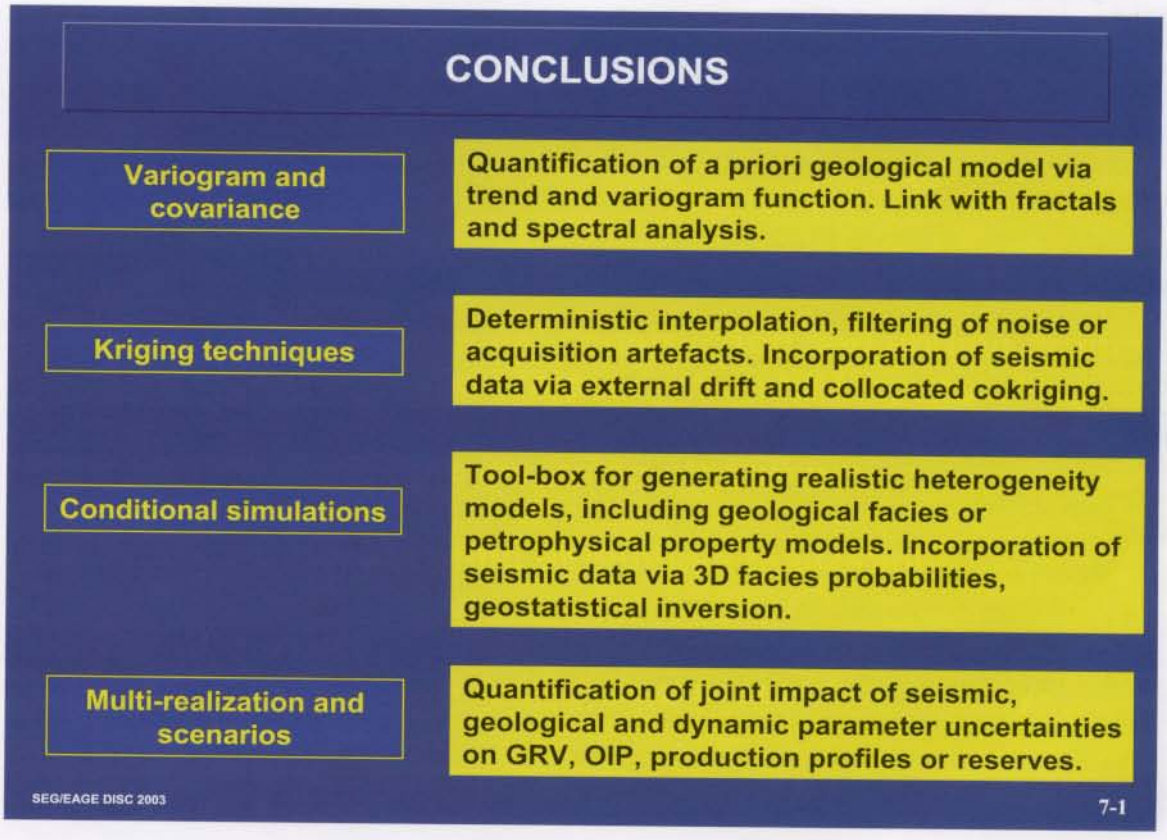
7 Conclusions

7.1 What We Have Learned

We have presented the model used by geostatistics for 2D and 3D petroleum applications. This quantified geological model assumes that the variable of interest is the sum of a deterministic trend plus a residual characterized by its variogram or its covariance. When there are enough data, the variogram model can be fitted to the experimental function. In other cases, assumptions based on geological knowledge, combined with a variogram analysis of seismic data, will be used to define the variogram model. Variograms can be related to fractal models and to a priori information used by geophysicists in seismic inversion or in Fourier analysis (Fig. 7-1).

The geostatistical model can address the problem of deterministic interpolation through kriging. A degree of smoothing can be applied to kriging through the error-cokriging approach, which allows the filtering of random noise, whereas the factorial kriging approach allows the filtering of short-range — or high-frequency — terms due (for instance) to seismic-acquisition artifacts. Kriging based on well data can also incorporate extra information coming from seismic data, through the external drift or the collocated cokriging approach. Kriging is closely related to other interpolation techniques, such as splines or radial-basis functions. Specifying a kriging model amounts to specifying the regularization term of energy-based inversion techniques.

Kriging and all its family of associated techniques remain a deterministic method.



The production of a minimum variance estimate results in an interpolation that is very smooth away from the data points. In practice, however, geology has no reason to become smoother away from the wells! The method of conditional simulation allows the generation of models that show everywhere, including away from wells, a similar degree of variability, as quantified by the variogram. All kriging techniques presented here can be generalized into continuous parameter conditional simulation, thanks to algorithms such as SGS.

Conditional simulation can also be used to generate models of discrete variables, such as depositional facies, lithology, or rock type. The most popular techniques are object-based models and indicator simulation. They can be applied jointly, in a hierarchical manner, and combined with continuous parameter-simulation techniques, using tools available today in earth modeling software. Indicator simulation realizations can be constrained by 3D probability models derived from seismic data using the Bayes theorem and SIS. Object-based models can also be constrained by 3D probability models, but this requires the use of MCMC algorithms.

Conditional simulation can be generalized into geostatistical inversion, a sequential technique that allows the generation of 3D acoustic-impedance realizations constrained by seismic data. The realizations can be produced at the reservoir model scale, that is, at higher resolution than the seismic data. The lower the resolution of the seismic data, the more variability there is from one inverted realization to another. The acoustic-impedance realizations can then be used to constrain 3D realizations of lithology or petrophysical parameters.

Joint conditional simulation of the different features of a 3D earth model can also lead to quantification of the 3D earth model's uncertainty. Realizations of the geometrical model, accounting for the uncertainty affecting interpretation picks and time-to-depth conversion, are combined with realizations of the geological model, including lithology and petrophysical parameters. This leads to the quantification of the GRV, OIP, or reserves uncertainty resulting from uncertainties affecting the static model. The impact of the uncertainties affecting dynamic parameters can be quantified using experimental design. Scenario-based and geostatistical approaches can be combined, as long as a probability is attached to each possible scenario.

7.2 Future Topics

The nature of geostatistical research has changed since the 1980s. At that time, petroleum geostatistics was still new, and there was an explosion of new techniques. Some of them proved difficult to understand or to apply, while many others — in spite of the jargon used — proved redundant with each other. Now, the dust has settled and a natural selection process has occurred. Geostatistical modeling has become part of the standard reservoir-modeling workflow, and the interest of geologists, geophysicists, and reservoir engineers in these techniques has grown. Instead of considering the techniques to be remote from their day-to-day activities, geoscientists now understand that they may have a strong impact on their workflow.

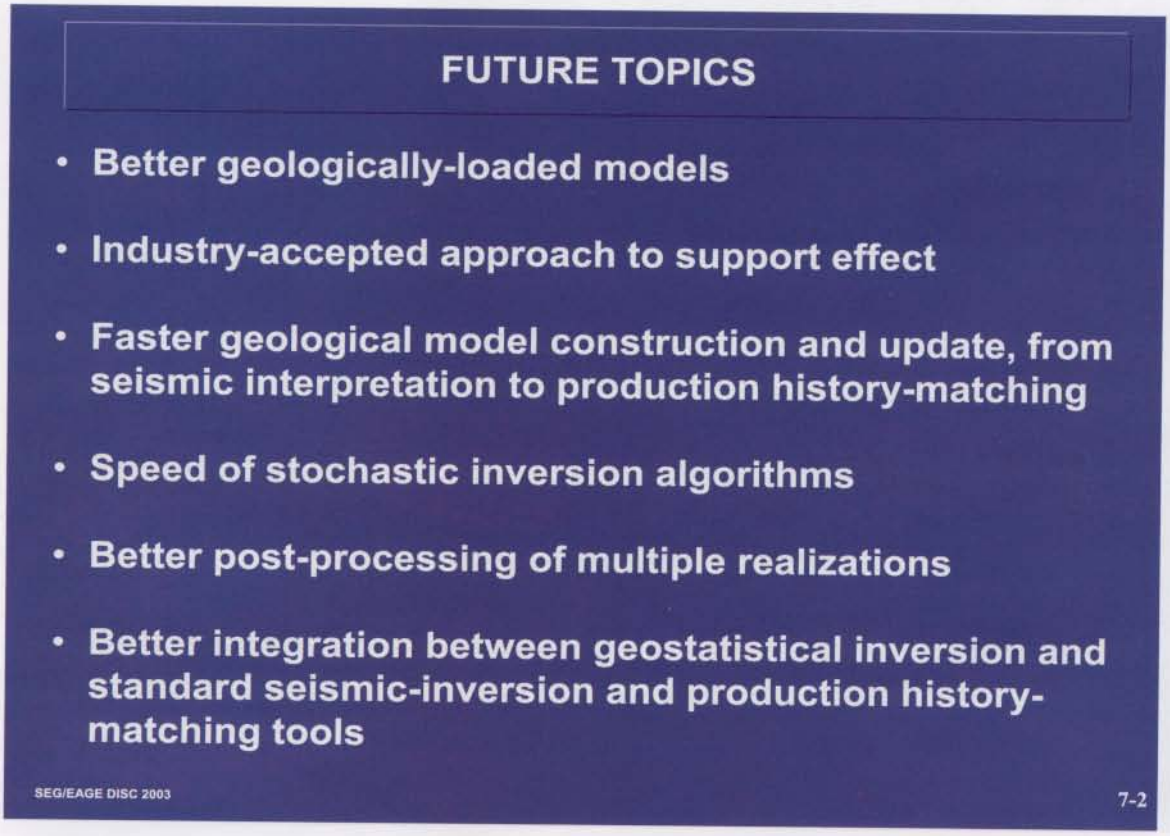
The development of earth modeling software has dramatically impacted the work of the sedimentologist and the reservoir geologist. Fewer models are generated by hand,

and geologists want better geologically loaded models that are able to incorporate their a priori knowledge about reservoir architecture (Fig. 7-2). This is one of the drivers behind the research in multiple-point geostatistics.

The border between geostatistics and seismic inversion, or between geostatistical uncertainty quantification and production history matching, is also disappearing. Reservoir engineers now see that the multirealization approach may help them tackle this old problem of the nonuniqueness of history matching. The success met by geostatistical inversion is generating new questions among geophysicists. How can we make better use of the multiple realizations? Can we apply this stochastic paradigm to prestack data? Academics, contractors, and petroleum companies work together on these topics, which may require strong knowledge of mathematics and an understanding of the business issues, and which may lead — when the methods are successful — to their integration into earth modeling software.

The role of the petroleum geostatistician is also changing. More and more, he/she must build a dialog with the various disciplines and understand the tools of the trade in order to understand the added value that geostatistics can bring. The time when the geostatistician worked sequentially with the geophysicist, the geologist, and the reservoir engineer is over. The geostatistician needs to spend less time developing new geostatistical methods, but more time understanding how the existing ones fit into the multidisciplinary integration workflow.

Thus, geostatistics offers a range of tools for building 3D models that are consistent with all data available, and for quantifying the associated uncertainty. The different



disciplines meet each other around the geostatistical model, which acts like a glue between them, thanks to the algorithms and models it provides. In spite of the recent progress made by earth modeling software, it still takes too much time to go from seismic interpretation to flow simulation, or to update models as new data are acquired. Today, new data often require complete reconstruction of the model, rather than a fast update. There is room for much progress in this area.

The multidisciplinary reservoir characterization process often remains long and tedious. Issues such as the support effect, the speed of geostatistical inversion algorithms, or the processing of multiple realizations are yet to be addressed properly and with industry-accepted solutions.

7.3 Websites and Software

Fig. 7-3 provides a list of active geostatistics websites, all of which are run by academic institutions that are leaders in petroleum geostatistics research: Stanford University, Ecole des Mines de Paris, University of Alberta, Norwegian Computing Centre, and the University of Trondheim. This is definitely not an exhaustive list, but it is a good entry point.

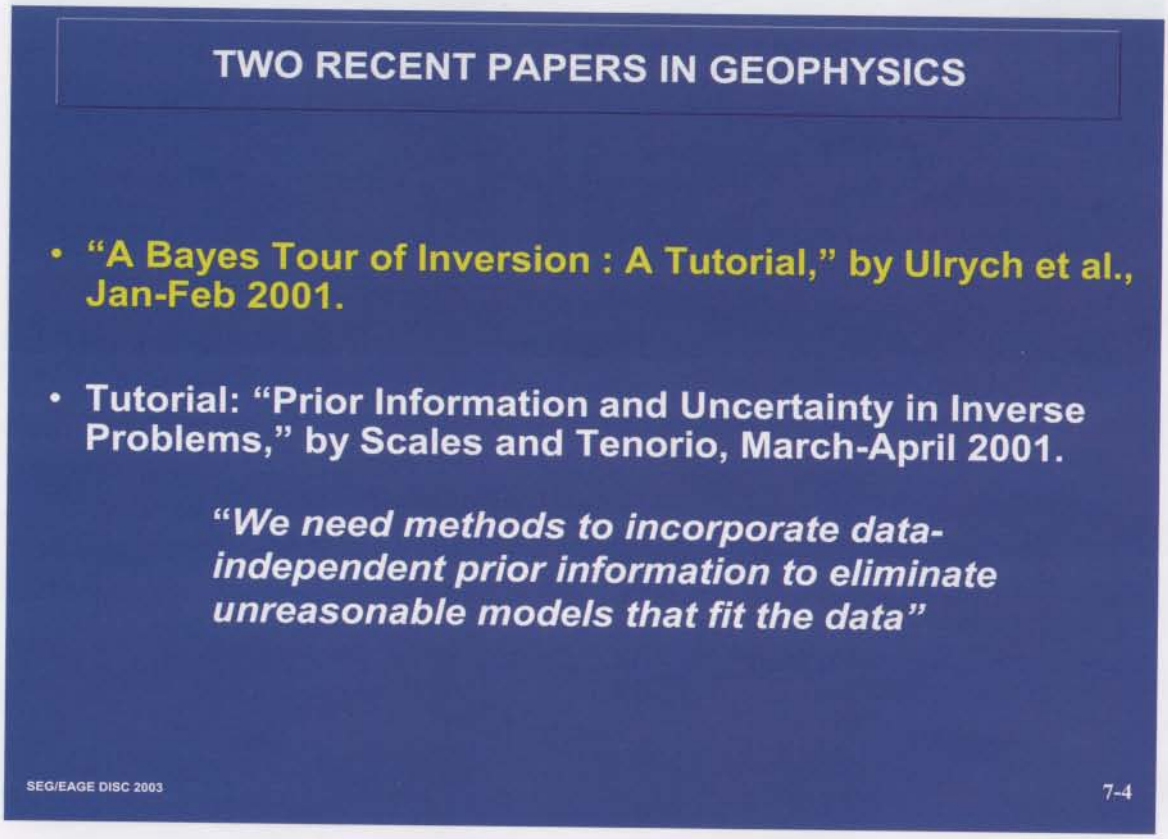
We do not specifically discuss software here, to avoid commercialism. Today, new geostatistical techniques are usually developed by companies as in-house prototypes, or by academics using public-domain software libraries (the most successful example today is GSLIB; see Deutsch and Journel, 1992). There are very few stand-alone com-

mercial geostatistical products. Once new geostatistical technologies have been proved, whether by academics, contractors, or petroleum companies, they are usually incorporated into earth modeling software, because such software (see Chapter 1) provides the geometrical and stratigraphic grid framework required for modeling studies.

7.4 The Role of Geostatistics in Geophysics

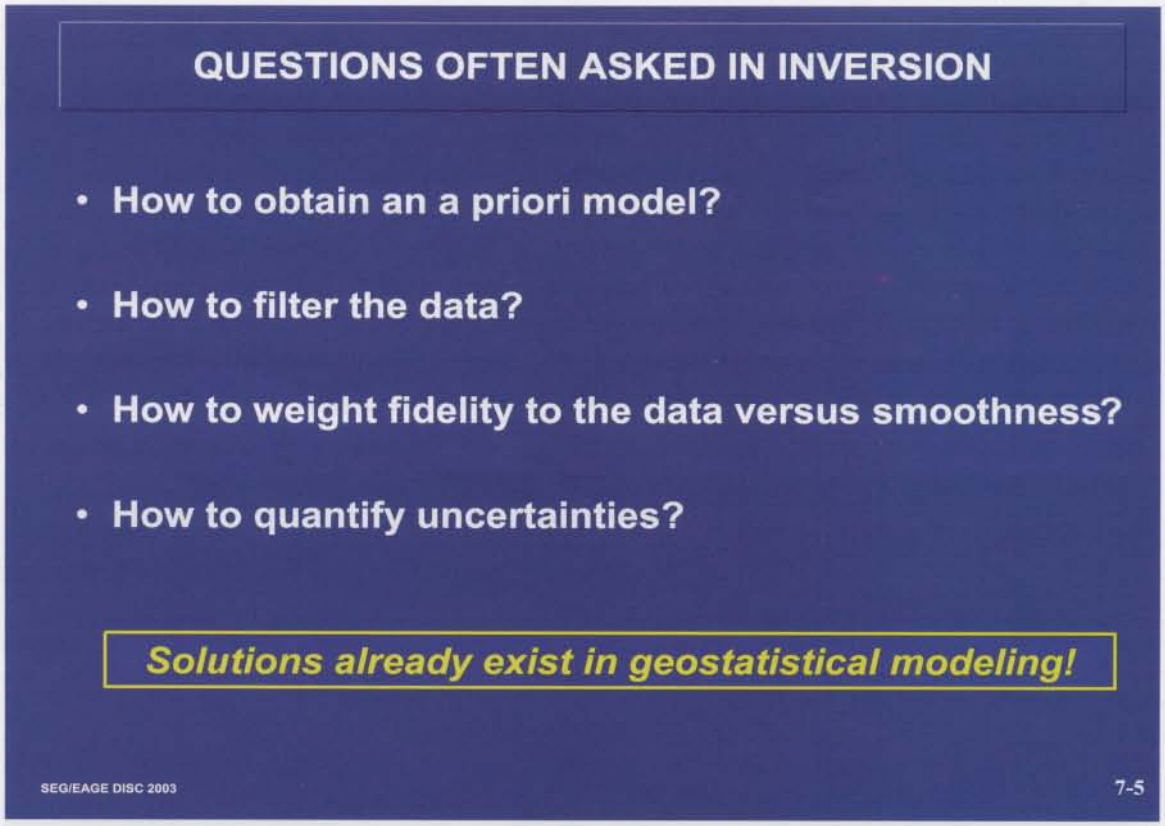
It seems appropriate to conclude this course with considerations about the role of geostatistics in geophysics. Hopefully, we have now made clear that the a priori geostatistical model, as quantified by the variogram and the trend model, is just an extension of the a priori model used in Bayesian inversion. We have also seen in this course that, if an energy-based inversion approach is used, the specification of a regularization term is equivalent to that of an a priori model. Of course, the a priori model of the geostatistician should not be different from that of the geophysicist! Thus, geostatistics formalizes and quantifies the a priori geological model and makes a significant step toward filling the need stated by Scales and Tenorio (2001) (Fig. 7-4).

Geostatistics also provides a large number of tools that extend the standard 3D modeling tool-box of geophysicists. Deterministic techniques such as kriging, error cokriging, factorial kriging, collocated cokriging, and external drift provide new ways of filtering acquisition artifacts or combining different kinds of information. These tools can be applied to scattered data without going through the intermediate step of interpolation on a regular grid, as spectral methods do.



Even more interesting, geostatistical simulation provides a means of generating samples of the a posteriori distribution, whether it is constrained by wells, seismic, or dynamic data (Fig. 7-5). The recent developments of geostatistical inversion, even if algorithms need to be optimized, show that these techniques have the potential to change the way seismic-constrained modeling was done in the past.

Thus, we hope that course attendees and readers of these notes will go home with an understanding that the gap between deterministic and probabilistic techniques has now almost disappeared, as Wadsworth et al. (Fig. 7-6) had already envisoned in . . . 1953!



THE LAST WORD . . .

THE DETERMINISTIC AND THE PROBABILISTIC APPROACH

"There are two basic approaches to treating data observed in nature, and in particular the data represented on a seismogram. One is the deterministic approach and the other is the probabilistic approach. Many people think of these two approaches as conflicting, but actually this is not the case. Recent investigations indicate that each approach is fundamentally equivalent to the other."

Wadsworth, Robinson, Bryan and Hurley, 1953.

8 Exercises

The exercises are modified from some of the exercises already presented in Dubrule (1998).

8.1 Exercise 1: Fitting a Variogram Model

The goal of this exercise is to show how a 3D model is fitted to an anisotropic experimental variogram.

The example used is from Chu et al. (1994). This paper shows (in its Fig. 7) experimental porosity variograms calculated in the North-South, East-West, and vertical directions. The data are from an Amoco west-Texas carbonate field of Permian age.

The experimental variograms were calculated from the well data in their stratigraphic coordinates, within a layer of thickness ranging from 11 to 27 m. A total of 4697 elementary porosity-log data were available in 90 wells. Experimental variograms are displayed in Fig. 2-26.

The model fitted by Chu et al. is shown in Fig. 2-26. A screen copy of the spreadsheet used is shown in Fig. E1-1. Thanks to this spreadsheet, it is possible to evaluate the impact of a parameter change on the variogram model, and thus to better understand the meaning of each parameter. A particular point to discuss is the use of a short-range model to represent the nugget effect of the lateral variograms, without impacting the vertical variogram fit.

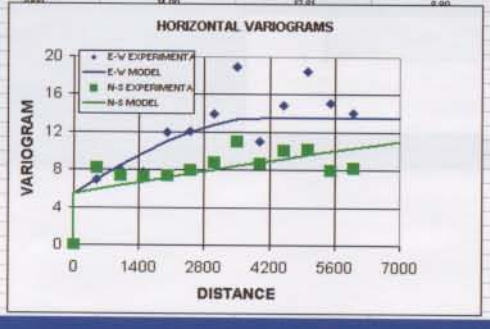
EXERCISE 1 (CHU ET AL, 1994)

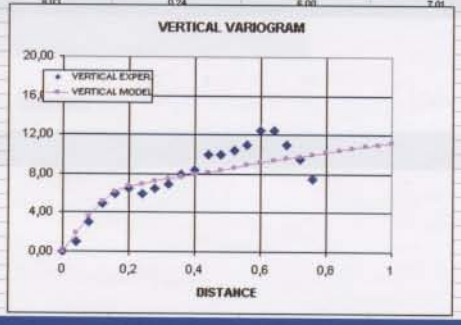
VARIOGRAM FITTING EXAMPLE

EXPERIMENTAL VARIOGRAM (EXAMPLE TAKEN FROM CHU ET AL, AAPG94-16)

		SILL	RANGE E-W	RANGE N-S	RANGE VERTICAL
NUGGET	0,000				
MODEL1		5,500	0,2	0,2	
MODEL 2		8	4000	14000	1,9

HORIZONTAL DISTANC	E-W EXPERIMENTAL	E-W MODEL	N-S EXPERIMENTAL	N-S MODEL	VERTICAL DISTANCE	VERTICAL EXPER.	VERTICAL MODEL
0	0	0,00	0,00	0,00	0	0,00	0,00
0,2		5,50		5,50	0,001		0,05
500	7,00	6,99	8,20	5,93	0,04	1,00	1,88
1000	8,00	8,44	7,50	6,36	0,09	3,00	3,63
1500	7,80	9,79	7,30	6,78	0,12	5,00	5,11
2000	12,00	11,00	7,50	7,20	0,16	6,00	6,20
2500	12,08	12,02	9,00	7,62	0,2	6,50	6,76





SEG/EAGE DISC 2003

E1-1

8.2 Exercise 2: Understanding the Kriging System

The goal of this exercise is to evaluate the impact of the variogram choice on the results of an elementary kriging system.

We assume that we work in 2D, and that the value of a variable z at a data point $\mathbf{x}_0$ is kriged using the values at four data points located in the neighborhood of $\mathbf{x}_0$ (Fig. E2-1). We make assumptions about the variogram model (which can be Gaussian, spherical, or exponential), about the practical range and the sill, and about the ratio and direction of anisotropy. The size of the color circles is proportional to the value of the four corresponding kriging weights. If a kriging weight is negative, no circle is plotted.

By moving the data points around the estimated point and changing the parameters of the variogram model, the impact of each parameter of the variogram on the kriging system is better understood.

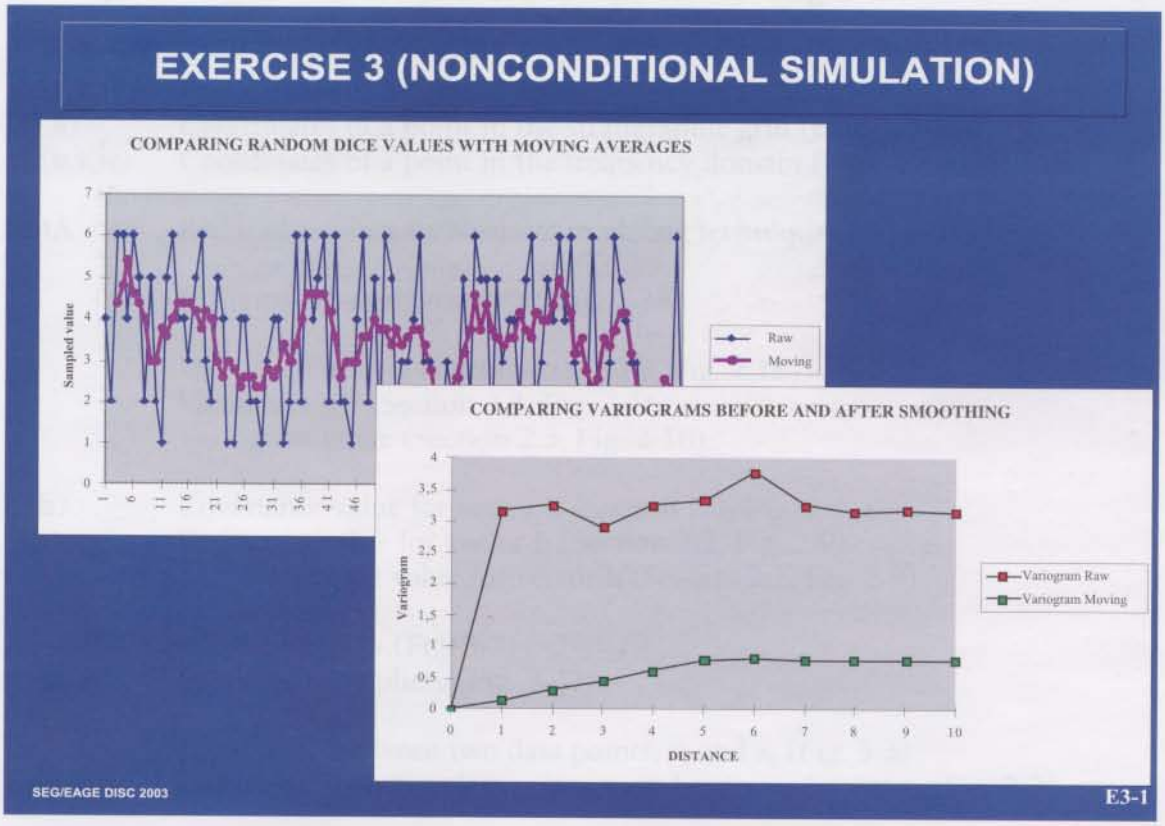
8.3 Exercise 3: Generating a Nonconditional Simulation in 1D

The goal of this exercise is to explain, using a simple example, how a nonconditional simulation can be generated.

A throw of a die generates a uniformly distributed, random variable taking the values 1, 2, 3, 4, 5, or 6. If the die is thrown repeatedly, a sequence of uncorrelated values is generated. In geostatistical terms, this is called an uncorrelated random function, or "white noise," the variogram of which is a "pure nugget effect."

The exercise consists of experimenting with a simple approach for introducing spatial correlation into such a sequence. The value of z at each location is averaged with the two preceding and the two following values. A spreadsheet simulates 100 throws of a die and calculates this moving average at each location.

Fig. E3-1 shows an example of a sequence of values z (Raw) and the corresponding average (moving) at 100 locations. How do the sequences differ? Fig. E3-1 also shows the variograms calculated on the data displayed in Fig. E3-1. What has been the impact of averaging the initial results of the die's throw? How does this relate to the previously discussed moving simulation technique of Oliver (1995)?



9 Notation

$\mathbf{x}$	A point in 1D, 2D, or 3D space
$z(\mathbf{x})$	A property measured in 1D, 2D, or 3D space
$Z(\mathbf{x})$	A random function in 1D, 2D, or 3D space (Section 2.1)
z_i	The value of variable z at location $\mathbf{x}_i$
$Z_1(\mathbf{x})$	The primary variable (Section 2.8)
$Z_2(\mathbf{x})$	The secondary variable (Section 2.8)
x	A real number
$f(x)$	A probability density function (pdf) (Section 1.4.1, Fig. 1-13)
$F(x)$	A cumulative density function (cdf) (Section 1.4.1, Fig. 1-13)
X,Y, or Z	A random variable (Fig. 1-12)
m	The mean of a random variable (Fig. 1-15)
σ	The standard deviation of a random variable (Fig. 1-15)
q_x	The x% quantile of a random variable (Fig. 1-24)
$\mathbf{x}=(x,y,z)$	Coordinates of a point in space
(i,j,k)	Coordinates of a point in the stratigraphic grid (Figs. 1-5 and 1-9)
$\mathbf{u}=(u,v,w)$	Coordinates of a point in the frequency domain (Figs. 2-40 and 2-41)
RMA	Reduced-major-axis bivariate modeling technique (Fig. 1-37)
Σ	Variance-covariance matrix (Fig. 1-38)
C_0	Variogram nugget effect (Section 2.3, Fig. 2-16)
C	Variogram sill (Section 2.3, Fig. 2-16)
a	Variogram range (Section 2.3, Fig. 2-16)
$C(\mathbf{h})$	Covariance value for vector $\mathbf{h}$ (Section 2.2, Fig. 2-9)
$\gamma(\mathbf{h})$	Variogram value for vector $\mathbf{h}$ (Section 2.2, Fig. 2-9)
$\rho(\mathbf{h})$	Autocorrelation value for vector $\mathbf{h}$ (Section 2.2, Fig. 2-9)
$\lambda_1,\ldots,\lambda_N$	Kriging weights (Fig. 3-3)
μ, μ_x, μ_y	Lagrange multipliers (Fig. 3-3)
C_{ij}	Covariance between two data points, $\mathbf{x}_i$ and $\mathbf{x}_j$ (Fig. 3-3)
C_{i0}	Covariance between data points $\mathbf{x}_i$ and estimated point $\mathbf{x}_0$ (Fig. 3-3)
$\varepsilon\ (x,y)$	Measurement error at location (x,y), for instance, with V_{stack} data (Fig. 3-35)
e_i	Measurement error affecting measurement z_i at location, $\mathbf{x}_i$ (Fig. 3-91)
V_{stack}	Stacking velocity (Section 3.3.1)

V_{app}	Apparent velocity (Fig. 3-14)
UK	Universal kriging (Section 3.2.2, Fig. 3-3)
$Z_{uk}(\mathbf{x})$	Universal kriging interpolation at location **x** (Section 3.2.2, Fig. 3-3)
SK	Simple kriging (Section 3.2.2)
OK	Ordinary kriging (Section 3.2.2)
$\sigma_k^2(\mathbf{x})$	Kriging variance at location **x** (Section 3.2.2, Fig. 3-17)
FK	Factorial kriging (Section 3.3.2, Fig. 3-49)
$Z_{fk}(\mathbf{x})$	Factorial kriging interpolation at location **x** (Section 3.3.2, Fig. 3-49)
KED	Kriging with an external drift (Section 3.4, Fig. 3-56)
$Z_{ked}(\mathbf{x})$	Kriging with an external drift interpolation at location **x**
BK	Bayesian kriging (Section 3.5, Fig. 3-65)
COK	Cokriging (Section 3.6, Fig. 3-67)
$Z_{cok}(\mathbf{x})$	Cokriging interpolation at location **x** (Fig. 3-67)
$\sigma_{cok}^2(\mathbf{x})$	Cokriging variance at location **x** (Fig. 3-67)
CCK	Collocated cokriging (Section 3.6.3)
$Z_{cck}(\mathbf{x})$	Collocated cokriging interpolation at location **x** (Fig. 3-75)
MCCK	Multicollocated cokriging (Section 3.6.3)
k-IRF	Intrinsic random functions of order k (Section 3.2.3)
GC-k	Generalized covariance of order k (Section 3.2.3)
$z_{hs}(\mathbf{x})$	Harmonic spline interpolation at location **x** (Section 3.8.2, Fig. 3-90)
$z_{bhs}(\mathbf{x})$	Biharmonic spline interpolation at location **x** (Section 3.8.2, Fig. 3-90)
$z_{ss}(\mathbf{x})$	Smoothing spline interpolation at location **x** (Section 3.8.2, Fig. 3-91)
$z_{rbf}(\mathbf{x})$	Radial basis function interpolation at location **x** (Section 3.8.3, Fig. 3-91)
θ	Weighting parameter used by smoothing splines (Section 3.8.2, Fig. 3-91)
GCS	Geostatistical conditional simulation (Chapter 4, Fig. 4-1)
MCS	Monte-Carlo simulation (Section 4.2, Figs. 4-3 and 4-4)
OIP	Oil-in-place volume
STOOIP	Standard original oil-in-place volume
GRV	Gross-rock volume
Φ	Porosity
N/G	Net/gross
S_w	Water-saturation
B_o	Formation volume factor
SGS	Sequential Gaussian simulation (Section 4.3.2, Fig. 4-22)
SIS	Sequential indicator simulation (Section 4.5.2, Fig. 4-42)
MCMC	Markov-chain Monte Carlo (Section 4.3.2, Figs. 4-24 and 4-25)
GI	Geostatistical inversion (Chapter 5, Fig. 5-3)

10 Acknowledgments

Before all, I would like to thank TotalFinaElf for enthusiastically supporting me in both the preparation and the actual delivery of this 2003 DISC course manual and presentation. Thank you also to affiliate companies, who provided some of their data.

A very special thank you to *Pierre Delfiner, Jean-Luc Piazza,* and *Vincent Bigault de Cazanove*, who helped me a lot in the preparation of the manuscript.

Within the company, many other people were kind enough to provide me with some technical material. They are acknowledged by having their name written down at the bottom of each figure they contributed. These people are: *Leon Barens (from the TotalFinaElf Exploration UK Geoscience Research Centre in London), Wafik Beydoun, Pierre Biver, Bernard Corre, Pierre Delfiner, Vivien de Feraudy, Xavier Freulon, Jean-Michel Guéméné, André Haas, Gérard Massonnat, Jean-Luc Piazza, Etienne Robein, Philippe Samson, Pierre Thore.*

The following people from TotalFinaElf also helped me in the preparation of the manuscript, by producing slides, making bibliographic search, or reviewing the English: *Hélène Boyer, Cathy Brewerton, Nadine Calderoni, Geneviève Desbordes, Jean-Louis Poulot-Cadet, Aurélien Tendant.*

I would also like to thank many people outside TotalFinaElf, who kindly contributed some material:
Hisham Al Qassab from Saudi Aramco.
Philippe Doyen, formerly with Western Geophysical, now with Schlumberger.
Peter Frykman, from the Geological Survey of Denmark and Greenland.
Henning Omre, from the Norwegian University of Science and Technology.
Peter Swaby, formerly with the TotalFinaElf Geoscience Research Centre.
Philippe Lamy, from Earth Decision Sciences.
Laurent de Chambure, Sarah Boudon, Stéphanie Légeron, Fleur Bouramoué , from ERM.S.
Thierry Coleou and David Le Meur, from CGG.
Paul van Riel, from Jason, and Carlos Torres-Verdin, from The University of Texas at Austin.

Thank you to *SEG* and *EAGE* for choosing me to give this course and supporting me all along. Special thanks to *SEG'S Dean Clark, Marcia Burnett,* and *Ted Bakamjian; Anne Thomas* and *ProType,* for helping me in the preparation of the manuscript; and to *Lisa Sutliff* from *SEG*, and *Eveline Schut* from *EAGE* for helping me in the organization of the course.

Last but not least, I would like to thank my wife *Anne*, and my children *Pauline, Noémie, Arthur,* and *Thibaud* for their love and support during these sunny weekends in Pau, or the last summer vacations in beautiful "Ile de Ré," which I spent preparing the DISC lecture, stuck in front of my computer.

11 References

Abrahamsen, P., Hauge, R., Heggland, K., and Mostad, P., 2000, Estimation of gross–rock volume of filled geological structures with uncertainty measures: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE Reservoir Evaluation and Engineering, 3(4), Aug., 304–309.

Abrahamsen, P., Omre, H., and Lia, O., 1991, Stochastic models for seismic depth conversion of geological horizons: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 23138.

Agterberg, F. K, 1994, Fractals, multifractals, and change of support, *in* Dimitrakopoulos, R., Ed., Geostatistics for the next century: Kluwer Academic Publ., 223–234.

Ahmed, S., and Murthy, P.S.N., 1997, Could radial basis function estimator replace ordinary kriging, *in* Baafi, E. Y., and Schofield, N. A., Eds., Geostatistics Wollongong 96, Kluwer Academic Publ., 314–323.

Aki, K., and Richards, P. G., 1980, Quantitative seismology: W.H. Freeman and Co.

Alabert, F. G., 1987, The practice of fast conditional simulations through the LU decomposition of the covariance matrix: Mathematical Geology, **19** (5), 369–387.

Alabert, F.G., and Massonnat, G.J., 1990, Heterogeneity in a complex turbiditic reservoir: Stochastic modeling of facies and petrophysical variability: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 20604.

Almeida, A. S., and Frykman, P., 1994, Geostatistical modeling of chalk reservoir properties in the Dan Field, Danish North Sea, *in* Yarus, J. M., and Chambers, R. L. Eds., Stochastic modeling and geostatistics: Am. Assn. Petr. Geol., AAPG Computer Applications in Geology, No. 3, 273–286.

Al Qassab, H., Fitzmaurice, J., Al Ali, Z., Al Khalifa, M., Aktas, G., and Glover, P., 2000, Cross–discipline integration in reservoir modeling: the impact on fluid flow simulation and reservoir management: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 62902.

Arakawa, K. and Krotkov, E., 1996, Fractal modeling of natural terrain: analysis and surface reconstruction with range data: Graphical Models and Image Processing, **58**(5), 413–436.

Bashore, W. M., Araktingi, U. G., Levy, M., and Schweller, W. J., 1993, The importance of the geological model for reservoir characterization using geostatistical techniques and the impact on subsequent fluid flow: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 26474.

Behrens, R. A., MacLeod, M. K., and Tran, T. T., 1998, incorporating seismic attribute maps in 3D reservoir models: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE reservoir Evaluation and Engineering, April, 122–126.

Besag, J., 1974, Spatial interaction and the statistical analysis of lattice systems: Journal of the Royal Statistical Society, Series B, **36**(2), p. 192–225.

Billings, S. D., Beaston, R. K., and Newsam, G. N., 2002a, Interpolation of geophysical data using continuous global surfaces: Geophysics, **67**(6), 1810–1822.

Billings, S.D., Newsam, G.N., and Beaston, R.K., 2002b, Smooth fitting of geophysical data using continuous global surfaces: Geophysics, **67**(6), 1823–1834.

Bolondi, G., Rocca, F., and Zanoletti, S., 1976, Automatic contouring of faulted subsurfaces: Geophysics, **41**(6), 1377–1393.

Bortoli, L. J., Alabert, F., Haas, A., and Journel, A. G., 1992, Constraining stochastic images to seismic data, *in* Soares A., Ed., Proceedings of the International Geostatistics Congress, Troia 1992: Kluwer Academic Publ.

Briggs, I. C., 1974, Machine contouring using minimum curvature: Geophysics, **39**(1) 39–48.

Bryant, I. D., and Flint, S. S., 1993, Quantitative clastic reservoir geological modeling: problems and perspectives: Internat. Assn. Sediment. Spec. Publs., **15**, 3–20.

Buland, A., Kolbjornsen, O., Omre, H., 2003, Rapid Spatially Coupled AVO Inversion in the Fourier Domain: Geophysics, **68**, in process.

Buland, A., and Omre, H., 2003, Bayesian linearized AVO inversion: Geophysics, **68**(1), 185–198.

Capen, E. C., 1976, The difficulty of quantifying uncertainty: Journal of Petroleum Technology, **18**(8), 843–850.

Carle, S. F., and Fogg, G. E., 1996, Transition probability–based geostatistics: Mathematical Geology, **28**(4), 453–476.

Charles, Th., Guéméné, J. M., Corre, B., Vincent, G., and Dubrule, O., 2001, Experience with the quantification of subsurface uncertainties: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 68703.

Chewaroungroaj, J., Varela, O. J., Lake, L. W., 2000, An evaluation of procedures to estimate uncertainty in hydrocarbon recovery predictions: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 59449.

Chilès, J. P., and Delfiner, P., 1999, Geostatistics modeling spatial uncertainty: Wiley Series in Probability and Statistics, Wiley & Sons.

Chilès, J. P., and Guillen, A., 1984, Variogrammes et Krigeages pour la Gravimétrie et le Magnétisme: Sciences de la Terre, No. **20**.

Christie, M. A., 1996, Upscaling for reservoir simulation: Journal of Petroleum Technology, **48**(11), 1004–1010.

Chu, J., Xu, W., and Journel, A. G., 1994, 3–D implementation of geostatistical analyses —the Amoco case study, *in* Yarus, J. M., and Chambers, R. L., Eds., Stochastic modeling and geostatistics: Am. Assn. Petr. Geol., AAPG Computer Applications in Geology, No. 3, 201–216.

Claerbout, J., and Brown, M., 1999, Two-dimensional textures and prediction-error filters: Presented at the EAGE 61st Conference and Technical Exhibition, Eur. Assn. Geosci. Eng., Extended Abstracts.

Claerbout, J., 2002, Image estimation by example: geophysical soundings image construction: multidimensional autoregression: <http://sepwww.stanford.edu/sep/prof> Accessed January 21, 2003.

Clemetsen, R., Hurst, A. R., Knarud, R., and Omre, H., 1990, A computer program for evaluation of fluvial reservoirs, North Sea oil and gas reservoirs II: Graham and Trotman, 373–385.

Coleou, Th., 1996, Tying well data to seismic data: the missing link between statistics and geostatistics: Presented at the Norsk Petroleumforening (NPF) Conference on Geophysics for Lithology Prediction.

Coleou, Th., 2001, On the use of seismic velocities in model building for depth conversion:

Presented at the EAGE 63rd Conference and Technical Exhibition, Eur. Assn. Geosci. Eng., Extended Abstracts, IV-1.

Coleou, Th., 2002a, Links between external drift, Bayesian kriging, collocated cokriging: <www.cgg.com/proserv/software/hints/KR1.html> Accessed January 21, 2003.

Coleou, Th., 2002b, Time-lapse filtering and improved repeatability with automatic factorial co-kriging: Presented at the EAGE 64th Conference and Technical Exhibition, Eur. Assn. Geosci. Eng.

Corre, B., Thore, P., de Feraudy, V., and Vincent, G., 2000, Integrated uncertainty assessment for project evaluation and risk analysis: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., Soc. Petr. Eng., SPE 65205.

Crane, S.D. and Tubman, K.M., 1990, Reservoir variability and modeling with fractals: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 20606.

Damsleth, E., Tjolsen, C.B., Omre, H., and Haldorsen, H. H., 1990, A two-stage stochastic model applied to a North Sea reservoir: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 20605.

Damsleth, E., Hage, A., and Volden, R., 1992, Maximum information at minimum cost, a North Sea field development study with experimental design: Journal of Petroleum Technology, **44**(12), 1350–1356.

Damsleth, E., and Omre, H., 1997, Geostatistical approaches in reservoir evaluation: Journal of Petroleum Technology, **49**(5), 498–501.

Davis, M. W., 1987, Production of conditional simulations via the LU triangular decomposition of the covariance matrix: Mathematical Geology, **19**(2), 91–98.

Deutsch, C. V., and Journel, A. G., 1992, GSLIB, geostatistical software library and user's guide: Oxford University Press.

Deutsch, C. V., 1992, Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data: PhD Thesis, Stanford University.

Deutsch, C. V., Srinivasan, S., and Mo, Y., 1996, Geostatistical reservoir modeling accounting for precision and scale of seismic data: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., ,SPE 36497.

Deutsch, C. V., 2002, Geostatistical reservoir modeling: Oxford University Press.

Doyen, P.M., 1988, Porosity from seismic data: a geostatistical approach: Geophysics, **53**(10), 1263–1275.

Doyen, P. M., Psaila, D. E., and Strandenes, S., 1994, Bayesian sequential indicator simulation of channel sands from 3D seismic data in the Oseberg field, Norwegian North Sea: Soc. Petr. Eng., SPE 28382.

Doyen, P. M., den Boer, L.D., and Pillet, W. R., 1996, Seismic porosity mapping in the Ekofisk field using a new form of collocated cokriging: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 36498.

Doyen, P. M., Psaila, D. E., den Boer, L. D., 1997, Reconciling data at seismic and well log scales in 3-D earth modeling: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE38698.

Dubrule, O., 1983, Cross-validation of kriging in a unique neighborhood: Mathematical Geology, **15**(6), 687–699.

Dubrule, O., and Haldorsen, H. H., 1984, Geostatistics for permeability estimation, *in* Reservoir characterization: Academic Press, p. 223–247.

Dubrule, O., 1994, Estimating or choosing a geostatistical model? *in* Dimitrakopoulos, R., Ed., Geostatistics for the next century: Kluwer Academic Publ., 3–14.

Dubrule, O., Dromgoole, P., and van Kruijsdijk, C., 1996, Workshop report: "Uncertainty in reserve estimates," EAGE Conference: Petroleum Geoscience, 2(4): 1996.

Dubrule, O., Basire, C., Bombarde, S., Samson, Ph., Segonds, D., and Wonham, J., 1997, reservoir geology using 3-D modeling tools: Soc. Petr. Eng., SPE 38659.

Dubrule, O., 1998, Geostatistics in petroleum geology: Am. Assn. Petr. Geol., AAPG Continuing Education Course Notes Series, No. 38.

Dubrule, O., Thibaut, M., Lamy, Ph., and Haas, A., 1998, Geostatistical reservoir characterization constrained by 3D seismic data: Petroleum Geoscience, 4, 121–128.

Dubrule, O. and Damsleth, E., 2001, Achievements and challenges in petroleum geostatistics: Petroleum Geoscience, 7, 1–7.

Duchon, J., 1975, Fonctions splines du type plaque mince en dimension 2: Séminaire d'analyse numérique, No. 231, U.S.M.G.

Duijndam, A .J. W., 1988, Bayesian estimation in seismic inversion, part ii: uncertainty analysis: Geophysical Prospecting, 36, 899–918.

Eide, A. L., Omre, H., and Ursin, B, 1997a, Stochastic reservoir characterization conditioned on seismic data, *in* Baafi, E. Y., and Schofield, N. A., Eds., Geostatistics Wollongong 96: Kluwer Academic Publ., 442–453.

Eide, A. L., Ursin, B., and Omre, H., 1997b, Stochastic simulation of porosity and acoustic impedance conditioned to seismic data and well data: 67th Ann. Mtg., Soc. Expl. Geophs., Expanded Abstracts, 1614–1617.

Farmer, C. L. 1988, The generation of stochastic fields of reservoir parameters with specified geostatistical distributions: Mathematics of oil production: Oxford Science Publications, Clarendon Press, 235–252.

Farmer, C. L, 1992, Numerical rocks, *in* King, M., Ed., Mathematics of oil recovery: Oxford University Press.

Feynmann, R. P., Leighton, R. B., and Sands, M., 1964, The Feynmann lectures on physics, Volume II: Addison Wesley.

Franke, R., and Nielson, G. M., 1991, Scattered data interpolation and applications: a tutorial and survey, *in* Hagen, H., and Roller, D., Eds, Geometric modeling methods and applications: Computer Graphics Series, Springer–Verlag, 131–159.

Frykman, P., and Deutsch, C. V., 2002, Practical application of geostatistical scaling laws for data integration: Petrophysics, 43(3), 153–171.

Gamerman, D., 1997, Markov chain Monte Carlo: Chapman and Hall.

Gastaldi, C., Roy, D., Doyen, Ph., and Den Boer, L., 1998, Using Bayesian simulations to predict reservoir thickness under tuning conditions: The Leading Edge, April, 589–593.

Geman, D., and Geman, S., 1984, Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images: IEEE transactions on pattern analysis and machine intelligence: PAMI–6(6), 721–741.

Gomez–Hernandez, J. J., and Journel, A. G., 1993, Joint sequential simulation of multigaussian fields, *in* Geostatistics Troia '92, Kluwer Academic Publ., 85–94.

Grijalba–Cuenca, A., Torres–Verdin, C., and van der Made, P., 2000, Geostatistical inversion of 3D seismic data to extrapolate wireline petrophysical variables away from the well: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 63283.

Guéméné, J. M., Thore, P., and Meesemaecker, R., 2002, Structural uncertainties and their impact on a Gulf of Guinea field and satellite prospects: Presented at the EAGE 64th Conference and Exhibition, Eur. Assn. Geol. Eng.

Gundeso, R. O., and Egeland, O., 1990, Sesimira—A new geological tool for 3D modelling of heterogeneous reservoirs, *in* North Sea oil and gas reservoirs, II: Graham and Trotman, 363–371.

Gunning, J. and Paterson, L., 1999, Conditoning of Levy-stable fractal reservoir models to seismic data: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 56823.

Haas, A., and Viallix, J. R., 1974, Krigeage applied to geophysics, the answer to the problem of estimates and contouring: Geophysical Prospecting, **24**, 49–69.

Haas, A., and Jousselin, C., 1976, Geostatistics in the petroleum industry, *in* Advanced geostatistics in the mining industry: D. Reidel Publ., 333–347.

Haas, A., and Dubrule, O., 1994, Geostatistical inversion—a sequential method of stochastic reservoir modeling constrained by seismic data: First Break, **12** (11), 561–569.

Haas, A., Biver, P., and Mouliere, D, 1998, Simulations Stochastiques en Cascade, Cahiers de Géostatistique: Ecole des Mines de Paris, **6**, 31–43.

Haldorsen, H. H., and Damsleth, E., 1990, Stochastic modeling: Journal of Petroleum Technology, **42**(4), 404–412.

Hansen, R. O., 1993, Interpretive gridding by anisotropic kriging: Geophysics, **58**(10), 1491–1497.

Hardy, R.L., 1990, Theory and applications of the multiquadric–biharmonic method, 20 years of discovery, 1968–1988: Computers Math. Applic., **19**(8/9), 163–208.

Hegstad, B. K., Omre, H., Tjelemeland, H., and Tyler, K., 1994, Stochastic simulation and conditioning by annealing in reservoir description, *in* Amstrong, M., and Dowd, P. A., Eds, Geostatistical simulations: Kluwer Academic Publ., 43–55.

Hegstad, B. K., and Omre, H., 2001, Uncertainty in production forecasts based on well observations, seismic data, and production history: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE Journal, Dec., 409–424.

Hewett, T. A., 1986, Fractal distributions of reservoir heterogeneity and their influence on fluid transport: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 15386.

Hirsche, K., Boerner, S., Kalkomey, C., Gastaldi, C., 1998, Avoiding pitfalls in geostatistical reservoir characterization: a survival guide: The Leading Edge, April, 493–504.

Hohn, M. E., 1988, Geostatistics and petroleum geology, computer methods in the geosciences: van Nostrand Reinhold.

Holden, L., Hauge, R., Skare, O., and Skorstad, A., 1998, Modeling of fluvial reservoirs with object models : Mathematical Geology, **30**(5), —.

Insalaco, E., Boisseau, Y., Marion, D., Michel, B., Rowbotham, P., 2001, Reservoir–scale 3D sedimentary modeling: approaches and impact of integrating sedimentology into the

reservoir characterization workflow: Presented at the Am. Assn. Petr. Geol. International Meeting.

Isaaks, E. H., and Srivastava, R. M., 1989, Applied geostatistics: Oxford University Press.

Isenberg, C., 1992, The science of soap films and soap bubbles: Dover.

Jeffery, R. W., Stewart, I. C, and Alexander, D. W., 1996, Geostatistical estimation of depth conversion velocity using well control and gravity data: First Break **14**(8), 313–320.

Johnson, H. D., and D. E. Krol, 1984, Geological modeling of a heterogeneous sandstone reservoir: Lower Jurassic Statfjord Formation, Brent field: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 13050.

Johnson, M. E., 1987, Multivariate statistical simulation: Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc.

Journel, A. G., and Huijbregts, Ch. J., 1978, Mining geostatistics: Academic Press, Inc.

Journel, A. G., and Gomez Hernandez, J. J., 1989, Stochastic imaging of the Wilmington Clastic Sequence: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 19857.

Journel, A. G., 1993, modeling uncertainty: some conceptual thoughts, *in* Dimitrakopoulos, R., Ed., Geostatistics for the next century: Kluwer Academic Publ.

Journel, A. G., 1999, Markov models for cross-covariances: Mathematical Geology, **33**(2), 117–131.

Kane, J., Rodi, W., and Toksoz, N., 2001, Simultaneous least squares deconvolution and kriging using conjugate gradients: EAGE 63rd Conference and Technical Exhibition, Extended Abstracts, A-024.

Kelkar, M., 2000, Application of geostatistics for reservoir characterization—accomplishments and challenges: Journal of Canadian Petroleum Techology, **39**(7), 25–29.

Kimeldorf, G. S., and Wahba, G., 1970, A correspondence bteween Bayesian estimation on stochastic processes and smoothing by splines: Annals of Mathematical Statistics, **41**(2), 495–502.

Lamy, Ph., Swaby, P. A, Rowbotham, P. S., and Marion, D., 1998a, Faulted reservoir models from high–resolution geostatistical inversion: EAGE/SPE International Symposium on Petroleum Geostatistics, Extended Abstract.

Lamy, Ph., Swaby, P. A, Rowbotham, P. S., Dubrule, O., and Haas. A., 1998b, From seismic to reservoir properties using geostatistical inversion: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 49147.

Lamy, Ph., Swaby, P. A, Rowbotham, P. S., Dubrule, O., and Haas, A., 1999, From seismic to reservoir properties with geostatistical inversion: SPE Reservoir Evaluation and Engineering, **2**(4), 334–340.

Lia, O., Omre, H., Tjelmeland, H., Holden, L., and Egeland, T., 1997, Uncertainties in reservoir production forecasts: AAPG Bull., **81**(5), 775–802.

Lo, T. W., and Bashore, W. M., 1999, seismic constrained facies modeling using stochastic seismic inversion and indicator simulation, a North Sea Example: 69th Annual Meeting, Soc. Expl. Geophys., Expanded Abstracts, 923–926.

Macdonald, A. C., Berg, J. I., Skare, O., Holden, L., 1995, Constraining a stochastic model of channel geometries using seismic data: 57th EAGE Conference, Eur. Assn. Geosci. Eng., Extended Abstracts, F-052.

Mandelbrot, B., 1982, The fractal geometry of nature: Freeman and Co.

Marion, D., Insalaco, E., Rowbotham, P., Lamy, Ph., and Michel, B., 2000, Constraining 3D static models to seismic and sedimentological data: a further step towards reduction of uncertainties: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 65132.

Massonnat, G. J., 2000, Can we sample the complete geological uncertainty space in reservoir modeling uncertainty estimates?: SPE Journal, 5(1), 46–59.

Matheron, G., 1962, Traité de Géostatistique Appliquée: Technip.

Matheron, G., 1970, The theory of regionalized variables and its applications: Centre de Géostatistique de l'Ecole des Mines de Paris, Les Cahiers du Centre de Morphologie Mathématique, Fasc. 5.

Matheron, G., 1973, The intrinsic random functions and their applications: Advances in Applied Probability, 5, 439–468.

Matheron, G., 1981a, Splines and kriging, their formal equivalence: Syracuse University Geol. Contrib., 8.

Matheron, G., 1981b, Remarques sur le Krigeage et son Dual: Centre de Géostatistique de l'Ecole des Mines de Paris, Internal Report N–695.

Matheron, G., 1987, A simple answer to an elementary question, letter to the editor: Mathematical Geology, 19(5), 455–457.

Mathieu, G., and Nutt, L., 1985, A geostatistical approach to velocity mapping: Schlumberger External Publication.

Maus, S., 1999, Variogram analysis of magnetic and gravity data: Geophysics, 64(3), 776–784.

Maus S., Sengpiel, K. P., Röttger, B., Siemon, B. and Tordiffe, E. A. W, 1999, Variogram analysis of helicopter magnetic data to identify paleochannels of the Omaruru River, Namibia: Geophysics, 64(3), 785–794.

Mitas L., and Mitasova, H., 1988, General variational approach to the interpolation problem: Comput. Math. Applic., 16(12), 983–992.

Mitasova, H., and Mitas, L., 1993, Interpolation by regularised spline with tension: theory and implementation: Mathematical Geology, 25(6), 641–655.

Mosegaard, K., and Tarantola, A., 1995, Monte Carlo sampling of solutions to inverse problems: J. Geophys. Res., 100(B7), 12,431–12,447.

Mundim, E. C., Johann, P., and Remacre, A. Z., 1999, Factorial kriging analysis: geostatistical filtering applied to reservoir characterization: The Leading Edge, July, 787–788.

Oliver, D. S., 1998, Calculation of the inverse of the covariance: Mathematical Geology, 30(7), 911–933.

Oliver, D.S., 1995, Moving averages for Gaussian simulation in two and three dimensions: Mathematical Geology, 27(8), 939–960.

Omre, H., 1987, Bayesian kriging, merging observations and qualified guesses in kriging: Mathematical Geology, 19(1), 25–39.

Omre, H., and Halvorsen, B., 1989, The Bayesian bridge between simple and universal kriging: Mathematical Geology, 21(7), 767–786.

Omre, H., and Tjelmeland, H., 1997, Petroleum geostatistics, *in* Baafi E. Y., and Schofield, N.

A., Eds, Geostatistics Wollongong '96, Kluwer Academic Publishers, Netherlands, **I**, 41–52.

Ovreberg, O., Damsleth, E., and Haldorsen, H. H., 1992, Putting error bars on reservoir engineering forecasts: Journal of Petroleum Technology, **44**(6), 732–738.

Parks, K. P., Bentley, L. R., and Crowe, A. S., 2000, Capturing geological realism in stochastic simulations of rock systems with Markov statistics and simulated annealing: Journal of Sedimentary Research, **70**(4), 803–813.

Pawar, R. J., Edwards, E. B., and Whitney, E. M., 2001, Geostatistical characterization of the Carpinteria field, California: Journal of Petroleum Science and Engineering, **31**, 175–192.

Perez, G., and Chopra, A .K., 1991, Evaluation of fractal models to describe reservoir heterogeneity and performance: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 22694, p. 387–398.

Perez, G., Chopra, A. K., and Severson, C. D., 1997, Integrated geostatistics for modeling fluid contacts and shales in Prudhoe Bay: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE Formation Evaluation, Dec., 213–219.

Piazza, J. L., Légeron, S., and Sandjivy, L., 1997, Use of geostatistics to improve seismic velocities: case studies: 67th Annual Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1293–1296.

Ravenne, C., and Beucher, H., 1988, Recent developments in description of sedimentary bodies in a fluvio–deltaic reservoir and their 3D conditional simulations: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 18310.

Renard, D., 1990, Bluepack 3D and its use in the petroleum industry: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 20352.

Ripley, B. D., 1981, Spatial statistics: John Wiley & Sons, Inc.

Ritzi, R. W., 2000, Behavior of Indicator variograms and transition probabilities in relation to the variance in lengths of hydrofacies: Water Resources Research, **36**(11), 3375–3381.

Rivoirard, J., 2002, On the structural link between variables in kriging with an external drift: Mathematical Geology, **34**(7), 797–808.

Rose, P., 2001, Risk analysis and management of petroleum exploration ventures: Am. Assn. Petr. Geol., AAPG Methods in Exploration Series, No. 12.

Rowbotham, P., Marion, D., Lamy, Ph., Swaby, P., and Rabary, G., 2000, Detailed reservoir characterization of the Elgin field using geostatistical inversion: 62nd EAGE Conference, Eur. Assn. Geosci. Eng., Extended Abstracts, A-17.

Rudkiewicz, J. L., Guérillot, D., and Galli, A., 1990, An integrated software for stochastic modeling of reservoir lithology and property with an example from the Yorkshire Middle Jurassic, *in* Buller et al., eds., North Sea oil and gas reservoirs, II: Graham & Trotman Ltd., 399–406.

Sams, M.S., Atkins, D., Said, P. T., Parwito, E., and van Riel, P., 1999, Stochastic inversion for high resolution reservoir characterization in the Central Sumatra Basin: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 57260.

Samson, Ph., Dubrule O., and Euler, N., 1996, Quantifying the impact of structural uncertainties on gross–rock volume estimates: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 35535.

Sandjivy, L., 1987, Analyse Krigeante des données de Prospection Géochimique: Docteur–Ingénieur Thesis, Centre de Géostatistique Ecole Nationale Supérieure des Mines de Paris.

Sandsdalen, C., Barbieri, M., Tyler, K., and Aasen, J. O., 1996, Applied uncertainty analysis using stochastic modeling: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 35533.

Scales, J. A. and Tenorio, L., 2001, Prior information and uncertainty in inverse problems: Geophysics, **66**(2), 389–397.

Sen, M. K., and Stoffa, P. L., 1996, Bayesian inference, Gibbs' Sampler and uncertainty estimation in geophysical inversion: Geophysical Prospecting, **4**, 313–350.

Shrestha, R. K., and Boeckmann, M., 2002, High–resolution 3D impedance data for reservoir modeling: 64th EAGE Conference and Exhibition, Eur. Assn. Geosci. Eng., Extended Abstracts, H-37.

Skare, O., Skorstad, A., Hauge, R., and Holden, L., 1997, Conditioning a fluvial model on seismic data, , *in* Baafi E.Y., and Schofield N.A., Eds., Geostatistics Wollongong '96: Kluwer Academic Publ.

SPE/WPC, 1997, Reserves definitions approved: Journal of Petroleum Technology, **49**(5), 527–528.

Spector, A., and Grant, F. S., 1970, Statistical models for interpreting aeromagnetic data: Geophysics, **35**, 293–302.

Spencer, J. A., and Morgan, D. T. K, 1998, Application of forecasting and uncertainty methods to production: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 49092.

Strebelle, S., and Payrazyan, K., 2002, Modeling of a deepwater turbidite reservoir conditional to seismic data using multiple–point geostatistics: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 77425.

Szeliski, R., and Terzopoulos, D., 1989, From splines to fractals: Computer Graphics, **23**(3), 51–60.

Szerbiak, R. B., McMechan, G. A., Corbeanu, R., Forster, C., and Snelgrove, S. H., 2001, 3D characterization of a clastic reservoir analog: from 3D GPR data to a 3D fluid permeability model: Geophysics, **66**(4), 1026–1037.

Tarantola, A., 1987, Inverse problem theory: methods for data fitting and model parameter estimation: Elsevier Science Publ. Co., Inc.

Taylor, S. R., 1996, 3D modeling to optimise production at the successive stages of field life: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 35501.

Thore, P., Shtuka, A., Lecour, M., Ait–Ettajer, T., and Cognot, R., 2002, Structural uncertainties: determination, management, and applications: Geophysics, **67**(3), 840–852.

Tinker, S. W., 1996, Building the 3D jigsaw puzzle: applications of sequence stratigraphy to 3D reservoir characterization, Permian Basin: AAPG Bull., **80**(4), 460–485.

Tjelmeland, H., and Omre, H., 1997, A Complex sand–shale facies model conditioned on observations from wells, seismics and production, *in* Baafi, E. Y., and Schofield, N. A., Eds, Geostatistics Wollongong '96: Kluwer Academic Publ., **I**, p. 634–643.

Tran, T. T., Deutsch, C. V., and Xie, Y., 2001, Direct geostatistical simulation with multiscale well, seismic and production data: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 71323.

Tyler, K., Sandsdalen, L., Maeland, L., Aasen, J. O., Siring, E., and Barbieri, M., 1996, Integrated stochastic modeling in reservoir evaluation to project evaluation and risk assessment: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 36706.

Tyler, N., Barton, M. D., and Finley, R. J., 1991, Outcrop characterization of flow unit and seal properties and geometries, Ferron Sandstone, Utah: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 22670.

Ulrych, J., Sacchi, M. D., and Woodbury, A., 2001, A Bayes tour of inversion: a tutorial: Geophysics, **66**(1), 55–69.

Vincent, G., Corre, B., and Thore, P., 1998, Managing structural uncertainty in a mature field for optimal well placement: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 48953.

Wadsworth, G. P., Robinson, E. A., Bryan, J. G., and Hurley, P. M., 1953, Detection of reflections on seismic records by linear operators: Geophysics, **18**(3), 539–586.

Weber, K. J., and van Geuns, L. C., 1990, Framework for constructing clastic reservoir simulation models: Journal of Petroleum Technology, **42**(10), 1248–1297.

Wessel, P., and Bercovici, D., 1998, Interpolation with splines in tension: A Green's Function approach: Mathematical Geology, **30**(1), 77–93.

Xu, W., Tran, T. T., Srivastava, R. M., and Journel, A. G., 1992, Integrating seismic data in reservoir modeling: the collocated cokriging alternative: Soc. Petr. Eng., Am. Inst. Min., Metall. Petr. Eng., SPE 24742.

Yao, T, 1998, Automatic covariance modeling and conditional spectral simulation with fast Fourier transform: PhD Dissertation, Stanford University.

Yarus, J. M., Yang, K., Sriisraporn, S., Chuemthaisong, N., and Sangwongwanich, K., 2000, Integrating 3D seismic and geostatistics: building a 3D model of a Tertiary deltaic and shallow marine deposit, Malay Basin, offshore Gulf of Thailand: Presented at the 2000 OTC Conference, OTC 11961.